

PROSPECTIVE D'ÉVOLUTION DES CONSOMMATIONS DES CENTRES DE DONNÉES EN FRANCE DE 2024 À 2060




EXPERTISES

RAPPORT FINAL

Janvier 2026

REMERCIEMENTS

Nous tenons à remercier chaleureusement les membres du COPIL de cette étude pour leurs contacts, les mises en relation qu'ils ont pu faciliter, leurs relectures, leurs avis qui nous ont été d'une grande utilité. Nous remercions aussi toutes les personnes ressources que nous avons interviewées, nos relecteurs et relectrices, et l'ADEME pour sa confiance dans la réalisation de cette étude.

CITATION DE CE RAPPORT

Groupement CLIK, Lorraine DE MONTENAY, Benoit PETIT, Cécile DIGUET, Eric FOURBOUL. 2026/01. Prospective d'évolution des consommations des centres de données en France de 2024 à 2060. 215 pages.

Cet ouvrage est disponible en ligne <https://librairie.ademe.fr/>

Toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite selon le Code de la propriété intellectuelle (art. L 122-4) et constitue une contrefaçon réprimée par le Code pénal. Seules sont autorisées (art. 122-5) les copies ou reproductions strictement réservées à l'usage privé de copiste et non destinées à une utilisation collective, ainsi que les analyses et courtes citations justifiées par le caractère critique, pédagogique ou d'information de l'œuvre à laquelle elles sont incorporées, sous réserve, toutefois, du respect des dispositions des articles L 122-10 à L 122-12 du même Code, relatives à la reproduction par reprographie.

Ce document est diffusé par l'ADEME

ADEME

20, avenue du Grésillé

BP 90 406 | 49004 Angers Cedex 01

Numéro de contrat : 2024MA000251

Étude réalisée pour le compte de l'ADEME par : groupement CLIK (Omnicité - Lorraine de Montenay / Hubblo / Studio Dégel / Sinok)

Coordination technique - ADEME : LAFITTE Bruno

Direction/Service : DVTD / Service Bâtiments

Photo de couverture : Jardins partagés au bord du canal St Denis (93) et centre de données d'Equinix. Crédits : Cécile Diguët

SOMMAIRE

SOMMAIRE	3
Résumé Exécutif	6
Glossaire.....	16
1. Contexte des évolutions actuelles des usages et consommations des centres de données	21
1.1. Contexte	21
1.2. Évolutions du marché	27
1.2.1. Les principaux acteurs de la colocation et du IAAS	30
1.2.2. Les fournisseurs de hardware	31
1.3. Prospective sur les usages	31
1.3.1. La "migration" vers le Cloud : une tendance de fond	32
1.3.2. Une migration aux airs d'accumulation, plus que de transition.....	35
1.3.3. Évolutions des usages du grand public	36
1.4. Évolutions des technologies et des pratiques du secteur	46
1.4.1. Évolution des techniques et montages liés à l'énergie.....	46
1.4.2. Évolutions matériel IT	54
1.5. Évolutions environnementales et sociales	84
1.5.1. Energie.....	84
1.5.2. Eau	105
2. Analyse qualitative des études prospectives sur l'évolution des consommations des centres de données	130
2.1. Études institutionnelles	130
2.1.1. "Transition(s) 2050. Choisir maintenant. Agir pour le climat"	130
2.1.2. "Évaluation de l'impact environnemental du numérique en France et analyse prospective"	131
2.1.3. Études RTE.....	135
2.1.4. Rapport de l'AIE "Electricity 2024, Analysis and forecast to 2026"	137
2.1.5. Les rapports d'orientation stratégique du CIGREF de 2020, 2022 et 2023	137
2.1.6. Energy-efficient cloud computing technologies and policies for an eco-friendly cloud market	138
2.2. Les principaux modèles utilisés	138
2.2.1. Modèle Masanet.....	138
2.2.2. Modèle Andrae	139
2.2.3. Modèle Belkhir	139

2.2.4. Modèle Malmodyn	139
2.2.5. Modèle Borderstep	140
2.2.6. Le modèle Total Energy Model (TEM) de l'EDNA.....	141
2.3. Perspectives climatiques	142
2.4. Études prospectives sur la ressource en eau.....	143
3. Méthodologie du modèle.....	144
3.1. Définition de l'objectif et du périmètre de l'étude	144
3.2. Indicateurs de consommation considérés	145
3.2.1. Consommation d'électricité.....	145
3.2.2. Émissions de Gaz à Effet de Serre	145
3.2.3. Consommation foncière.....	145
3.3. Principes généraux du modèle	146
3.3.1. Évolution usage vs efficacité.....	146
3.3.2. Choix de modélisation	148
3.3.3. Architecture globale.....	149
3.3.4. Base de données 2024.....	150
3.3.5. Modélisation bottom-up.....	157
3.4. Approche méthodologique détaillée.....	160
3.4.1. Modélisation de la situation initiale.....	160
3.4.2. Court, moyen, long terme : définition de séquences d'évolution	165
3.4.3. Modélisation du potentiel de récupération de chaleur fatale.....	167
3.5. Modèle 2035-2060	168
4. Scénarios : approche de révision des scénarios Transition 2050 et scénario tendanciel	168
4.1. Approche globale	168
4.2. Scénario tendanciel	169
4.3. Scénario Génération Frugale (S1).....	173
4.4. Scénario Coopération Territoriale (S2)	176
4.5. Scénario Technologies Vertes (S3).....	179
4.6. Scénario Pari Réparateur (S4).....	182
5. Hypothèses	186
5.1. Évolution de l'efficacité des serveurs.....	186
5.2. Usages de l'IA de "nouvelle génération"	187
5.2.1. Inférences vs Apprentissage	187
5.2.2. Les tendances initiales.....	189
5.3. Blockchain	189
6. Résultats.....	192

6.1. Scénario tendanciel	192
6.1.1. Période 2024-3025	192
6.1.2. Répartitions par type de centres de données en pourcentage des consommations.....	193
6.1.3. Modélisation post-2035	193
6.2. Scénario Génération Frugale	194
6.2.1. Période 2024-3025.....	194
6.2.2. Répartitions par type de centres de données en pourcentage des consommations.....	195
6.2.3. Modélisation post-2035	196
6.3. Scénario Coopérations Territoriales	197
6.3.1. Période 2024-3025.....	197
6.3.2. Répartitions par type de centres de données en pourcentage des consommations.....	198
6.3.3. Modélisation post-2035	198
6.4. Scénario Technologies Vertes	199
6.4.1. Période 2024-3025.....	199
6.4.2. Répartitions par type de centres de données en pourcentage des consommations.....	201
6.4.3. Modélisation post-2035	201
6.5. Scénario Pari Réparateur	202
6.5.1. Période 2024-3025.....	202
6.5.2. Répartitions par type de centres de données en pourcentage des consommations.....	204
6.5.3. Modélisation post-2035	204
6.6. Comparaison des scénarios	206
6.6.1. Comparaison des consommations électriques 2024-2035.....	208
6.6.2. Comparaison des émissions de GES 2024-2035.....	210
6.6.3. Comparaison des consommations électriques jusqu'à 2060	211
6.6.4. Comparaison des émissions de GES jusqu'à 2060	212
6.6.5. Modélisation du potentiel de récupération de chaleur fatale.....	213
7. Limites de l'étude.....	216
7.1. Limites liées au périmètre de l'étude	216
7.2. Limites liées à la maille géographique définie	216
7.3. Limites liées à la disponibilité des données	217
7.4. Limites liées aux indicateurs	219
7.5. Limites liées à l'exercice prospectif.....	219
Conclusion.....	221

Résumé Exécutif

Des usages croissants aux infrastructures stratégiques, le numérique doit être intégré dans la transition écologique

Les usages numériques actuels connaissent une croissance forte, portée par l'essor de l'intelligence artificielle, du cloud, de l'e-commerce, de l'IoT, du divertissement, et des services publics dématérialisés. Le développement de cette demande est soutenu par la construction de centres de données de plus en plus nombreux et de taille croissante. Au-delà de leurs capacités de stockage des données, ils se distinguent par leurs capacités de calcul, leur très faible latence et leur disponibilité permanente. Ces infrastructures sont devenues stratégiques, comparables à des réseaux routiers ou électriques : en témoigne la récente adoption le 17 juin 2025 de l'article 15 du projet de loi de simplification de la vie économique.¹ L'implantation géographique de nouveaux centres de données est aujourd'hui l'objet de nombreux enjeux économiques - compétitivité, investissements considérables, promesse de nouveaux pôles technologiques - et stratégiques - souveraineté numérique, leadership. Toutefois, la transition numérique doit être intégrée dans la transition écologique et répondre aux défis environnementaux d'aujourd'hui et de demain.

En effet, les centres de données sont parmi les infrastructures les plus concentrées spatialement et intensives énergétiquement² : ils nécessitent non seulement de l'électricité pour alimenter les serveurs et autres équipements, mais aussi pour assurer leur refroidissement permanent. À l'échelle mondiale, leur consommation électrique représentait déjà 415 TWh en 2024³, et les projections convergent vers une croissance importante de cette consommation dans les prochaines années.⁴ Dans le contexte de la transition énergétique et de tensions à anticiper sur les réseaux électriques, leur impact doit être mesuré et maîtrisé à l'aune de la compatibilité du développement de ces infrastructures avec les objectifs climatiques et énergétiques français et européens.

Enfin, disposer de données fiables sur la consommation électrique des centres de données contribue à réduire leurs externalités négatives, dont les émissions de gaz à effet de serre, leur éventuelle contribution à la pression sur les ressources en eau, la concurrence avec les usages industriels et résidentiels de l'électricité, ou encore l'artificialisation des sols. Pour les pouvoirs publics, cela constitue un levier indispensable afin de stimuler les efforts d'efficacité énergétique du secteur, d'imposer des normes adaptées et de promouvoir des innovations plus durables.

Les principaux objectifs de cette étude

Ainsi, la présente étude de l'ADEME sur l'évolution des consommations des centres de données en France, vise 3 objectifs :

1. Dresser **l'état des lieux de la consommation électrique actuelle des centres de données en France** ;
2. Proposer un **modèle prospectif détaillé** permettant de modéliser des scénarios d'évolution des consommations des centres de données dans le temps, qui tienne compte des évolutions des

¹ qui permet à certains projets de centres de données d'être qualifiés de projets d'intérêt national majeur (PINM), leur permettant, dans la limite de 10 000 hectares au niveau national, de ne pas être comptabilisé dans la consommation d'espaces naturels, agricoles et forestiers https://www.assemblee-nationale.fr/dyn/17/textes/17t0144_texte-adopte-seance#Article_15 ; NB : à la date d'écriture de ce rapport, la loi de simplification de la vie économique n'est pas encore promulguée, une commission mixte paritaire devant se réunir concernant les dispositions restant en discussion sur le projet de loi.

² voir à ce titre la section [Le sujet de la concentration spatiale](#) de la présente étude

³ IEA (2025), Energy and AI, IEA, Paris <https://www.iea.org/reports/energy-and-ai>

⁴ voir à ce titre notre analyse au chapitre [Analyse qualitative des études prospectives sur l'évolution des consommations des centres de données](#) de la présente étude

usages des Français et distingue la consommation des centres de données présents sur le territoire français des centres de données présents ailleurs dans le monde mais répondant aux usages des Français ;

3. Modéliser et analyser **5 scénarios prospectifs jusqu'à 2060** : un scénario tendanciel et 4 scénarios envisageant les 4 chemins possibles de transition écologique déjà travaillés dans le cadre du programme Transition(s) 2050 de l'ADEME - Génération Frugale, Coopérations Territoriales, Technologies Vertes, Pari Réparateur.

A noter que cette étude propre aux data centers n'évalue pas les besoins en ressources ni l'électricité nécessaire à la fabrication des serveurs.

Si rien n'est fait, la consommation électrique des centres de données pourrait être x3,7 France d'ici 2035 et x4,4 si on tient compte de leur consommation importée pour répondre aux usages des Français.

Les principaux jalons de cette étude

Cette étude a rassemblé un large Comité de Pilotage composé de représentants du ministère de la transition écologique, du ministère de l'économie et des finances, de l'ARCEP, de RTE, d'Enedis, de France Data Center et de différents services de l'ADEME.

Les résultats présentés dans cette étude sont le fruit du travail de collecte, d'analyse, de synthèse, de modélisation, de scénarisation et de concertation réalisé par le consortium CLIK pour l'ADEME :

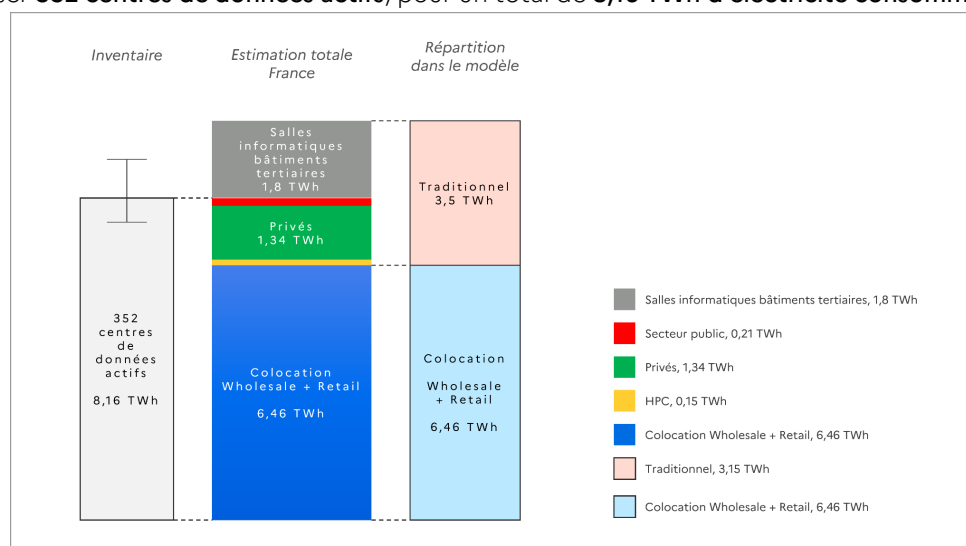
1. Une **collecte** à la fois qualitative et quantitative : qualitative afin de rendre compte de l'état de l'art des connaissances actuelles sur les consommations des centres de données et leurs évolutions en cours, quantitative afin de constituer une base de données des centres de données actuellement actifs et en projet sur le territoire français et estimer leur consommation électrique annuelle sur la base des données disponibles.
2. Une **analyse et une synthèse** de l'ensemble foisonnant et complexe de l'ensemble des constituants de la consommation des centres de données : les évolutions des usages des Français et l'augmentation de l'offre de services numériques (IA générative, blockchain, ...), les évolutions en matières de technologies de l'information et d'architecture et de technologies de refroidissement (développement de technologies de refroidissement adaptées aux GPU pour l'IA, récupération de chaleur fatale, ...), les enjeux liés aux consommations de ressources (principalement énergies, eau, foncier), ainsi que les évolutions réglementaires, et les signaux forts ou faibles comme potentiels marqueurs des tendances à venir.
3. Un travail de **modélisation** en réalisant un modèle Excel paramétré pour permettre de scénariser année après année les évolutions des consommations électriques des centres de données, au regard de l'évolution des usages des Français, en distinguant différentes typologies de centres de données et de nombreuses variables.
4. Un travail de **scénarisation** pour imaginer, distinguer et narrer via une vidéo, les 5 voies possibles d'évolution des usages du numérique et des évolutions des technologies des centres de données, modéliser et estimer leurs impacts en termes de consommation électrique des centres de données, en cohérence avec les travaux réalisés par l'ADEME dans le cadre de Transition(s) 2050 et avec les enseignements tirés de la revue documentaire.

Les principaux enseignements de cette étude

Les principaux enseignements tirés de ce travail sont les suivants :

1. L'accroissement de l'offre de services numériques, combiné à une adoption massive des usages et un encadrement faible des services numériques concernant les enjeux environnementaux conduit à une montée en puissance inéluctable et rapide de la consommation des centres de données en France et dans le monde, menaçant la tenue des engagements environnementaux.
2. Parmi les tendances actuelles, les principaux signaux fort d'évolution du numérique qui influencent la consommation des centres de données sont :
 - a. L'intelligence artificielle, en particulier les IA génératives et agentielles⁵, dont les applications et les usages se généralisent.
 - b. La blockchain, chaînon des consommations souvent invisibilisé du fait de sa nature décentralisée : son développement reste encore limité sur le sol français mais la progression de ses usages pourrait s'accroître plus fortement dans les années à venir.
 - c. Le dialogue machine à machine (M2M) : après l'atteinte de deux formes de plafond (1) du taux de pénétration du numérique dans la population et (2) du temps par jour en interaction numérique possible pour un individu, le numérique continue de capter, générer et analyser de la données en s'affranchissant de la barrière de l'individu (IoT, M2M, ...).
 - d. La consommation électrique des centres de données en France pourrait aussi s'accroître du fait du rôle stratégique que joue aujourd'hui ce type d'infrastructure, au carrefour des enjeux de la souveraineté et de la défense dans des conflits hybrides de plus en plus électroniques.

Pour l'année 2024, les résultats de notre travail d'inventaire des centres de données en France a permis de recenser **352 centres de données actifs**, pour un total de **8,16 TWh d'électricité consommés**.⁶



Représentation schématisée de la répartition des centres de données par typologie dans l'établissement de la baseline du modèle : à gauche, l'inventaire réalisé sur la base de données primaires ; au centre, la répartition sur la base de l'inventaire et en complément l'estimation des baies de bureau selon le modèle Masanet ; à droite, la répartition conservée dans le modèle CLIK.

⁵ L'IA agentielle correspond aux agents virtuels alimentés par l'intelligence artificielle ; c'est une sous-catégorie de l'IA générative.

⁶ Ce travail d'inventaire pour établir la baseline nous a permis d'identifier avec un bon degré de fiabilité 352 centres de données actuellement en exploitation, ainsi que d'autres en projet. Une marge d'erreur subsiste toutefois concernant les centres de données qui n'ont pas pu être répertoriés ainsi que d'éventuelles lacunes sur les informations disponibles concernant leur puissance et leur surface.

Dans ce décompte les centres de données de colocation (Wholesale et Retail)⁷ pèsent pour 6,46 TWh, les centres de données privés pour 1,34 TWh, les centres de données du secteur public pour 0,21 TWh et les centres de données spécialisés dans le calcul haute performance (HPC) pour 0,15 TWh. Le périmètre concerné par cette collecte exclut les salles informatiques présentes dans les bâtiments tertiaires. Cette partie est prise en compte en s'appuyant sur la méthode utilisée dans l'étude ADEME/Arcep parue en 2022, prenant en compte les besoins en surface d'hébergement IT des entreprises. Ce faisant, en appliquant les hypothèses du modèle Masanet, c'est-à-dire, une baisse annuelle de 5% pour le volume de workloads opérés par les centres de données "traditionnels", nous prenons en compte une consommation électrique annuelle de 3,5 TWh. Cela inclut des centres de données privés, des centres de données publics et des baies informatiques dans les bâtiments tertiaires. Le cumul de consommation d'électricité sur l'ensemble du périmètre, prenant en compte les centres de données de colocation issus de la collecte et l'extrapolation surfacique issue du rapport ADEME/Arcep, donne une consommation de 10 TWh en 2024, qui sert de point de départ au modèle.⁸

Les 5 scénarios modélisés

Basée sur les données de notre inventaire supra citées, notre modélisation prospective nous a permis de distinguer les évolutions pour 5 scénarios, sur 2 périodes clés (pré-2035 et post-2035). En résumé :

1. **Tendanciel** : ce scénario rend compte de l'évolution potentielle des consommations électriques des centres de données pour les usages des Français, dans le cas où les tendances actuelles perdureraient : forte croissance liée à l'adoption massive de nouveaux usages et la diversification de l'offre de services numériques, principalement liée à l'IA, la blockchain, et l'évolution de la typologie des centres de données vers de très grands centres appelés hyperscale. Dans ce scénario, la croissance des centres de données sur le territoire français est soutenue, mais la consommation des workloads des Français se développe proportionnellement plus rapidement encore.
2. **Génération Frugale** : scénario le plus engagé dans la sobriété parmi les 4 scénarios Transition(s) 2050, Génération Frugale envisage une dénumérisation de certains usages actuels du numérique, en particulier les usages les plus gourmands et ceux qui contraignent les usagers à l'adoption du tout numérique, et en proposant des services collectifs alternatifs non numériques ou faiblement numériques ; ce scénario envisage aussi la mise en place de politiques environnementales contraignantes comme un moratoire sur la construction de nouveaux centres de données en France. Les politiques publiques mises en place dans ce scénario visent une réorientation complète de nos modes de vie allant bien au-delà d'une politique qui serait purement incitative, en la structurant. L'écoconception de rupture est recherchée et valorisée. Dans ce scénario, les usages de l'IA continuent de se développer, mais moins rapidement que dans les autres scénarios. Dans Génération Frugale, la réorientation de nos modes de vie et l'écoconception permettent de réduire de façon très significative la consommation de workloads des Français, au point où, malgré le moratoire sur la construction de nouveaux centres de données, la dépendance de la France aux workloads calculés hors de son territoire se réduit elle aussi très fortement.

⁷ voir section [Les typologies de centres de données retenues pour cette étude](#)

⁸ En complément des typologies Traditionnel, Colocation Wholesale et Retail dans le modèle prospectif, les typologies Hyperscale, IA et Blockchain ont été ajoutées. Les hyperscales, très développés aux Etats-Unis, ne sont pas en exploitation en France au moment de la réalisation de cette étude, bien que certains soient en projet ; dans le modèle prospectif, les hyperscales se répartissent entre import et traitement en France en proportion variée selon les scénarios. La blockchain restant anecdotique en France et son usage amené à rester dépendant d'instances de calcul traitées massivement à l'étranger, la blockchain est traitée en import dans le modèle. Concernant l'IA dans le modèle, nous avons considéré séparément les nouveaux usages de l'IA générative, afin de pouvoir mettre leur évolution en exergue dans l'évolution prospective, sans qu'il nous soit toutefois possible d'établir cette distinction pour les usages passés dans la baseline.

3. **Coopérations Territoriales** : dans ce scénario, les projets de centres de données sont développés en concertation avec les territoires et encadrés par des normes qui limitent leur implantation, sans toutefois aller jusqu'à la mise en place d'un moratoire ; comme dans le scénario précédent, les centres de données sont incités à adresser avant tout les usages considérés comme prioritaires : bénéfiques pour la santé, l'environnement et la société dans son ensemble. Les pouvoirs publics facilitent l'adoption de pratiques plus sobres, tout en laissant à chacune et chacun la liberté de rester connecté. Le scénario Coopérations Territoriales est celui qui favorise le plus les synergies entre l'implantation des centres de données et les enjeux d'un territoire : récupération de la chaleur fatale, éviter l'accroissement des tensions sur les ressources telles que l'eau ou le foncier.
4. **Technologies Vertes** : dans ce scénario, les centres de données sont un des points d'appui stratégiques majeurs de la transition, en favorisant l'innovation technologique pour réduire les émissions d'autres secteurs d'activité et optimiser les consommations du numérique. Dans ce scénario, les politiques publiques incitent à l'implantation de centres de données sur le territoire français, à la fois pour des questions de souveraineté numérique, mais aussi en valorisant le mix électrique français, peu carboné. Dans Technologies Vertes, sont particulièrement favorisés les projets d'implantation de centres de données ayant vocation à héberger des projets technologiques porteurs de solutions pour la transition écologique. Dans cette logique, ce scénario voit un développement relativement important de la blockchain et de l'IA.
5. **Pari Réparateur** : dans ce scénario, le pari est fait que les innovations technologiques nous permettraient de conserver notre mode de vie à travers une compensation de ses impacts ; ce scénario envisage le développement des usages des centres de données avec une réglementation allégée et une organisation des territoires pour satisfaire la demande plutôt que l'inverse ; les centres de données entrent en compétition avec les autres industries pour l'usage de l'électricité dans le cadre de la transition. Là où le scénario Technologies Vertes envisage des politiques publiques incitant à l'implantation de centres de données sur le territoire français dans une volonté de développer sa souveraineté, Pari Réparateur envisage des politiques publiques similaires mais à l'échelle européenne : est pris pour hypothèse dans ce scénario une dilution relative de l'implantation des centres de données, avec un développement moindre sur le sol français dans ce scénario en comparaison du scénario précédent. Ce scénario explore aussi l'hypothèse d'un développement très important de la blockchain dans les usages des Français.

Une évaluation prospective scindée en deux périodes : pré 2035 et post-2035 jusqu'à 2060

La distinction entre deux périodes clés (pré-2035 et post-2035) est liée à l'évolution technologique extrêmement rapide du secteur numérique : le degré d'incertitude est ainsi plus fort dans la période post-2035, plus lointaine et plus sujette à l'émergence de technologies de rupture, mais aussi aux aléas climatiques et tensions sur les ressources, que pour la période pré-2035 pour laquelle il est plus simple d'anticiper les éventuels sauts technologiques et modéliser la maturation des technologies actuelles. Notre modèle distingue donc les deux périodes avec une méthode d'évaluation différente.⁹

Ainsi, les résultats de notre modélisation pour la période 2024-2035, montrent que **si aucune politique de transition écologique n'est mise en place, la consommation électrique des centres de données risque d'être multipliée par 3,7 rien que sur le territoire français, et multipliée par 4,4 si on tient compte de la consommation électrique des centres de données à l'étranger pour répondre aux usages des Français** (scénario Tendanciel) :

⁹ voir à ce titre le chapitre [Approche méthodologique détaillée](#)

Long terme - Evolution de la consommation électrique pour l'usage français (TWh / an)

	2024	2030	2035	2040	2045	2050	2055	2060
Tendanciel - en France	9,91	20,75	36,73	48,04	51,61	54,73	59,03	63,58
Tendanciel - import	13,79	35,61	68,53	127,66	168,68	194,26	213,30	234,97
Tendanciel - Total	23,71	56,36	105,25	175,70	220,29	248,99	272,33	298,55
Génération Frugale - en France	9,91	14,87	11,58	6,95	5,58	5,77	4,99	4,59
Génération Frugale - import	13,79	15,26	10,54	5,91	2,69	0,57	0,48	0,42
Génération Frugale - Total	23,71	30,13	22,12	12,86	8,27	6,34	5,47	5,01
Coopérations Territoriales - en France	9,91	17,10	17,54	11,69	8,58	6,81	5,88	5,46
Coopérations Territoriales - import	13,79	17,04	11,92	10,14	9,32	8,90	8,86	9,10
Coopérations Territoriales - Total	23,71	34,14	29,46	21,83	17,89	15,72	14,74	14,56
Technologies Vertes - en France	9,91	19,34	32,14	48,85	55,50	64,19	69,14	74,89
Technologies Vertes - import	13,79	24,01	29,53	31,77	31,67	30,39	33,85	37,70
Technologies Vertes - Total	23,71	43,34	61,67	80,62	87,17	94,58	102,99	112,59
Pari Réparateur - en France	9,91	19,28	33,04	50,65	56,47	61,31	64,07	67,46
Pari Réparateur - import	13,79	37,00	77,31	145,79	190,12	230,57	254,40	280,95
Pari Réparateur - Total	23,71	56,28	110,35	196,44	246,59	291,88	318,48	348,41

Sur une échelle temporelle plus longue, en 2050, le scénario Tendanciel pourrait voir la consommation électrique des centres de données multipliée par 5,5 rien que sur le territoire français, et multipliée par 10,5 si on tient compte de la consommation électrique des centres de données à l'étranger pour répondre aux usages des Français.

NB : A noter que cette prospective de très long terme est sujette à précautions, les prévisions sectorielles pour les centres de données s'arrêtant généralement à une perspective de 5 à 10 ans.

En fonction des scénarios, l'évolution des consommations électriques des centres de données varie considérablement, rendant compte de l'efficacité substantielle ou relative des différentes mesures adoptées.

Ainsi, si le scénario **Génération Frugale** était adopté, la consommation électrique des centres de données continuerait de croître pendant quelques années, le temps que les politiques publiques de rupture envisagées prennent effet, puis après avoir marqué un ralentissement, ces consommations diminueraient progressivement. Ainsi on assisterait dans ce scénario à **une augmentation limitée à 17% seulement de la consommation des centres de données en France en 2035 par rapport à leur consommation en 2024, et**

de -7% en tenant compte des consommations importées. En 2050, la réduction de la consommation des centres de données en France serait de -42% par rapport à leur consommation en 2025, et de -73% en tenant compte des consommations importées.

Dans l'intervalle, si le scénario Coopérations Territorial était adopté, on assisterait à une augmentation de 77% des consommations électriques des centres de données en France entre 2024 et 2035, puis une stabilisation, avec en 2050 -31% de consommations électriques des centres de données en France par rapport à 2024. Pour Technologies Vertes, autre scénario intermédiaire, on verrait une augmentation de 224% des consommations électriques des centres de données en France entre 2035 et 2024, et à long terme, une multiplication par 7,5 des consommations électriques des centres de données en France entre 2050 et 2024.¹⁰

A l'autre bout du spectre, si le scénario **Pari Réparateur** était adopté, la consommation des centres de données poursuivrait la forte croissance entamée actuellement pour aboutir à un total de **plus de 33 TWh consommés par les centres de données en France en 2035, soit 3,3 fois plus qu'en 2024, et près de 110 TWh en incluant les centres de données à l'étranger répondant aux usages des Français, soit 4,7 fois plus qu'en 2024**. Cette croissance serait principalement portée par le développement de la blockchain, l'IA et le recours croissant aux centres de données de type hyperscale pour les usages des IA et hors IA. Cette croissance resterait cependant légèrement inférieure au scénario Tendanciel sur le territoire français en 2035 (la situation s'inverse à plus long terme), en raison d'une sortie de terre plus progressive des projets de centres de données dans ce scénario.

En termes d'émissions de gaz à effet de serre (GES)¹¹, le scénario Tendanciel verrait les émissions provoquées par l'utilisation des centres de données présents sur le territoire français être multipliées par 3,3 entre 2024 et 2035, et par 4,8 d'ici 2050. En tenant compte des émissions provoquées aussi par les centres de données présents ailleurs dans le monde pour répondre aux usages des Français, les émissions passeraient de 7,31 Mt CO₂eq en 2024 à 23,44 Mt CO₂eq en 2035, soit une multiplication par 3,2. En 2050, les émissions du scénario Tendanciel monteraient à 55,63 Mt CO₂eq, soit une multiplication par 7,6 depuis 2024.

Long terme - Evolution des émissions de gaz à effet de serre (GES) pour l'usage français (Mega Tonnes CO₂eq / an)

	2024	2030	2035	2040	2045	2050	2055	2060
Tendanciel - en France	0,38	0,73	1,25	1,61	1,73	1,83	1,98	2,13
Tendanciel - import	6,93	14,03	22,19	38,23	48,57	48,36	56,80	60,38
Tendanciel - Total	7,31	14,76	23,44	39,84	50,30	55,63	58,78	62,51
Génération Frugale - en France	0,38	0,52	0,39	0,23	0,19	0,19	0,17	0,15
Génération Frugale - import	6,93	6,01	3,41	1,77	0,77	0,16	0,13	0,11
Génération Frugale - Total	7,31	6,53	3,81	2,00	0,96	0,35	0,29	0,26
Coopérations Territoriales - in France	0,38	0,60	0,59	0,39	0,29	0,23	0,20	0,18
Coopérations Territoriales - import	6,93	6,71	3,86	3,04	2,68	2,47	2,36	2,34
Coopérations Territoriales - Total	7,31	7,31	4,46	3,43	2,97	2,69	2,56	2,52

¹⁰ Pour plus de détails, voir à ce titre le chapitre [Résultats](#) de l'étude

¹¹ Estimation des émissions de gaz à effet de serre expliquées section [Émissions de Gaz à Effet de Serre](#)

Total								
Technologies Vertes - en France	0,38	0,68	1,09	1,64	1,86	2,15	2,32	2,51
Technologies Vertes - import	6,93	9,46	9,56	9,51	9,12	8,41	9,01	9,69
Technologies Vertes - Total	7,31	10,14	10,65	11,15	10,98	10,56	11,33	12,20
Pari Réparateur - en France	0,38	0,68	1,12	1,70	1,89	2,05	2,15	2,26
Pari Réparateur - import	6,93	14,58	25,04	43,65	54,74	63,85	67,75	72,20
Pari Réparateur - Total	7,31	15,25	26,16	45,35	56,64	65,90	69,89	74,46

Concernant les scénarios de Transition(s), il existe une variabilité significative des émissions provoquées par l'usage des centres de données pour les différents scénarios. Ainsi, concernant le scénario Pari Réparateur, les émissions de GES des centres de données présents sur le sol français seraient multipliées par 2,9 entre 2024 et 2035, et atteindraient 2,05 Mt CO₂eq en 2050, soit une multiplication par 5,4 par rapport à 2024. En tenant compte des émissions importées, le scénario Pari Réparateur monterait à 26,16 Mt CO₂eq au total en 2035, soit 3,6 fois le total des émissions en 2024, et 65,9 Mt CO₂eq en 2050, soit une multiplication par 9 du total des émissions.

A contrario, concernant le scénario Génération Frugale, les émissions de GES des centres de données présents sur le sol français seraient presque identiques en 2035 par rapport à 2024. En tenant compte des émissions importées, le total des émissions des centres de données pour répondre aux usages des Français serait réduit de moitié (-48%) en 2035 par rapport à 2024. En 2050, les émissions en France de ce scénario atteindraient 0,19 Mt CO₂eq, soit -50% par rapport à 2024. En tenant compte des émissions importées, le total des émissions atteindrait 0,35 Mt CO₂eq pour Génération Frugale en 2050, soit une réduction de -95% par rapport à 2024.

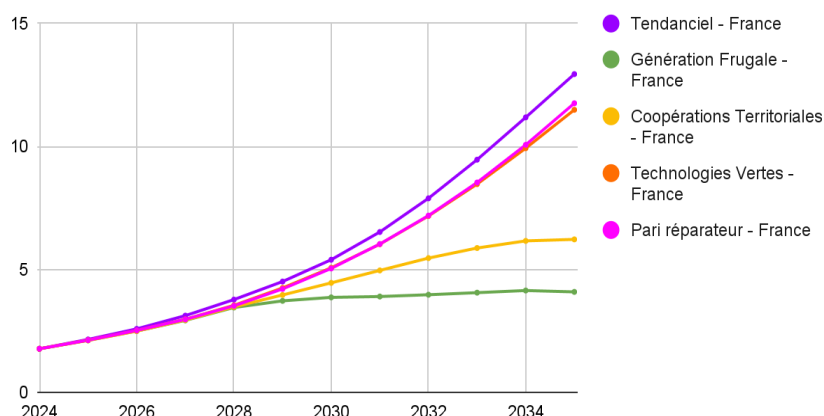
Le potentiel de récupération de la chaleur fatale

Quid de la récupération de la chaleur émise par les centres de données ? Nous avons estimé le potentiel de récupération de la chaleur émise par les équipements IT des centres de données, communément appelée "chaleur fatale", en fonction des scénarios définis dans le modèle prospectif. Le potentiel brut correspond à la quantité totale d'énergie émise sous forme de chaleur dans un centre de données et correspond à la consommation électrique de la partie IT. Le **potentiel net** de récupération de la chaleur fatale correspond au potentiel brut, pondéré par l'efficacité, actuelle et future, des différentes technologies de refroidissement présentes sur le marché, du point de vue de la récupération de chaleur et la capacité à trouver des débouchés. Cette efficacité augmente avec l'amélioration des techniques de captation de la chaleur et des débouchés plus atteignables, ce qui permet une amélioration du potentiel net de récupération. Ainsi l'utilisation du *liquid cooling* qui apparaît comme solution dans certaines salles informatiques permet de récupérer plus facilement la chaleur des serveurs et à une température plus élevée plus facilement valorisable. **Précisons également qu'il s'agit uniquement d'un potentiel réalisable d'un point de vue technique, qui ne correspond pas à la réalité à date puisqu'un nombre très restreint de centres de données français récupèrent réellement la chaleur fatale à ce jour.** Ces projections ne tiennent pas compte de l'intérêt économique réel de la récupération de cette chaleur et des particularités locales propres à chaque centre de données, ces éléments réduisant la probabilité d'une

telle récupération.¹² Cette projection passe de 1,77 TWh de potentiel net de récupération de chaleur sur le territoire français pour 2024, aux valeurs suivantes en 2035 :

- 12,94 TWh pour le scénario Tendanciel
- 4,09 TWh pour le scénario Génération Frugale
- 6,22 TWh pour le scénario Coopérations Territoriales
- 11,49 TWh pour le scénario Technologies Vertes
- 11,76 TWh pour le scénario Pari Réparateur

Projection de potentiel net de récupération de chaleur fatale en valeur absolue (TWh)



Ainsi, si la totalité du potentiel net de récupération de chaleur fatale était effectivement récupérée - hypothèse purement théorique et ne tenant pas compte de la faisabilité économique ni des réalités locales propres à chaque centre de données - cette récupération pourrait atteindre jusqu'à 23,6% de la consommation électrique des centres de données pour les usages des Français¹³ en 2035, dans le cas du scénario Coopérations Territoriales. *A contrario*, le scénario Pari Réparateur permettrait de récupérer en net jusqu'à 11,4% de la chaleur fatale de l'ensemble des centres de données du scénario, en 2035. Dans tous les scénarios, à l'échelle de consommations des centres de données présents sur le sol français uniquement, ce potentiel net de récupération de la chaleur fatale passerait d'un plafond de 18,8% en 2024 à un plafond de 38% en 2035. **Ainsi, si le potentiel de récupération de chaleur fatale est un levier intéressant à mobiliser pour optimiser l'usage de l'énergie, son intérêt est à mettre en regard du niveau de sobriété, du total de la consommation d'énergie en valeur absolue et de sa faisabilité.**

Quelle trajectoire de transition vers le Zéro Émissions Net ?

En résumé, une forte expansion tendancielle des consommations des centres de données entraînerait un accroissement important des émissions du secteur, incompatible avec l'Accord de Paris et la trajectoire SBTi du secteur du numérique. Concernant les scénarios Transition(s) 2050, l'ensemble de ces scénarios visent à atteindre le Zéro Émissions Net en 2050 sur les émissions du périmètre France¹⁴, ceci en tenant compte du fait que les centres de données ne sont qu'une des composantes parmi les secteurs calculés pour atteindre le Zéro Émissions Net dans une approche globale tous secteurs. Pour rappel, parmi les 4 scénarios Transition(s) 2050, l'atteinte du Zéro Émissions Net en 2050 se fait à travers des

¹² Pour plus de détails sur la méthodologie appliquée concernant l'estimation du potentiel de récupération de la chaleur fatale, voir la section correspondante dans le rapport

¹³ On tient compte ici de la consommation d'électricité totale nécessaire pour répondre aux usages des Français (celle des centres de données sur le sol français et celle de la part des centres de données étrangers sollicités pour cet usage)

¹⁴ c'est-à-dire, les émissions produites depuis le territoire français

mesures diverses et complémentaires, auquel chaque scénario a recours de manière variée.¹⁵ Dans le cadre de cette étude sur le périmètre des centres de données, nous avons pu voir que pour les cas des scénarios Technologies Vertes et Pari Réparateur, l'augmentation de la consommation électrique doit être anticipée et planifiée pour satisfaire des usages toujours croissants. La transition numérique est encouragée et ses bénéfices sont visibles dans la croissance du nombre de centres de données et la quantité de workloads (instances de calcul) traités. **Si les scénarios Technologies Vertes et Pari Réparateur soutiennent aussi la réduction des émissions d'autres secteurs d'activité, les émissions des centres de données dans ces deux scénarios sont toutefois très significatives et nécessitent une compensation importante des émissions.** Ainsi, à l'échelle des consommations des centres de données, les optimisations technologiques telles que l'augmentation de l'efficacité et la récupération de la chaleur fatale ne suffiront pas à stabiliser la croissance des consommations électriques des centres de données en France sans la mise en place d'une politique active de sobriété pour les services numériques et les usages. Par ailleurs, l'empreinte liée aux émissions importées sur ces scénarios est proportionnellement très importante au regard des émissions des centres de données en France.

Les cas des scénarios Génération Frugal et Coopérations Territoriales nous ont permis d'explorer que seule une politique de sobriété très volontariste, modifiant en profondeur nos modes de vie, permettrait un inflexion permettant de réduire les consommations futures des centres de données, permettant aux centres de données de contribuer moins fortement aux besoins de compensation pour atteindre le Zéro Émissions Net pour chacun des scénarios de Transition(s), et de réduire au global de façon importante leur empreinte. En raison de leur dimension systémique, de telles politiques ne peuvent réussir qu'à travers la mobilisation conjointe des pouvoirs publics, des entreprises et des citoyens. La modélisation des scénarios nous montre que leurs effets sur les consommations ne sont pas immédiats mais graduels sur plusieurs années.

¹⁵ Ces mesures (sobriété, efficacité, compensation par d'autres secteurs d'activités ou par captation, puits technologiques ou naturels...) n'ont pas été évaluées à l'échelle inter-sectorielle dans cette étude : nous nous focalisons ici sur les scénarios Transition(s) 2050 sur le périmètre des centres de données sans évaluation de leur contribution relative à la baisse des émissions sur d'autres secteurs (hors périmètre de cette étude).

Glossaire

Acronyme ou terme	Définition	Source
ACV	Analyse du Cycle de Vie	
AIE	Agence Internationale de l'Energie	
ASIC	Application Specific Integrated Circuit. Une puce de calcul conçue spécifiquement pour un usage donné, dès la phase de conception et de gravure, pour obtenir de meilleures performances pour cet usage.	https://en.wikipedia.org/wiki/Application-specific_integrated_circuit
Blockchain	"Développée à partir de 2008, la chaîne de blocs est, en premier lieu, une technologie de stockage et de transmission d'information. Elle est un registre ou encore une grande base de données, offrant de hauts standards de transparence et de sécurité. En effet, cette dernière est partagée simultanément avec tous ses utilisateurs, également détenteurs de ce registre, et ayant la capacité d'y inscrire des données, selon des règles spécifiques fixées par un protocole informatique sécurisé. La chaîne de blocs permet, par exemple, à ses utilisateurs - connectés en réseau à travers le monde - de soutenir des transactions, de cryptoactifs (le Bitcoin et l'Ethereum sont parmi les plus connus), sans passer par un organisme centralisateur, tel qu'une banque centrale."	https://www.economie.gouv.fr/entreprises/developper-son-entreprise/innovet-et-numeriser-son-entreprise/quest-ce-que-la-chaîne-de#
Cloud	"Le cloud est la possibilité de consommer des ressources informatiques (puissance de calcul, stockage, applications) comme on consomme de l'eau ou de l'électricité à la demande, sans posséder l'infrastructure technique, accessible via Internet. Cette approche permet aux entreprises d'accéder à des services performants tout en maîtrisant leurs coûts et leur flexibilité"	https://www.francenum.gouv.fr/formations/quest-ce-que-le-cloud-computing-ou-informatique-en-nuage-7-minutes-pour-comprendre
Co-hosting	désigne un centre de données dans lequel un ou plusieurs clients profitent d'un accès ou de plusieurs accès réseau, de serveurs et/ou d'équipements de stockage, sur lesquels ils opèrent leurs propres services et applications	rapport CLIK
Colocation	un centre de données dans lequel un ou plusieurs clients installent et opèrent leurs propres réseaux, serveurs et équipements de stockage, ainsi que leurs propres services	rapport CLIK
Compute	Calcul	
CPU	Central Processing Unit: processeur informatique classique	

Crypto actifs	Selon la Banque de France, un cryptoactif est « un actif numérique (ou digital), créé grâce à l'utilisation de technologies de cryptographie. Ils sont nommés ainsi car ils s'apparentent à des actifs financiers et sont créés et utilisés via des technologies de cryptage ». De manière globale, les cryptoactifs représentent des actifs virtuels stockés sur un support électronique permettant à une communauté d'utilisateurs les acceptant en paiement de réaliser des transactions sans avoir à recourir à la monnaie légale. La technologie employée est celle de la blockchain (chaîne de blocs ou registre de transactions, en français), qui permet de garder la trace d'un ensemble de transactions, de manière décentralisée, sécurisée et transparente.	https://www.economie.gouv.fr/particuliers/gerer-mon-argent/gerer-mon-budget-et-mon-epargne/crypto-actifs-cryptomonnaie-s-comment#
CUE	Carbon Usage Effectiveness	
Deep learning	Le deep learning ou apprentissage profond est un type d'intelligence artificielle dérivé du machine learning (apprentissage automatique) où la machine est capable d'apprendre par elle-même, contrairement à la programmation où elle se contente d'exécuter des règles prédéterminées.	rapport CLIK
DLC	Direct Liquid Cooling	
DLcC	Direct Liquid Cooling to Chip	
DRIEAT	Direction régionale et interdépartementale de l'énergie, de l'aménagement et des transports	
Edge	petits centres de données de proximité, la définition reste cependant floue	
EDNA	Efficient, Demand Flexible Networked Appliances	
EED	Energy Efficiency Directive (de l'Union Européenne)	
GO	Guarantee of origin: Un GO est un certificat permettant à l'acheteur d'1 MWh d'électricité d'attester que la production de cette énergie a été effectuée à partir d'énergies renouvelables.	rapport CLIK
GPU	Graphic Processing Unit: processeur plus puissant que les CPU pour usages graphiques et IA	
Hash	Les fonctions de hachage (hash functions en anglais) sont des suites d'opérations mathématiques et cryptographiques produisant un résultat, que l'on appelle également empreinte ou signature. Elles servent généralement à sécuriser un transfert d'informations entre deux systèmes informatiques.	https://cryptoast.fr/hash-hachage-bitcoin-blockchain/
HDD drive	les disques durs de type mécanique / Hard Disk Drive	rapport CLIK
High-end server	les serveurs hauts de gamme	rapport CLIK
Hyperscale	Dans le domaine informatique, l'hyperscale (en anglais hyperscale computing ou hyperscaling) est la capacité d'une architecture technique à s'adapter rapidement à des demandes importantes de ressources (dans des systèmes à grande échelle). L'approche consiste à permettre aux centres informatiques de prendre en charge des milliers de serveurs	Wikipedia

	virtuels, et adopter les outils permettant d'adapter rapidement les ressources processeur, mémoire et stockage aux besoins des applications. Par extension, un centre informatique, un datacenter, est dit hyperscale quand il utilise ces techniques pour héberger et exploiter plusieurs milliers de serveurs.	
IA	Ensemble des théories et des techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine.	
IA générative	L'IA générative est issue du Deep Learning. Sa fonctionnalité est de répondre à une demande ou requête appelée "prompt" en générant du contenu sur la base de contenus préexistants.	rapport CLIK
IA générative optimisée	Dans le cadre de cette étude, nous avons considéré que la consommation énergétique de l'IA générative, actuellement très importante, pourrait être optimisée. Nous avons donc introduit la notion d'IA générative optimisée, en ce sens.	rapport CLIK
IAAS	"L'infrastructure as a service (IaaS) ou infrastructure en tant que service est un modèle de service cloud qui permet aux entreprises d'héberger ses logiciels et bases de données. Avec l'IaaS, vous pouvez utiliser des services comme le calcul, le stockage et les bases de données d'un fournisseur cloud, faire héberger vos logiciels et matériels, assurer leurs maintenances."	https://www.franccenum.gouv.fr/formations/quel-est-ce-que-le-cloud-computing-ou-informatique-en-nuage-7-minutes-pour-comprendre
IEA	International Energy Agency	
IoT	Internet of Things: les objets connectés	
LLM	Large. Language Model	
M2M	Machine à Machine, se dit à propos de la communication entre les objets connectés, ou tout équipement informatique interagissant avec un ou plusieurs autres équipements informatiques pour créer de la donnée sans interaction humaine	
Machine learning	Le Machine Learning ou apprentissage automatique est un domaine scientifique, et plus particulièrement une sous-catégorie de l'intelligence artificielle. Il consiste à laisser des algorithmes découvrir des "patterns", à savoir des motifs récurrents, dans les ensembles de données.	rapport CLIK
Mid-range server	les serveurs de milieu de gamme	
MRAE	Mission Régionale de l'Autorité Environnementale	
On-premise	Terme désignant des infrastructures d'hébergement opérées physiquement par la même entreprise qui les utilise pour le fonctionnement de ses propres services numériques. Ce terme désigne de manière indifférenciée les équipements IT qui se trouvent dans une baie informatique ou une salle dans un datacenter de colocation, ou ceux présents dans un	

	datacenter également opéré par ou pour le compte de la même entreprise (datacenter dédié privé ou public).	
PDU	power distribution unit	
Port	le port réseau moyen présent sur un serveur	
PPA	Power Purchase Agreement	
PPE	Programmation Pluriannuelle de l'Energie	
Proof of Stake	La preuve d'enjeu (Proof of Stake ou PoS en anglais) est un algorithme alternatif permettant d'atteindre un consensus sur un réseau distribué, utilisé dans les systèmes de blockchain pour vérifier les transactions et ajouter de nouveaux blocs à la chaîne de blocs. Contrairement à la preuve de travail, la preuve d'enjeu ne nécessite que peu de puissance informatique, mais plutôt un montant spécifié de cryptomonnaie qui sert d'« enjeu », afin de participer au processus de création et d'ajout des blocs.	https://journalducoin.com/lexique/proof-of-stake/
Proof of Work	Le bitcoin est basé sur un protocole de Blockchain appelé "Proof Of Work". Il se base sur le minage, qui est un processus permettant de vérifier les transactions sur un réseau Blockchain à travers la résolution d'équations mathématiques très complexes et donc énergivores.	rapport CLIK
PUE	Power Usage Effectiveness	
Ramp-up	Rythme de remplissage d'un data center dont l'unité peut être la consommation électrique, tout comme le remplissage en mètres carré de surfaces commercialisables.	rapport CLIK
REC	Renewable Energy Certificate: équivalent des GO mais aux Etats-Unis	
Retail	Quand on parle de centre de données retail ou au détail, cela concerne les centres de données dont la puissance raccordée est en dessous de 40 MW	
SaaS	Le <i>software as a service</i> (SaaS) ou logiciel en tant que service est un modèle d'exploitation commerciale des logiciels dans lequel ceux-ci sont installés sur des serveurs distants plutôt que sur la machine de l'utilisateur.	Wikipedia
SDDR	Schéma Décennal de Développement du Réseau	
SNBC	Stratégie Nationale Bas Carbone	
SSD drive	les disques durs de type Solid State Drive	
Storage	stockage (qui se différencie du compute ou calcul)	
TIFCE	Taxe intérieure sur la consommation finale d'électricité, aussi appelée Accise sur l'électricité	
TPU	Un Tensor Processing Unit (TPU, unité de traitement de tenseur) est un circuit intégré spécifique pour une application (ASIC), développé par Google spécifiquement pour accélérer les systèmes d'intelligence artificielle par réseaux de neurones.	Wikipedia
UPS	Uninterruptible Power Supply	

Volume server	les serveurs d'entrée de gamme	
Wholesale	Quand on parle de centre de données wholesale ou en gros, cela concerne les centres de données dont la puissance raccordée est au dessus de 40 MW	
Workload ou compute instance	La notion de workload est définie par Cisco comme "un ensemble de ressources, virtuelles ou physiques, affectées à l'exécution d'une application ou à la fourniture d'un service numérique, à un ou plusieurs utilisateurs."	Cisco
WUE	Water Usage Effectiveness	
ZAN	Zéro Artificialisation Nette	
ZEN	Zéro Émissions Nettes	

1. Contexte des évolutions actuelles des usages et consommations des centres de données

1.1. Contexte

L'évolution de la consommation de ressources par les centres de données résulte de multiples facteurs, notamment l'évolution des usages numériques, du marché et des technologies. Ces infrastructures sont particulièrement consommatrices d'électricité, de foncier et dans certains cas, d'eau. La consommation d'électricité des salles informatiques s'accompagne d'une génération importante de chaleur, qui a son tour demande de consommer de l'électricité pour permettre le refroidissement des équipements. Ce processus, ainsi que la construction de ces centres de données et des équipements qui les composent sont source de divers impacts environnementaux. Elles s'inscrivent dans des cadres réglementaires, sociaux et politiques spécifiques et soulèvent des enjeux majeurs en matière de résilience face aux crises environnementales ainsi qu'aux risques de pénuries de ressources. A ce titre, les centres de données font l'objet de nombreuses études de référence et de modélisations.

1.1.1. Qu'est-ce qu'un Centre de données, "Data Centre" ou "Datacenter" ?

Il convient avant tout de s'accorder sur la définition d'un datacenter, "Data Centre", ou centre de données. Les normes EN 50600-1:2019¹⁶ et ISO/IEC 30134-1¹⁷ décrivent un centre de données comme suit :

"Structure ou groupe de structures dédiées à l'hébergement centralisé, à l'interconnexion et à l'exploitation d'équipements informatiques et de télécommunications en réseau fournissant des services de stockage, de traitement et de transport de données ainsi que toutes les installations et infrastructures de distribution d'énergie et de contrôle environnemental ainsi que les niveaux de résilience et de sécurité nécessaires pour assurer la disponibilité de service souhaitée".

Dans cette définition, une structure peut regrouper plusieurs bâtiments, ou inclure des espaces mutualisés, à condition que ceux-ci remplissent des fonctions spécifiques au service de la fonction principale. Les limites du périmètre considéré comme le centre de données peuvent ainsi être définies au sein d'une structure ou d'un bâtiment plus grand que les salles informatiques et les espaces techniques dédiés à l'hébergement.¹⁸

La directive 2024/1364 de la commission européenne¹⁹ propose plusieurs définitions supplémentaires :

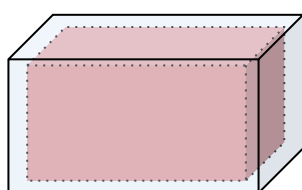
¹⁶ Norme EN 50600-1

¹⁷ Norme ISO/IEC 30134-1

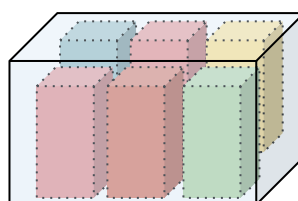
¹⁸ EN 50600 standards for data centres and their international implementation, Mike Gilmore, e-Ready Building Limited

¹⁹ COMMISSION DELEGATED REGULATION (EU) 2024/1364 of 14 March 2024 on the first phase of the establishment of a common Union rating scheme for data centres

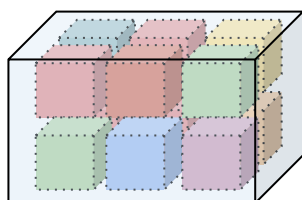
- Un “parc” ou un “campus” de centres de données est entendu comme un bâtiment ou une enceinte, qui comprend plus d’un centre de données.
- Un “centre de données d’entreprise”, désigne un centre de données opéré par une entreprise, pour le seul compte et usage d’héberger les infrastructures répondant aux besoins IT de cette entreprise.
- Un centre de données de colocation désigne un centre de données dans lequel un ou plusieurs clients installent et opèrent leurs propres réseaux, serveurs et équipements de stockage, ainsi que leurs propres services.
- Un centre de données de “co-hébergement” (“co-hosting”) désigne un centre de données dans lequel un ou plusieurs clients profitent d’un accès ou de plusieurs accès réseau, de serveurs et/ou d’équipements de stockage, sur lesquels ils opèrent leurs propres services et applications, dans le contexte où le matériel IT et l’infrastructure de support du bâtiment (sécurité, refroidissement, etc.) sont fournis comme un service par l’opérateur du centre de données.



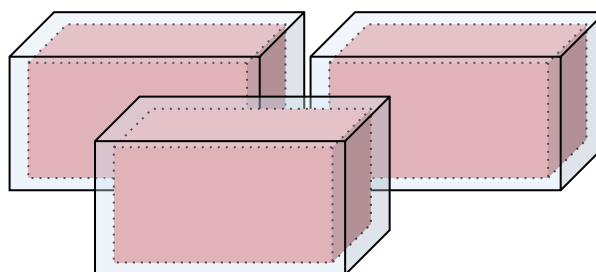
“Entreprise”



“Co-location”



“Co-hosting”



“Campus”

*Illustration non-officielle des définitions apportées par la directive 2024/1364
de la commission européenne*

Comme évoqué, la définition communément utilisée, provenant de l’EN 50600-1, ne se restreint pas aux bâtiments dédiés au seul usage de l’hébergement de matériel IT et de services numériques. Elle n’est également pas restrictive concernant la taille du bâtiment concerné et peut concerner un sous-ensemble d’un bâtiment aux multiples usages, qui aurait été équipé pour permettre l’hébergement de ces matériels et services. Peuvent donc être également considérés dans cette définition, les conteneurs métalliques (voir [la section sur les centres de données modulaires](#)) ou les sous-sols qui seraient utilisés à cet effet.

D’autres termes et définitions sont couramment employées, pour désigner des centres de données de tailles et de capacités différentes. Nous les parcourons dans la section suivante.

1.1.2. Description d'un centre de données

Un centre de données peut prendre plusieurs formes, tailles et s'insérer de différentes manières dans l'environnement urbain et péri-urbain. Nous détaillons les différentes typologies de centres de données que nous avons identifiées dans la section [2. Types de centres de données](#). La section [1. Architecture d'un centre de données](#) décrit les éléments composant couramment un centre de données.

1.1.2.1. Architecture d'un centre de données

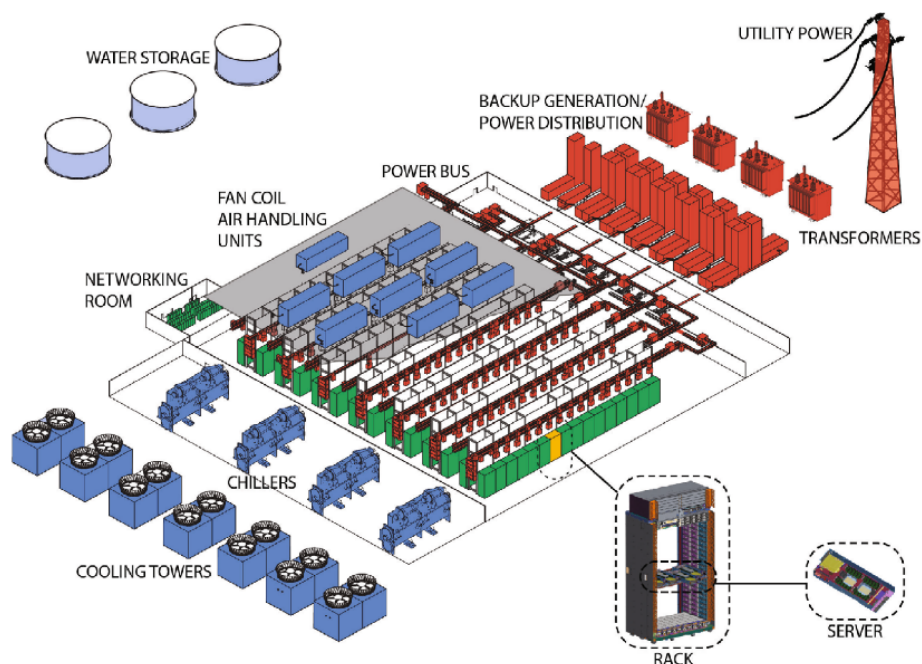
Un centre de données classique comprend :

- Une ou plusieurs salles spécialement conçues pour héberger des équipements informatiques, conformément à leur fonction principale (élément numéro 8 sur le schéma ci-dessous).
- Un système de refroidissement dont les ramifications commencent dès les salles IT. Celui-ci peut être de plusieurs sortes, détaillées dans la section [Techniques de refroidissement](#) (éléments numéro 4, 5 et 7 sur le schéma ci-dessous, dans un scénario de refroidissement par air, le plus commun).
- Un ensemble d'éléments dédiés à l'alimentation électrique, comprenant notamment :
 - Un "switch gear", équipement permettant de basculer automatiquement entre deux sources d'énergie électrique, couramment entre une ou plusieurs arrivées électriques provenant du réseau public et des génératrices de secours au fioul (numéro 1 sur le schéma ci-dessous).
 - Des génératrices d'électricité de secours (numéro 2 sur le schéma ci-dessous), qui fonctionnent la plupart du temps au diesel.
 - Des équipements permettant une alimentation électrique sans interruption ou UPS pour "Uninterruptible Power Supply" (numéro 3 sur le schéma ci-dessous). Les UPS nécessitent une capacité interne de stockage de l'énergie pour assurer l'alimentation électrique sans interruption (ils sont le premier rempart à la coupure nette de l'alimentation électrique, avant l'activation des génératrices de secours). Pour ce faire, ils sont reliés à des batteries, au lithium ou au plomb, le plus souvent (numéro 6 sur le schéma ci-dessous).



Vue aérienne d'un centre de données de Google, à Council Bluffs, Iowa, Source: The datacenter as a computer, Third edition, Barroso, Hölzle, Ranganathan²⁰

Certains centres de données comportent également des éléments externes supplémentaires, pour refroidir l'air, qui utilisent de l'eau : des conteneurs d'eau et des tours d'évaporation s'ajoutent aux éléments visibles de l'extérieur du centre de données, comme illustré sur le schéma ci-dessous (une description plus détaillée de leur fonctionnement est faite dans la section [Techniques de refroidissement](#)). Il faut noter que ces systèmes comprenant des tours d'évaporation sont très présents aux USA, mais n'existent pas sur le sol français, ce qui est notamment dû aux contraintes réglementaires liées à la prévention du risque de contamination à la légionellose²¹.



²⁰ The datacenter as a computer, Third Edition, 2022 - Barroso, Hölzle, Ranganathan

²¹ Tour aéroréfrigérantes, Aida, Ineris

1.1.2.2. Les typologies de centres de données retenues pour cette étude

En complément des définitions données par la directive 2024/1364 de la commission européenne, des critères de puissance électrique raccordée (réservée auprès du fournisseur d'électricité) sont couramment utilisés²³ pour classer les centres de données. Si l'on croise ces critères de puissance et les définitions citées précédemment, nous obtenons la grille suivante, dans laquelle nous avons indiqué les termes qui servent de vocabulaire de référence dans cette étude.

	Capacité (puissance électrique raccordée)		
Type (au sens de la directive 2024/1364 de la commission européenne)	< 10 MW	10 < x < 40 MW	> 40 MW
"co-hosting" (ou co-hébergement)	retail	wholesale	hyperscale
"co-location"	retail	wholesale	hyperscale
"entreprise"	privé	privé	privé
"entreprise" appliqué à une organisation publique	public	public	public

Les centres de données de type **retail**, sont constitués de bâtiments de taille réduite, avec une puissance électrique raccordée de moins de 10 MW. Ils sont le plus souvent situés dans ou aux abords des villes. Leur positionnement est généralement de proposer une connectivité forte à plusieurs opérateurs télécom et à d'autres entreprises qui souhaitent un point de présence local. Souvent présentés comme des "centres de données de proximité" par les opérateurs de centres de données, leur définition semble croiser celle du centre de données "edge" tels que présentés au sens marketing.

Les centres de données de type **wholesale** présentent une puissance raccordée plus importante, entre 10 et 40 MW. La taille des bâtiments et l'emprise au sol sont également plus importantes que les centres de données retail. Dans ce type de centres de données, les services de colocation peuvent dépasser la location d'une ou de plusieurs baies et concernent également la location de "suites" dédiées (rangées entières de baies, dont l'accès est sécurisé et réservé au client), voire d'une salle entière dédiée.

Les centres de données de type **hyperscale** sont encore plus imposants et la puissance raccordée dépasse les 40 MW. Les futurs exemples les plus extrêmes pourraient dépasser 1 GW avec pour principale finalité l'intelligence artificielle²⁴ et même plus de 2 GW si l'on en croit l'annonce de la construction du futur centre de données de Meta en Louisiane²⁵.

Les centres de données à destination et opérés par des institutions publiques, ou bien à destination exclusive et opérées par des organisations privées (dont le métier n'est pas de proposer des services

²² [The datacenter as a computer, Third Edition, 2022 - Barroso, Hölzle, Ranganathan](#)
²³ [Datacenter Anatomy Part 1: Electrical Systems, SemiAnalysis, 2024](#)
²⁴ [ECL construit le premier datacenter d'IA off-grid de 1 GW, DCMag, 2024](#)
²⁵ [Day after nuclear power vow, Meta announces largest-ever datacenter powered by fossil fuels, The Register, 2024](#)

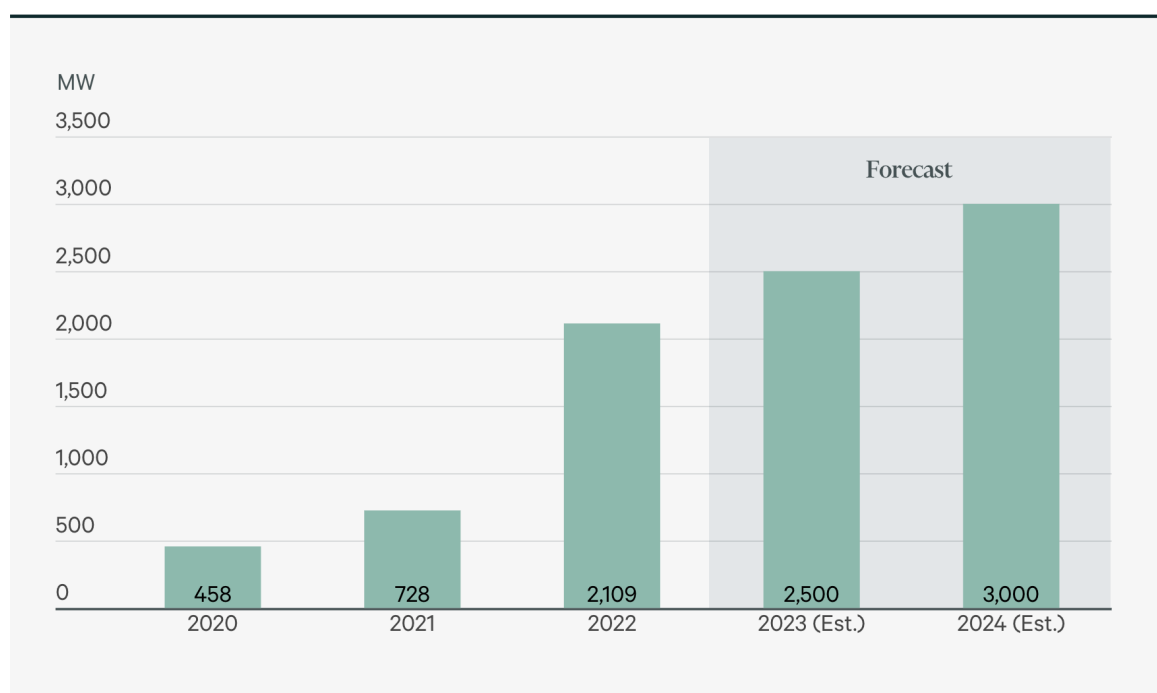
d'hébergement), semblent pour le moment proches des catégories "Retail" et "Wholesale" sur le territoire français, bien qu'ils puissent théoriquement également dépasser les 40 MW de puissance installée. Ils seront désignés par les catégories "Public" et "Privé", respectivement, lors de cette étude, puisqu'il s'agit de contextes spécifiques.

1.2. Évolutions du marché

Dans cette section, nous traitons de l'ensemble des signaux d'évolutions commerciales du secteur, quelques soient les métriques disponibles.

Les cabinets d'analyse marché montrent une accélération très visible dans le développement d'implantations de centres de données dans leurs publications récentes. Sur le marché états-unien, CBRE ²⁶ prévoit une croissance importante des constructions de nouveaux centres de données : de 458 MW de nouvelles capacités en 2020 à 3000 MW en 2024.

Figure 20: Under-Construction Capacity in Primary Markets



Source: CBRE Research, CBRE Data Center Solutions, 2023.

Des annonces faites à la fin de l'année 2024, pourraient fortement accélérer ces dynamiques. Meta prévoit la construction d'un centre de données d'une puissance installée supérieure à 2 GW, en Louisiane, pour une mise en service envisagée en 2029²⁷. Celui-ci pourrait fonctionner indépendamment du réseau électrique de l'État en étant alimenté par une production locale d'électricité, à base de gaz. De son côté, ECL projette la construction d'un centre de données dédié à l'IA, au Texas, doté d'une puissance initiale de 50 MW, extensible jusqu'à 2 GW grâce à une architecture modulaire ²⁸, appuyée par des [piles à hydrogène](#). Microsoft et OpenAI ont également annoncé leur intention de construire un campus de centres de données approchant les 5 GW, pour entraîner des modèles d'IA²⁹. Amazon, quant à lui, a déjà

²⁶ <https://www.cbre.com/insights/books/us-real-estate-market-outlook-2024/data-centers>

²⁷ Day after nuclear power vow, Meta announces largest-ever datacenter powered by fossil fuels, The Register, 2024

²⁸ ECL construit le premier datacenter d'IA off-grid de 1 GW, DCMag, 2024

²⁹ Microsoft & OpenAI consider \$100bn, 5GW 'Stargate' AI data center, Data Center Dynamics, 2024

fait l'acquisition d'un centre de données de 960 MW en Pennsylvanie³⁰. Cependant, dès juin 2025, certains acteurs comme Microsoft ont revu à la baisse leurs projets d'investissement dans de nouveaux centres de données.³¹

L'accélération de la croissance du développement des centres de données est aussi illustrée par l'Agence Internationale de l'Énergie (AIE) en prenant en compte l'évolution des investissements (valeur de 1 (index) est la valeur de référence des investissements réalisés dans les centres de données aux Etats-Unis en décembre 2019)³². Cette forte croissance des investissements lors de ces dernières années présage d'une forte croissance économique et énergétique à moyen terme.

Investment in data centres in the United States, January 2014 to August 2024

Open



IEA. Licence: CC BY 4.0

L'AIE note également qu'en 2023, l'investissement global de Google, Microsoft et Amazon, leaders du secteur en matière d'adoption de l'IA et d'installation de centres de données, a été supérieur à celui de l'ensemble de l'industrie pétrolière et gazière américaine, soit au total environ 0,5 % du PIB des États-Unis.

Le cabinet Dell'Oro estime que pour l'année 2023 les 11 principaux fournisseurs de services Cloud ont investi 260 milliards en centres de données.³³

En ligne avec cette tendance, FabricatedKnowledge illustre la croissance de plus en plus forte des investissements (CAPEX) de Google, AWS, Microsoft et Facebook entre 2017 et 2022³⁴ :

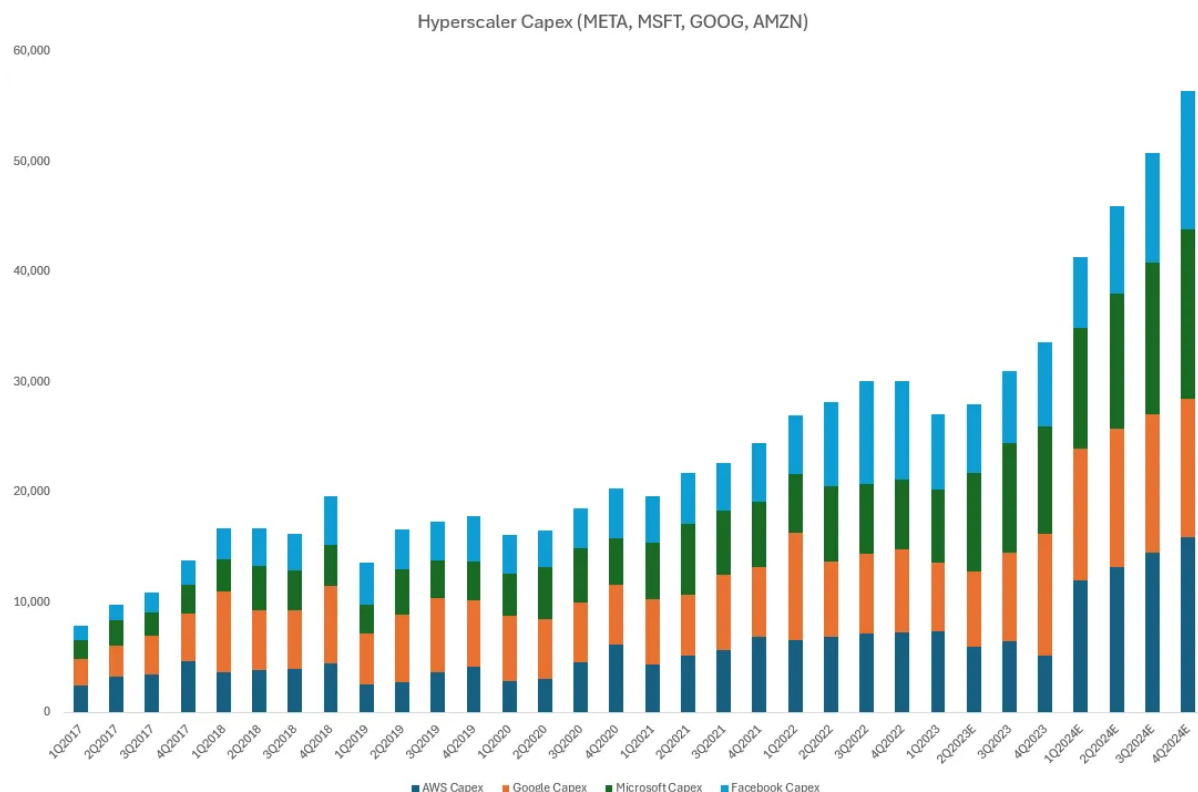
³⁰ [Amazon buys nuclear-powered data center from Talen, NuclearNewsWire, 2024](#)

³¹ <https://www.journaldunet.com/cloud/1541583-aws-et-microsoft-freinent-l-extension-de-leurs-data-centers-le-debut-d-une-crise/>

³² <https://www.iea.org/commentaries/what-the-data-centre-and-ai-boom-could-mean-for-the-energy-sector>

³³ <https://www.delloro.com/market-research/data-center-infrastructure/data-center-capex/>

³⁴ https://www.fabricatedknowledge.com/p/amds-mi300-disappointment-hyperscalers?publication_id=22108&post_id=144182985



D'un point de vue prospectif, sur le marché de la construction des centres de données, Global Market Insights projette une croissance annuelle moyenne de 6% pour passer de 212 milliards en 2023 à 367 milliards en 2032, soit plus de 73% en 9 ans ³⁵ :



³⁵ <https://www.gminsights.com/industry-analysis/data-center-construction-market>

1.2.1. Les principaux acteurs de la colocation et du IAAS

Si on analyse plus précisément les rapports financiers et extra-financiers des leaders du marché (Azure, Google cloud, AWS), on constate une croissance très importante de leur consommation électrique.

Concernant les acteurs de la colocation, la croissance annuelle moyenne de la consommation électrique entre 2019 et 2023 varie de 9 à 26% :

	Electricity consumption (TWh)					Croissance annuelle moyenne
	2019	2020	2021	2022	2023	
Equinix	5,7	6,4	7,14	7,75	8,17	
		12,28%	11,56%	8,54%	5,42%	9%
Digital Realty	6,9	8,3	9,3	10,3	11,05	
		20,29%	12,05%	10,75%	7,28%	13%
GDS		2,8	4,2	5,04	5,43	
			50,00%	20,00%	7,74%	26%
CyrusOne	2,51	2,8	3,2	3,4	4,2	
		11,55%	14,29%	6,25%	23,53%	14%

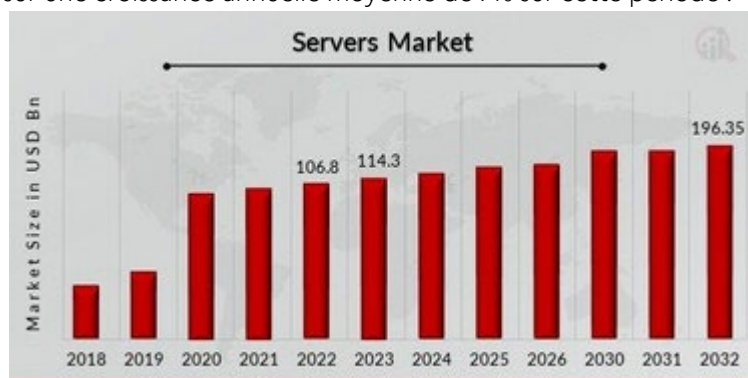
Pour les acteurs du IAAS la croissance annuelle moyenne de la consommation électrique entre 2020 et 2022 est encore plus importante. Elle varie de 22 à 30% comme indiqué également dans le rapport de l'ITU/Worldbank ³⁶ :

Electricity consumption (TWh)				
	2020	2021	2022	Croissance annuelle moyenne
AWS	12	15,4	17,7	
		28,33%	14,94%	22%
Azure	10,8	13,6	18,2	
		25,93%	33,82%	30%
Google Cloud	3,3	4,7	6	
		42,42%	27,66%	35%

³⁶ <https://documents1.worldbank.org/curated/en/099121223165540890/pdf/P17859702a98880540a4b70d57876048abb.pdf>

1.2.2. Les fournisseurs de hardware

Sur la partie hardware, le site MarketResearchFuture a effectué une étude de marché ³⁷ sur la période 2018-2032 et table sur une croissance annuelle moyenne de 7% sur cette période :



De son côté, TheInsightPartners projette une croissance annuelle moyenne de 9,7% qui aboutit à un doublement du marché entre 2023 et 2031. ³⁸

On retrouve des chiffres de croissance des revenus encore plus importants dans les derniers rapports de DELL, +38% en un an, et Lenovo, +65%, si on se focalise sur la partie équipements des centres de données. ³⁹

Plus précisément sur la partie IA, les chiffres de prévisions de croissance annuelle des revenus de Nvidia pour la partie centre de données sont spectaculaires : +427% entre 2025 et 2024 et +478% si on se focalise uniquement sur la partie "compute" ⁴⁰ :

Revenue by Market Platform					
(\$ in millions)	Q1 FY25	Q4 FY24	Q1 FY24	Q/Q	Y/Y
Data Center	\$22,563	\$18,404	\$4,284	Up 23%	Up 427%
Compute	19,392	15,073	3,357	Up 29%	Up 478%
Networking	3,171	3,331	927	Down 5%	Up 242%
Gaming	2,647	2,865	2,240	Down 8%	Up 18%
Professional Visualization	427	463	295	Down 8%	Up 45%
Automotive	329	281	296	Up 17%	Up 11%
OEM and Other	78	90	77	Down 13%	Up 1%
Total	\$26,044	\$22,103	\$7,192	Up 18%	Up 262%

1.3. Prospective sur les usages

Cette section explore les évolutions en cours ainsi que les premiers signes des changements à venir dans les usages. Les exemples présentés reposent fortement sur l'utilisation des centres de données.

³⁷ <https://www.marketresearchfuture.com/reports/servers-market-16189>

³⁸ <https://www.theinsightpartners.com/fr/reports/server-market>

³⁹ <https://www.lemagit.fr/actualites/366609998/Serveurs-les-ventes-de-Dell-et-HPE-progredient-sur-fond-dIA>

⁴⁰ https://www.fabricatedknowledge.com/p/nvidia-earnings-overview-networking?utm_source=profile&utm_medium=reader2

1.3.1. La “migration” vers le Cloud : une tendance de fond

De nombreuses études institutionnelles ou académiques se basent sur les données provenant de Cisco pour évaluer la tendance de migration vers le Cloud.

Selon Cisco, l'utilisation de services cloud par les entreprises et les particuliers est responsable depuis très longtemps de la majorité des activités de traitement, de stockage et de transmission de données - i.e. workload (*) - : ⁴¹

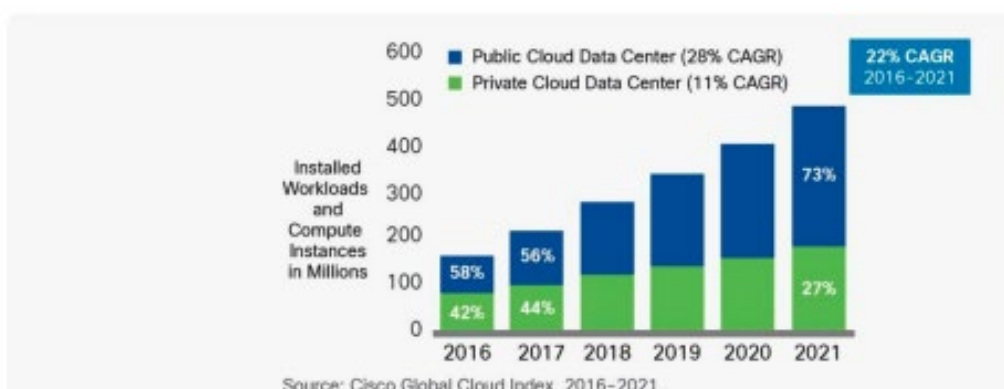
Figure 9. Workload and compute instance distribution: 2016–2021



(*) Selon Cisco, un “workload” et une “compute instance” désignent un ensemble de ressources, virtuelles ou physiques, affectées à l'exécution d'une application ou à la fourniture d'un service numérique, à un ou plusieurs utilisateurs.

Cisco fournit également des informations sur la manière dont les workload et compute instance des centres de données Cloud sont réparties entre services privés et publics. Selon leurs estimations, environ 70 % des workload et compute instance du cloud mondial seront opérées par des Cloud publics d'ici 2019.

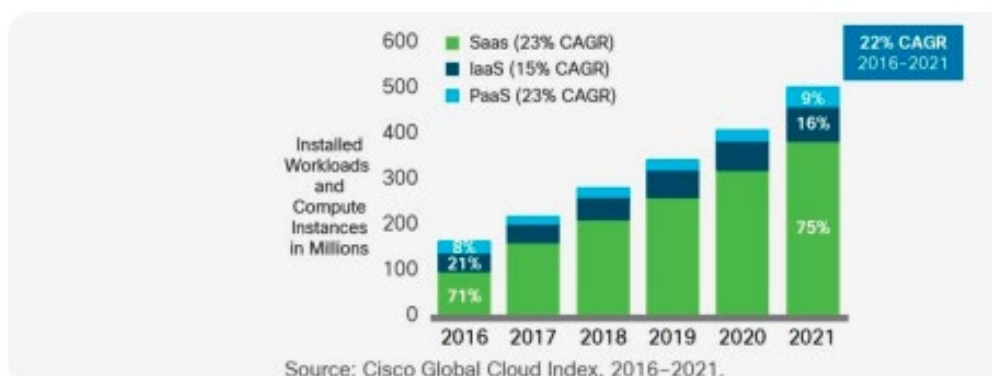
Figure 11. Public vs. private cloud growth



⁴¹ <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.pdf>

Cisco analyse aussi la nature des services Cloud entre les catégories IAAS, PAAS et SAAS. L'entreprise prévoit que le SAAS représentera en 2021 près de 75% des workloads.

Figure 13. SaaS most highly deployed global cloud service from 2016 to 2021



Avec un angle plus économique, le rapport "Energy-efficient Cloud Computing Technologies and Policies for an Eco-friendly Cloud Market" de l'Union Européenne ⁴² illustre cette tendance en compilant des études de marché des cabinets IDC, Gartner et Wikibon Research. Les croissances annuelles en revenus varient de 16 à 22%.

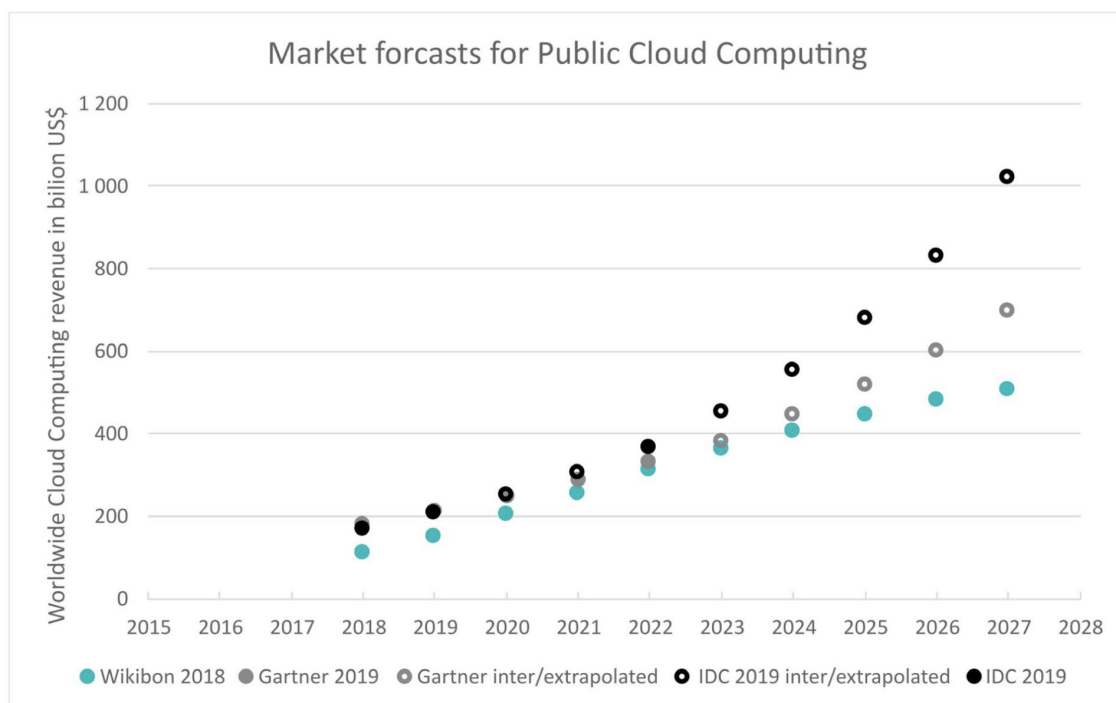


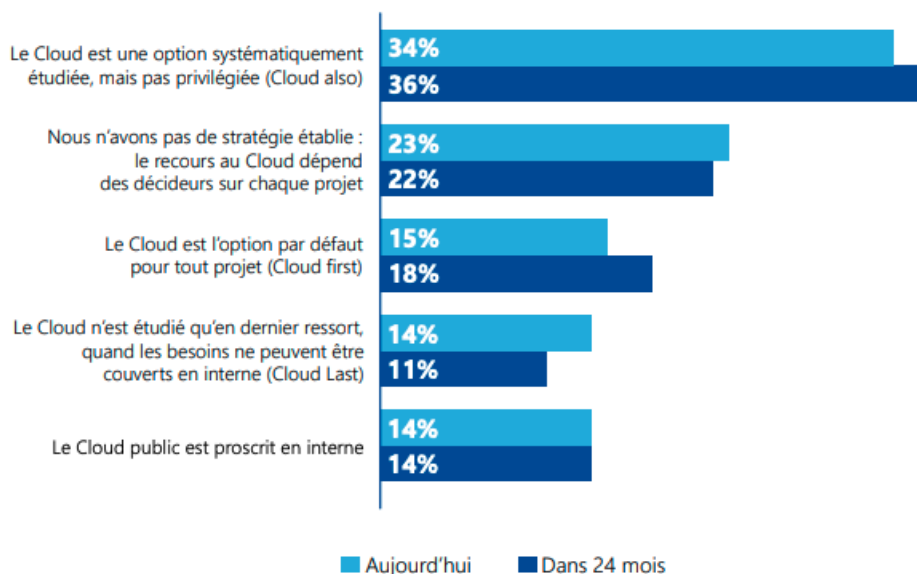
Figure 9 -Comparison of market forecasts for public Cloud Computing, Sources: (IDC, 2019a), (Gartner, 2019), (Wikibon Research, 2018), own inter- and extrapolation based on average CAGR

⁴² <https://digital-strategy.ec.europa.eu/en/library/energy-efficient-cloud-computing-technologies-and-policies-eco-friendly-cloud-market>

A l'échelle française, l'IDC a publié une étude menée auprès de 152 PME et ETI françaises. Cette étude met en évidence l'accélération des stratégies Cloud au sein des organisations.

GRAPHIQUE 2

Politique des entreprises concernant le recours au Cloud public

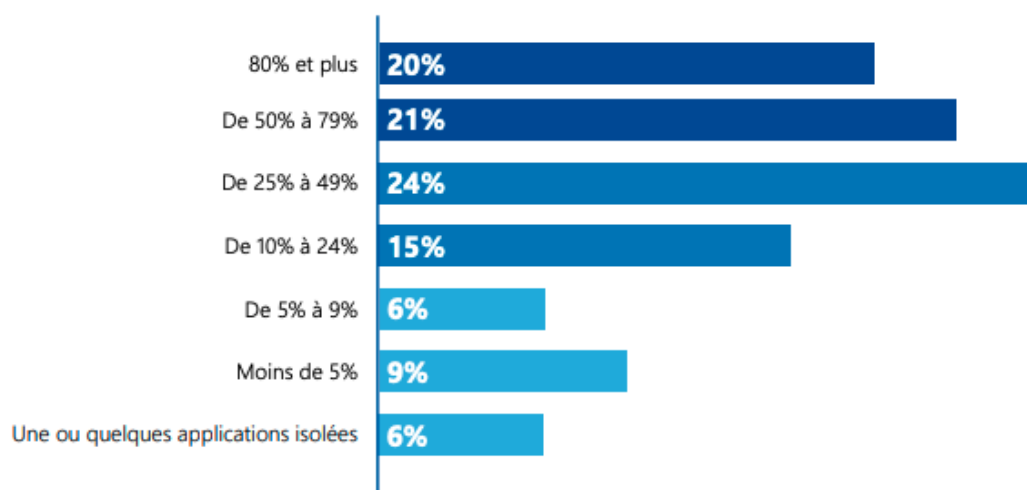


Source : IDC France, n=152

Selon IDC, cette tendance est appelée à s'élargir dans les 24 prochains mois, sous l'effet de deux dynamiques : l'utilisation de plus en plus fréquente d'applications SaaS et la migration d'applications critiques vers le Cloud.

GRAPHIQUE 3

Objectif de migration du portefeuille applicatif vers les architectures Cloud à 24 mois

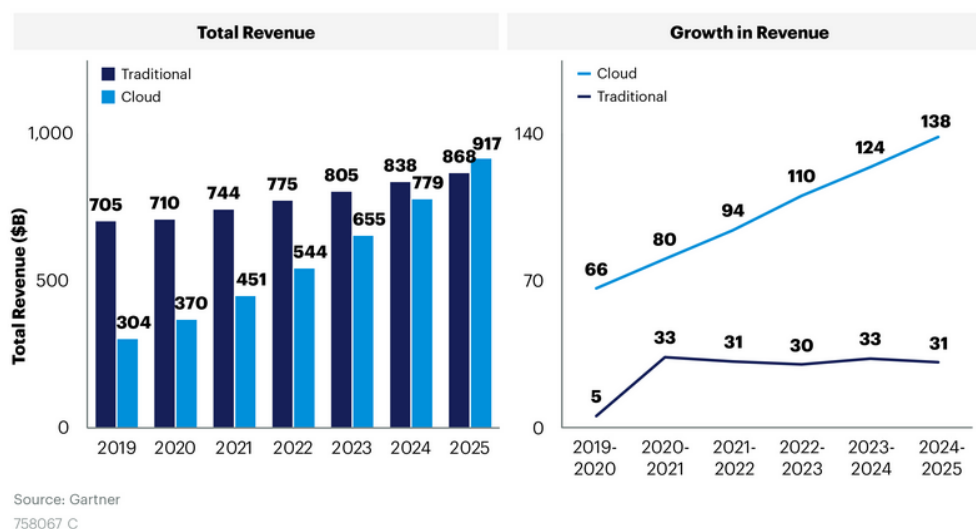


Source : IDC France, n=87 (entreprises ayant défini des objectifs de migration d'applications)

1.3.2. Une migration aux airs d'accumulation, plus que de transition

La notion de "migration vers le cloud" est largement relayée dans la presse. Toutefois, le marché ne s'est pas entièrement tourné vers le cloud à ce jour. Selon Gartner, si les infrastructures "traditionnelles" connaissent une croissance nettement plus modérée que celle du cloud, elles ne sont pas pour autant en déclin⁴³.

Figure 1: Sizing Cloud Shift, Worldwide, 2019 – 2025



Gartner

Comparaison des marchés du Cloud et des infrastructures traditionnelles avec projection, par Gartner, en février 2022

Nous assistons alors à une accumulation des infrastructures traditionnelles et des infrastructures Cloud, loin de la migration nette qui permet de décommissionner massivement les équipements physiques directement opérés par les entreprises qui pourraient être remplacées par des ressources Cloud.

Des initiatives inverses

Une partie du marché B2B a pourtant entamé le chemin inverse, du cloud à la colocation ou aux infrastructures en propre, principalement dans le but de mieux maîtriser les coûts d'infrastructure. C'est par exemple le cas de Dropbox⁴⁴, 37Signals et BaseCamp⁴⁵. Il semble cependant que ce mouvement inverse ait été minoritaire⁴⁶ au moins jusqu'en 2023. Plus récemment, les entreprises opérant des services d'IA générative ont semblé renforcer cette tendance en cherchant à réduire les coûts d'infrastructure, prohibitifs pour cet usage sur les plateformes Cloud⁴⁷.

Il faut noter que les limites du "Cloud", sont de plus en plus floues même dans un contexte B2B. De nombreuses entreprises désignent maintenant leurs infrastructures opérées en propre comme leur "Cloud privé". Les technologies permettant l'hybridation entre infrastructures on-premise et Cloud se

⁴³ [Gartner Says More Than Half of Enterprise IT Spending in Key Market Segments Will Shift to the Cloud by 2025](#)

⁴⁴ [Dropbox's Reverse Migration, From Cloud to Own Data Centers, Five Years On, DataCenter Knowledge, 2020](#)

⁴⁵ [Why we're leaving the cloud, 2022](#)

⁴⁶ [Is there really a march from the public cloud back on-prem?, TechCrunch, 2023](#)

⁴⁷ [Cloudy with a chance of GPU bills: AI's energy appetite has CIOs sweating, The Register, 2024](#)

multiplient (stockage des données hors-Cloud sur des supports propriétaires⁴⁸, systèmes d'orchestration de charges de travail pour les infrastructures on-premise⁴⁹ provenant de grands fournisseurs de services cloud, etc.). En France, des offres dites "Cloud de confiance" visent à répondre aux demandes croissantes de souveraineté et de sécurité des données (notamment face aux Cloud Act⁵⁰ des USA), ce qui n'améliore pas la lisibilité du marché en termes de répartition des usages entre plateformes Cloud et infrastructures opérées en colocation ou dans des centres de données dédiés.

1.3.3. Évolutions des usages du grand public

Selon la même étude Cisco mentionnée [plus tôt](#), cette tendance vers le Cloud s'accroît également pour le grand public.

Les deux schémas ci-dessous illustrent 1) que l'usage du cloud est très largement majoritaire dès 2016, par rapport aux centres de données traditionnels et par rapport à l'usage des entreprises 2) que la tendance s'accroît pour atteindre au moins 98% en 2021 pour l'ensemble des applications grand public :

Figure 18. Global data center workloads and compute instances by applications: Traditional vs. cloud (2016)

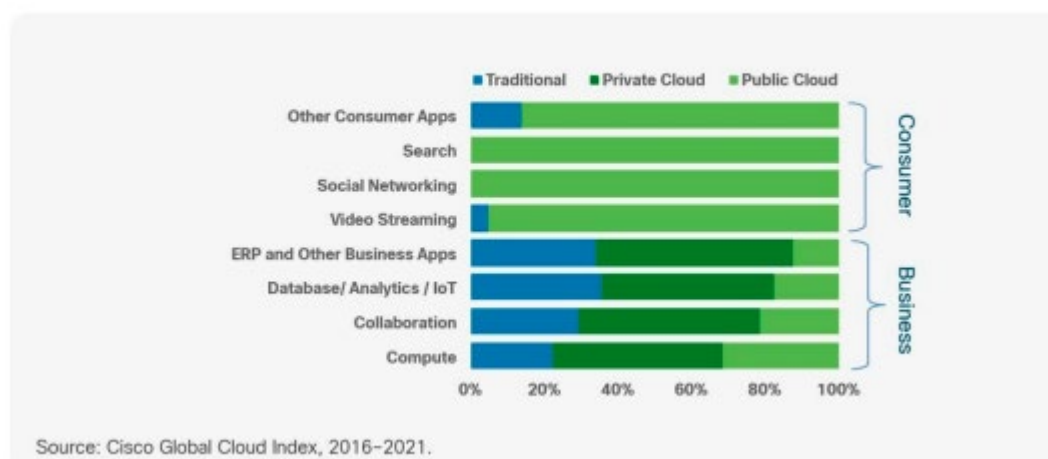
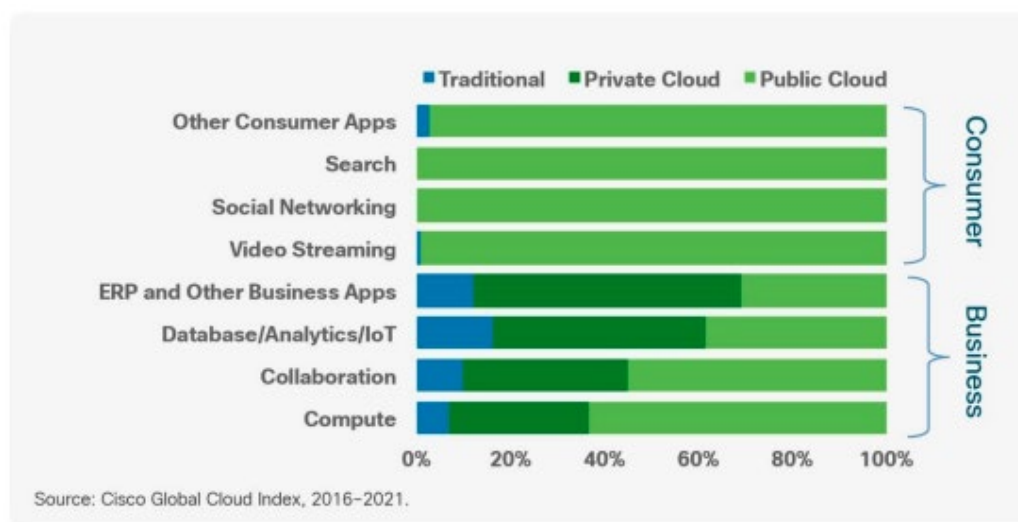


Figure 19. Global data center workloads and compute instances by applications: Traditional vs. cloud (2021)



⁴⁸ [AWS Snowball](#)

⁴⁹ [Bleu Cloud](#)

⁵⁰ [Cloud Act, US Department of Justice](#)

1.3.3.1. Le développement et l'adoption de l'IA

Les types d'IA

Si l'IA suscite autant d'attention depuis deux ans, c'est principalement en raison de l'essor spectaculaire de l'IA générative, et notamment du lancement de ChatGPT. Pourtant, l'IA existe depuis plusieurs décennies et on peut distinguer plusieurs types d'IA. Voici une version simplifiée des définitions proposées par Dominique Cardon ⁵¹ :

- **Intelligence artificielle** : Ensemble des théories et des techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine. ⁵²
- **Machine Learning** : Le Machine Learning ou apprentissage automatique est un domaine scientifique, et plus particulièrement une sous-catégorie de l'intelligence artificielle. Il consiste à laisser des algorithmes découvrir des "patterns", à savoir des motifs récurrents, dans les ensembles de données. Ces données peuvent être des chiffres, des mots, des images, des statistiques. Tout ce qui peut être stocké numériquement peut servir de donnée pour le Machine Learning. ⁵³
- **Deep Learning** : Le deep learning ou apprentissage profond est un type d'intelligence artificielle dérivé du machine learning (apprentissage automatique) où la machine est capable d'apprendre par elle-même, contrairement à la programmation où elle se contente d'exécuter des règles prédéterminées. ⁵⁴

Depuis les années **1950s** (Alan Turing, déchiffrement des messages chiffrés par d'Enigma). Terme "IA" utilisé pour la première fois en **1956**.

Ou bien depuis les **années 1980** ? avec les "**expert systems**" qui imitent la prise de décision humaine. (systèmes simples)

Le terme "Intelligence Artificielle" comprend aussi la robotique.

1er **réseau de neurones**, par [Marvin Minsky](#) and Dean Edmonds: le SNARC en **1951**

1ère **IA générative** en **1966**: le chatbot ELIZA, simule un psychothérapeute.

Deux "hivers de l'IA" dans les années 1970, puis les années 1990

Années **2010s**: **Deep Learning** et des **Generative Adversarial Networks**, modèle **Transformer**

Années **2020s**: essor des **Large Language Models**

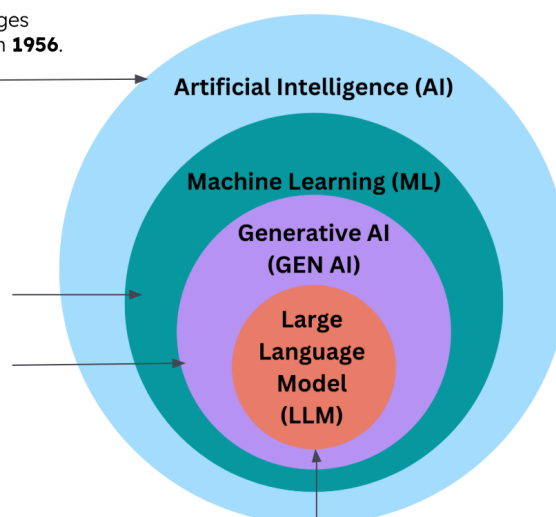


Illustration tentant d'illustrer l'histoire de la sémantique dans le domaine de l'intelligence artificielle. ⁵⁵ ⁵⁶

57 58

⁵¹ <https://www.youtube.com/watch?v=oCzqMkEaKh8>

⁵² https://www.larousse.fr/encyclopedie/divers/intelligence_artificielle/187257

⁵³ <https://datascientest.com/machine-learning-tout-savoir>

⁵⁴ <https://www.futura-sciences.com/tech/definitions/intelligence-artificielle-deep-learning-17262/>

⁵⁵ Expert systems, par E A Feigenbaum, 1980, archives de Stanford

⁵⁶ SNARC, Wikipedia

⁵⁷ ELIZA, Wikipedia

⁵⁸ [History of Generative AI](#), Brillworks, 2024

Les principales étapes de l'IA

En simplifiant, les deux principales étapes à considérer lorsque qu'il faut dimensionner et localiser des centres de données pour l'IA, sont :

- **L'apprentissage**

L'apprentissage ou l'entraînement dans l'IA consiste à alimenter les modèles avec de grands ensembles de données.⁵⁹ À mesure que l'apprentissage progresse, on peut avoir besoin d'affiner les modèles.

L'entraînement d'un modèle d'IA peut être très gourmand en puissance de calcul, mais de manière ponctuelle. Une fois qu'un modèle est correctement entraîné, il peut être utilisé. Si besoin, des mises à jour seront ensuite effectuées de manière ponctuelle. Ce fonctionnement asynchrone et ponctuel permet de localiser les phases d'apprentissage dans des centres de données qui ne sont pas forcément à proximité des utilisateurs.

- **L'inférence**

La phase d'inférence représente l'exécution d'un modèle une fois qu'il a déjà été entraîné. Ainsi l'inférence intervient après les étapes d'apprentissage. On peut également parler de phase de déploiement sur le terrain du modèle d'IA. Pour simplifier, à l'étape d'inférence, le modèle d'intelligence artificielle est déjà pré-calculé et a déjà été modélisé à l'aide de jeu de données (data set). Contrairement à l'apprentissage, l'inférence est une action continue. Cela nécessite une puissance de calcul assez importante et peut être très coûteux. Il est alors nécessaire que les moteurs d'inférence soient opérés dans des centres de données au plus près des utilisateurs.

L'IA générative est issue du Deep Learning. Sa fonctionnalité est de répondre à une demande ou requête appelée "prompt" en générant du contenu sur la base de contenus préexistants. En 2024, la page française de Wikipédia concernant l'Intelligence artificielle générative indique que l'IA générative "semble avoir des applications possibles dans presque tous les domaines, avec une balance des risques et des opportunités encore discutées : l'IA générative est en effet aussi source d'inquiétudes et de défis éthiques, techniques et socioéconomiques à la hauteur des espoirs qu'elle suscite."⁶⁰

Les évolutions en termes de services et d'usages étant très rapides dans ce secteur, il sera important de déterminer quels types d'IA vont être utilisées dans le futur, en ayant à l'esprit que la consommation en ressources peut varier très fortement en fonction du type d'IA utilisée.

Le boom de l'IA

Ces dernières années ont été marquées par une accélération extraordinaire de la croissance de l'intelligence artificielle (IA), qui promet de transformer en profondeur nos modes de vie, de travailler et d'interagir avec la technologie. L'IA générative, avec l'exemple de ChatGPT, joue un rôle central dans cette dynamique en tant que véritable catalyseur de croissance. Bien que des incertitudes subsistent quant à la place opérationnelle que va prendre l'IA, de nombreuses études sur son adoption dans des secteurs industriels ont déjà été publiées :

⁵⁹ Ces ensembles de données peuvent être structurés ou non, étiquetés ou non. Certains types de modèles peuvent nécessiter des exemples spécifiques d'entrées et de résultats souhaités. Pour d'autres modèles, les modèles de deep learning par exemple, des données brutes peuvent suffire. Les modèles finissent par apprendre à reconnaître des logiques ou des corrélations, et peuvent alors faire des déductions sur la base de nouvelles données.

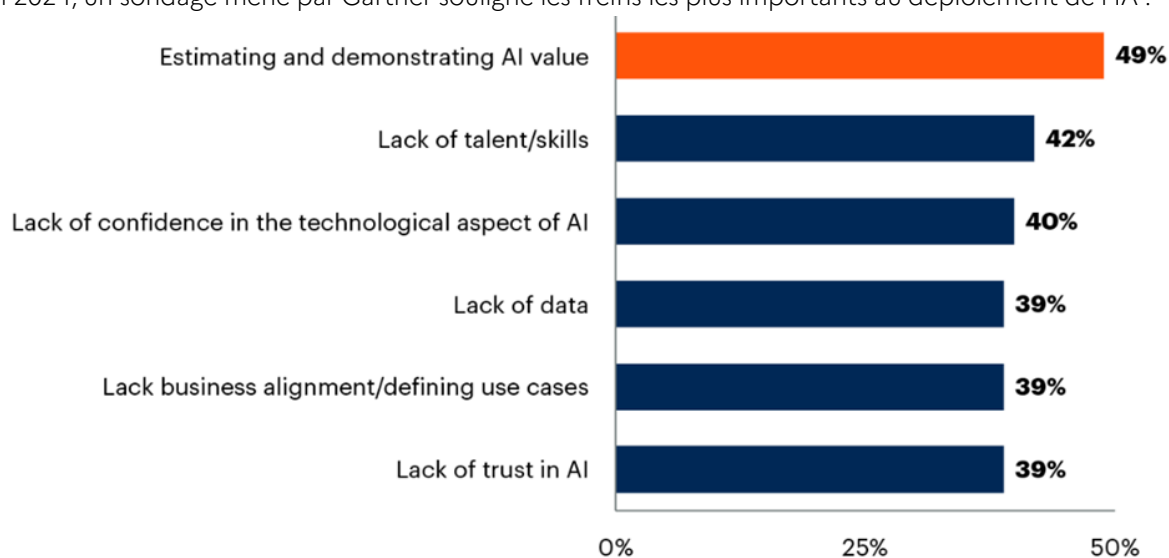
⁶⁰ https://fr.wikipedia.org/wiki/Intelligence_artificielle_g%C3%A9n%C3%A9rative

- Dans le domaine de la santé, il existe des travaux sur la convergence des intelligences humaine et artificielle qui mettent en avant la facilité et la rapidité d'interprétation ou la baisse potentielle des erreurs de diagnostic.^{61 62 63}
- Dans le secteur du manufacturing, l'utilisation de l'IA pour du monitoring et des tests d'assurance-qualité est déjà en place.⁶⁴
- Dans l'industrie du transport, l'usage de l'IA est en continuelle évolution pour adresser les nombreux challenges des véhicules autonomes tels que les interprétations en temps réel de l'environnement, la prévision du comportement humain, la gestion de scénarios de circulation complexes.⁶⁵
- En plus des transformations des secteurs d'activités actuels, il existe un scénario d'un déploiement massif et systématique de l'IA pour des usages qui ne le nécessitent pas forcément.

Les limites potentielles au développement de l'IA

En complément de l'évolution des usages et des scénarios de vitesse de déploiement, liés à la demande, il faut prendre en compte les nombreux freins potentiels dans ces différents secteurs : la garantie d'une cybersécurité solide, les problèmes de confidentialité des données, les cadres réglementaires etc... Ces limites sont notamment déjà discutées et documentées dans le secteur des véhicules autonomes.⁶⁶

Au-delà de ces limites sectorielles, il convient également de mettre en avant des limites opérationnelles. En 2024, un sondage mené par Gartner souligne les freins les plus importants au déploiement de l'IA :⁶⁷



n = 632, leaders highly involved in AI; excludes "unsure"

Q. What are or will be the top 3 barriers to the implementation of AI techniques within your organization?

Source: 2023 Gartner AI in the Enterprise Survey

810529_C

Gartner

⁶¹ <https://www.nature.com/articles/s41591-018-0300-7>

⁶² <https://pubs.rsna.org/doi/10.1148/radiol.210937>

⁶³ <https://www.sciencedirect.com/science/article/pii/S0933365723000209?via%3Dihub>

⁶⁴ <https://ijpem-st.org/journal/view.php?doi=10.57062/ijpem-st.2022.0017>

⁶⁵ <https://www.researchradicals.com/index.php/rr/article/view/62/59>

⁶⁶ <https://www.researchradicals.com/index.php/rr/article/view/62/59>

⁶⁷ <https://www.gartner.com/en/newsroom/press-releases/2024-05-07-gartner-survey-finds-generative-ai-is-now-the-most-frequently-deployed-ai-solution-in-organizations>

Un autre obstacle significatif au développement de l'IA générative réside dans le coût élevé des infrastructures nécessaires. L'investissement en termes d'infrastructure apparaît colossal (jusqu'à 40 000 \$ par GPU⁶⁸) et peut limiter drastiquement le nombre d'acteurs capables de se positionner sur ce marché. En conséquence, cette innovation reste largement entre les mains de quelques grandes entreprises privées et laissant peu de place à d'autres acteurs de l'innovation, notamment le monde académique. Alex Kantrowitz illustre cet écart en soulignant que l'université de Stanford ne dispose que d'environ 300 GPUs, à comparer aux 1,8 millions détenus par Microsoft.⁶⁹

La place de l'IA générative est structurante pour penser la prospective des usages numériques. Va-t-elle devenir la manière incontournable d'utiliser l'IA, quel que soit le besoin ? ou est-ce que nous adapterons le type d'IA au besoin ?

13.3.2. Le développement de l'usage de la blockchain et des cryptomonnaies

Maxime Gauthier, directeur de recherche, synthétise le marché et les enjeux des crypto-monnaies dans un article de Statista.⁷⁰ En synthèse, il soulignait en 2023 que la taille du marché était en constante augmentation et s'élevait à 12,7 milliards de dollars en 2022 soit quatre fois plus qu'en 2020, et devrait atteindre presque 40 milliards d'ici 2025.⁷¹

Le bitcoin leader du marché

Devenu un véritable phénomène attirant toujours plus d'investisseurs, le Bitcoin est la crypto-monnaie leader du marché et a vu sa valeur être multipliée par 30 entre avril 2017 et avril 2022. En avril 2013, le Bitcoin représentait plus de 94 % du marché des cryptomonnaies, et en avril 2022 malgré l'arrivée de nombreux concurrents, tel que l'Ethereum, il représente à lui seul toujours plus de 46,5 % du marché. À titre de comparaison, la quasi-totalité des autres principales cryptomonnaies possèdent moins de 3 % de part du marché.

Le Proof Of Work

Le bitcoin est basé sur un protocole de Blockchain appelé "Proof Of Work". Il se base sur le minage, qui est un processus permettant de vérifier les transactions sur un réseau Blockchain à travers la résolution d'équations mathématiques très complexes.

Chaque fois qu'un ensemble de transactions est validé, il constitue un bloc. Si ce bloc remplit certains critères spécifiques à la chaîne de blocs de la crypto monnaie, il est alors ajouté au sommet de la chaîne et le mineur qui a constitué ce bloc est récompensé pour son travail. Afin de participer à la blockchain il faut utiliser des logiciels spécifiques et rejoindre un "mining pool" qui dépendent de la crypto monnaie que l'on souhaite miner.⁷² Cela conduit naturellement à l'augmentation de la difficulté de minage. Cette dernière implique que les mineurs doivent posséder une capacité de traitement plus élevée pour résoudre les problèmes mathématiques et obtenir des récompenses de blocs.

⁶⁸ <https://qz.com/a-40-000-nvidia-chip-has-become-the-worlds-most-sought-1850746956>

⁶⁹ <https://kantrowitz.medium.com/universities-are-woefully-under-resourced-for-ai-research-theyre-fighting-to-change-that-ebe027cb538b>

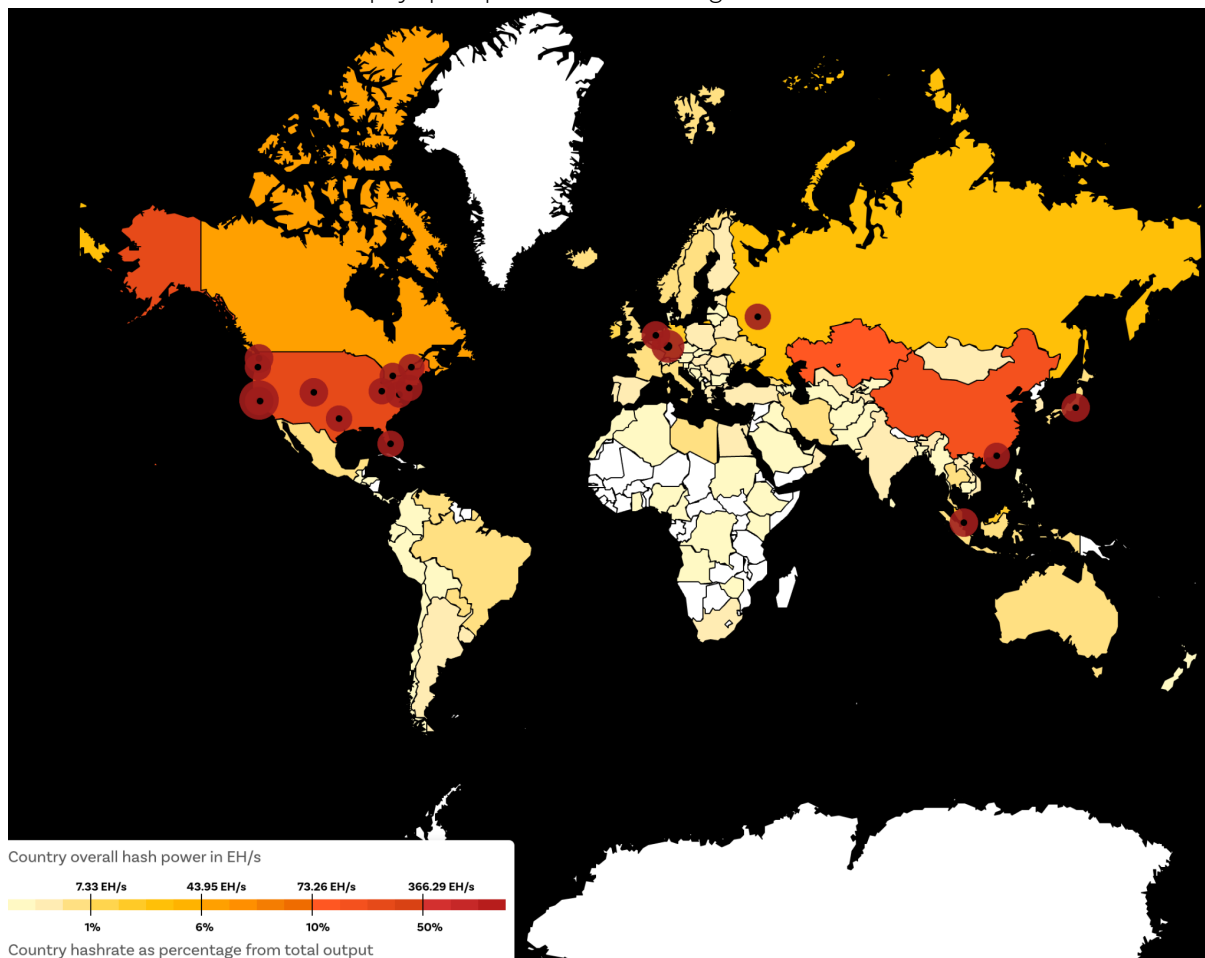
⁷⁰ <https://fr.statista.com/themes/9325/les-cryptomonnaies/#topFacts>

⁷¹ Aujourd'hui, il existe un nombre important de crypto-monnaies et différents moyens d'en acquérir : soit en les minant directement, soit par l'intermédiaire d'échanges réalisés sur des plateformes spécifiques avec des portefeuilles en ligne.

⁷² <https://www.cointribune.com/hashrate-btc-comprendre-son-importance-pour-la-securite-du-reseau-bitcoin/>

Des mines de plus en plus regroupées


En avril 2021, le plus grand site permettant le minage de Bitcoin était le "mining pool" chinois F2Pool. Néanmoins, en août 2021, le pays minant le plus de Bitcoin était les États-Unis, avec plus d'un tiers des minages mondiaux effectués. Pour avoir une vue plus actualisée, le site chainbulletin.com fournit une carte des mines en classant les pays par "puissance" de minage⁷³ :



La puissance de minage s'exprime en EH/s, c'est à dire en exa hash par seconde. Cela correspond à la puissance de calcul globale sur un réseau en vue de valider des nouveaux blocs d'une chaîne à un moment déterminé. Pour l'écosystème bitcoin, cela correspond donc à la capacité globale des ordinateurs de toutes les fermes de minage réunies. Le tableau suivant, également issu du site chainbulletin.com, illustre bien l'extrême concentration des pools de mineurs. En décembre 2024, les 5 premiers acteurs représentent 98% de la puissance disponible :

⁷³ <https://chainbulletin.com/bitcoin-mining-map/>

Pools

#	Pool	Hash Power 	Change
1	Foundry USA Pool	229.47 EH/s (31.3%)	2.66%
2	AntPool	161.24 EH/s (21.99%)	1.34%
3	ViaBTC	100.67 EH/s (13.73%)	7.67%
4	F2Pool	80.06 EH/s (10.92%)	11.92%
5	Unknown	77.18 EH/s (10.53%)	15.32%

Course à l'armement et enjeu énergétique

Le site Cointribune explique bien l'importance de "l'équipement minier".⁷⁴ L'évolution du matériel d'extraction a un impact significatif sur la puissance de calcul. Auparavant, les mineurs avaient essentiellement recours à des processeurs CPU puis à des cartes graphiques GPU. L'introduction des circuits intégrés spécifiques à l'application ou ASIC dans ce domaine a permis d'être encore plus performant. Ils offrent une capacité de traitement plus élevée, favorisant une "course à l'armement". L'adoption massive des nouveaux équipements de minage peut ainsi faire augmenter la puissance de minage globale. En plus de l'équipement minier, le coût de l'électricité est un facteur crucial pouvant influencer la variation de cette puissance. **Le processus de minage (en Proof-of-Work) étant très énergivore, le changement de prix de l'électricité dans un pays peut donc se répercuter de manière imprévisible et aléatoire sur la puissance de calcul disponible.**

Le recours à des sources d'énergie moins coûteuses peut augmenter la rentabilité du minage. Cette dernière permet d'attirer de nouveaux participants et donc de faire augmenter la puissance globale. **La localisation s'effectue souvent là où l'énergie est la moins chère.** Aux Etats-Unis par exemple, certains états comme le Texas hébergent de nombreuses mines gourmandes en électricité. À l'échelle mondiale, compte tenu de son faible coût de l'électricité, le Kazakhstan représente plus de 13% de la capacité de minage.⁷⁵ **Ce besoin vital d'électricité à prix compétitif peut aussi aboutir à la relance de centrales à charbon moribondes qui retrouvent une seconde vie.** Un article récent du Guardian documente en détails comment une centrale au charbon presque désaffectée du Montana, la centrale de Hardin, a été ramenée à la vie en 2020 par Marathon, une société de minage de bitcoins.⁷⁶

⁷⁴ <https://www.cointribune.com/>

⁷⁵ <https://chainbulletin.com/bitcoin-mining-map/>

⁷⁶ <https://www.theguardian.com/technology/2022/feb/18/bitcoin-miners-revive-fossil-fuel-plant-co2-emissions-soared>

Le Proof Of Stake comme solution ?

La spécificité du **"Proof Of Stake"** est de demander aux personnes qui veulent vérifier des transactions de **mettre en gage des tokens du projet (la preuve d'enjeu)**. Lorsqu'ils agissent "honnêtement", ils sont alors récompensés en obtenant la plupart du temps des jetons du projet. Par contre, lorsqu'ils tentent de tricher en n'approuvant pas correctement les transactions, alors ils peuvent être sanctionnés en perdant les jetons qu'ils ont mis en gage et en se faisant exclure de la liste des personnes autorisées à vérifier les transactions. Le Proof Of Work, avec le Bitcoin en tête, ayant une mauvaise réputation en termes de consommation énergétique, on pourrait penser que les solutions basées sur du Proof Of Stake - considérablement moins gourmandes en énergie - seront plébiscitées sur le marché des crypto-monnaies. Pour autant, les offres concurrentes du bitcoin en Proof Of Stake ne semblent pas gagner du terrain, bien au contraire. Ethereum, par exemple, numéro 2 du secteur, a réalisé une migration de protocole vers le PoS en 2022. Si l'opération a permis au réseau Ethereum de réduire la consommation énergétique de leur blockchain de 99%, il aurait dans le même temps sous performé Bitcoin de 44%.⁷⁷

L'impact énergétique global

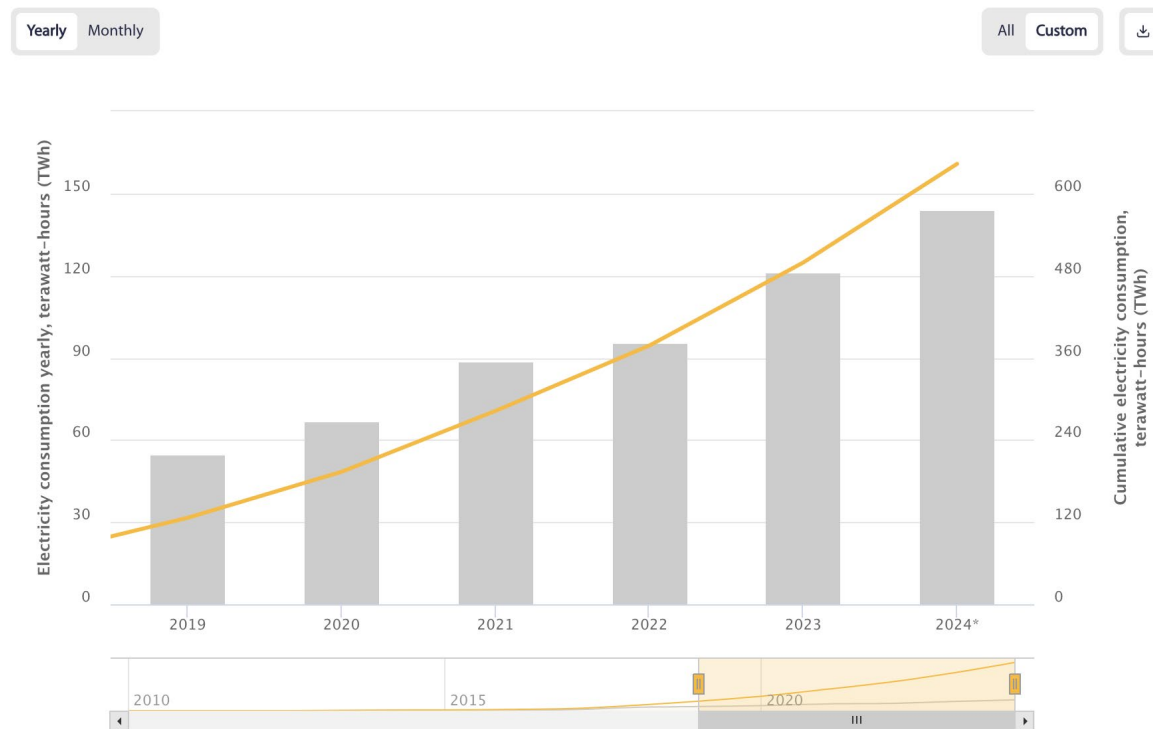
L'université de Cambridge a créé un site⁷⁸ pour suivre la consommation annualisée du Bitcoin via l'indice Cambridge Bitcoin Electricity Consumption Index (CBECI). La méthodologie appliquée repose sur une approche "top-down" hybride qui consiste à constituer un inventaire de matériel informatique en partant du principe que les nœuds de minage (« mineurs ») sont des agents économiques rationnels qui n'utilisent que du matériel rentable. L'estimation de la consommation annuelle du Bitcoin, au 3 décembre 2024, est de 173 TWh.

⁷⁷ <https://journalducoin.com/ethereum/depuis-passage-preuve-enjeu-pos-merge-ethereum-sous-performe-bitcoin-cryptoquant/>

⁷⁸ <https://ccaf.io/cbnsi/cbeci>

Total Bitcoin electricity consumption

Select a time window by clicking and dragging to define the start and end dates on the timeline below



* Year to Date (YTD)

Une étude américaine de whitehouse.gov estimait en 2022 que le bitcoin représenterait à lui seul 60 à 77% de la consommation totale d'électricité des crypto-monnaies et Ethereum 20% à 39%.⁷⁹ Bien que depuis, Ethereum soit passé au protocole de Proof Of Stake, on peut estimer aujourd'hui la consommation globale à 192 TWh en 2024 ⁸⁰

⁷⁹ <https://www.whitehouse.gov/ostp/news-updates/2022/09/08/fact-sheet-climate-and-energy-implications-of-crypto-assets-in-the-united-states/>

⁸⁰ https://blockchain-observatory.ec.europa.eu/document/download/3f78c885-d14e-47cb-b183-f22ef529a258_en?filename=EUBOF3.0_Ethereum_Merge_Trend_Report_final.pdf&prefLang=sl et calcul CLIK

1.3.3.3. Encadré: intérêt et limite des analyses marché

Est-ce l'évolution des usages qui guide l'évolution des technologies ou l'inverse ?

Les analyses de marché intègrent à la fois l'observation de nouveaux usages mais aussi l'émergence de nouvelles technologies. Ces analyses tentent d'anticiper le niveau d'adoption de l'un et l'autre. Cependant, la méthodologie d'analyse de marché permettant ces arbitrages est rarement transparente bien qu'il serait utile de pouvoir la connaître et en décomposer les paramètres pour modélisation.

Par ailleurs, les projections des analyses de marché ne tiennent pas compte d'un certain nombre de freins à la généralisation d'un nouvel usage.

Pour favoriser un regard critique sur l'analyse des nouveaux marchés liés à une innovation numérique, Gauthier Roussilhe propose un modèle ⁸¹ basé sur 4 questions:

Cette innovation nécessite-t-elle :

- du nouveau matériel (hardware) pour le grand public ?
- de nouveaux services ?
- de nouveaux usages ?
- de nouveaux modèles d'affaires ?

Son hypothèse est que "plus un système numérique lancé aujourd'hui à l'échelle globale nécessite de modifier matériel, services ou usages, moins il a de chances de se développer."

Dans le cas du metaverse de Meta, on peut dire que l'innovation nécessite trois changements : du nouveau matériel, des nouveaux services (Horizon Worlds) et des nouveaux usages (interactions homme-machine, micro-paiements, ...).

Cette approche pourrait aussi être appliquée à l'Intelligence Artificielle : on peut considérer que l'IA générative ne nécessite pas de nouveau matériel côté usagers grand public (c'est moins le cas du côté des centres de données des fournisseurs). Si de nombreux usages sont encore à inventer ou à finaliser, certains ont déjà vu le jour, comme le "prompt" via interface. Par contre, les nouveaux modèles d'affaires ne sont pas encore clairement identifiés au regard du coût croissant des infrastructures nécessaires.

De même, la promesse de la blockchain laisse entrevoir des améliorations technologiques autour de ses fonctions principales. Pour autant, c'est l'avènement de la technologie qui a entraîné des propositions de nouveaux usages (crypto monnaies, NFTs, ..), plutôt que l'émergence forte de nouveaux usages qui a conduit au développement de la blockchain.

D'un point de vue prospectif, si on applique l'approche proposée par Gauthier Roussilhe, la blockchain nécessite de nouveaux services et certainement également de nouveaux modèles d'affaires.⁸²

⁸¹ <https://gauthierroussilhe.com/articles/projeter-les-futurs-environnementaux-de-la-numerisation>

⁸² Note de la rédaction : Cette méthode pour évaluer l'adoption d'une nouvelle innovation pourrait être utilisée dans l'étude pour modéliser l'adoption d'une application émergente. Elle invite aussi à faire un usage distinct et clair des termes "projection"

1.4. Évolutions des technologies et des pratiques du secteur

1.4.1. Évolution des techniques et montages liés à l'énergie

Au-delà du recours au réseau électrique national, stable, fiable et décarboné, les opérateurs de centre de données souhaitent afficher un recours fort aux ENR via des certificats de garantie d'origine, mais anticipent également des innovations pouvant leur être utiles dans le futur : la pile à hydrogène et les petits réacteurs nucléaires notamment.

Énergies Renouvelables (ENR)

L'utilisation et la mise en avant de l'électricité produite à base d'énergies renouvelables est un sujet porteur dans la filière des centres de données. Si l'on peut noter certains cas particuliers de centres de données consommant principalement de l'électricité produite à partir de barrage hydroélectriques,⁸³ les ENR principalement citées dans ce secteur concernent les énergies solaires et éoliennes⁸⁴. L'utilisation d'électricité de provenance ENR est un véritable enjeu de communication pour les centres de données. Il est ainsi fréquent de lire des présentations de centres de données "alimenté par une énergie 100 % verte et renouvelable" comme chez Equinix⁸⁵. AWS annonce par exemple avoir atteint en 2023 l'objectif d'utiliser 100% d'énergies renouvelables pour l'électricité consommée par l'entreprise⁸⁶. Parmi les actions mises en avant par les entreprises signataires du Climate Neutral Data Centre Pact, initiative privée d'auto-incitation du secteur à une moindre empreinte carbone, "acheter de l'énergie 100% décarbonée" ("carbon-free energy"), est la première listée⁸⁷. Pour introduire ses lignes directrices de réduction de l'empreinte carbone du numérique, la Science Based Target Initiative (SBTI) indique que "historiquement, le secteur numérique a été en capacité de répondre aux demandes de capacité et de vitesse croissantes, grâce à l'avancement technologique et l'achat d'énergie renouvelable"⁸⁸. Derrière ces affirmations, une explication est nécessaire pour comprendre comment les énergies renouvelables sont effectivement utilisées par le secteur.

ENR locales en auto-consommation

À Marseille, le cinquième centre de données de Digital Reality (Interxion), MRS 5, qui est au centre de vives critiques (voir la section [Impacts territoriaux et oppositions locales](#)), est un cas d'étude intéressant. Le dossier soumis par Digital Reality à la Mission régionale d'autorité environnementale Provence-Alpes-Côte d'Azur, stipule "Afin de réduire l'impact sur l'émission de gaz à effet de serre, le groupe DIGITAL REALTY poursuit son engagement en faveur des énergies renouvelables à hauteur de 100% de la consommation d'électricité de l'ensemble de ses datacenters."⁸⁹. L'avis délibéré de la MRAE PACA sur le projet indique en ce sens "Le maître d'ouvrage s'engage à utiliser **de l'électricité en provenance exclusivement de sources renouvelables** mais aussi que si des panneaux photovoltaïques sont bien prévus en façade sud-ouest

prolongement ou inflexion dans le futur de tendances passées
diagnostic initial.

- et "prospective" - anticiper des évolutions possibles à partir d'un

⁸³ [How we run it with water, DatacenterLight](#)

⁸⁴ [MARA Acquires Wind Farm – Site features 240 MW of interconnection capacity, with 114 MW of operational wind generation to sustainably power data center, Data Centre Central, 2024](#)

⁸⁵ [Datacenters d'Equinix à Barcelone, c'est également le cas de la plupart des Datacenters de l'entreprise](#)

⁸⁶ [AWS favorise des solutions de développement durable, Amazon](#)

⁸⁷ [Climate Neutral Data Centre Pact](#)

⁸⁸ [Guidance for the ICT sector, SBTi](#)

⁸⁹ [Demande d'autorisation environnementale, Étude d'impact projet MRS5, Interxion](#)

et sur une partie du toit des bureaux, "le dossier *ne précise pas la consommation énergétique globale du projet*". La MRAE recommande donc "de détailler la répartition des consommations des installations" et "de démontrer *que la production d'énergie renouvelable est suffisante pour faire fonctionner la totalité du centre d'hébergement de données dans des conditions normales*".

Une estimation de la capacité de production annuelle des panneaux photovoltaïques locaux se trouve dans le dossier : 429 MWh, ce qui équivaldrait à une puissance utilisable de 0.05 MW. L'estimation de la consommation d'énergie totale du centre de données est de 241 133 856 MWh, ce qui demande à priori une puissance disponible d'au moins 27 MW.

Pour comprendre comment Digital Reality peut affirmer une utilisation d'électricité provenant à 100% de renouvelables, alors que les panneaux photovoltaïques installés localement ne permettent de fournir au mieux que 0.018 % de la puissance nécessaire sur site, il faut se référer à la mention de "**certificats de garanties d'origine de l'électricité**" présente dans le même document.

Achat d'électricité et mécanisme de marché : GoOs, REC, PPA

Les certificats de garantie d'origine, souvent labellisés "GO" ou "GoO" pour "Guarantee of Origin" sont un mécanisme défini dans la directive 2018/2001 de l'Union Européenne⁹⁰. **Un GO est un certificat permettant à l'acheteur d'1 MWh d'électricité d'attester que la production de cette énergie a été effectuée à partir d'énergies renouvelables.** Cependant, ce mécanisme est vivement critiqué pour plusieurs raisons.

En premier lieu, concernant spécifiquement les centres de données, l'affirmation de consommation d'électricité produite **exclusivement** à base d'énergies renouvelables est trompeuse, puisque le MWh d'électricité provenant d'ENR peut être produit n'importe où, indépendamment de la localisation de l'acheteur (sauf pour les PPA physical/on-site, voir plus loin). C'est donc un mécanisme de marché qui ne permet pas de s'assurer de la nature de l'énergie consommée (puisque le centre de données consomme bien 1 MWh provenant du réseau électrique commun), mais de **rémunérer le producteur pour avoir produit un autre MWh d'énergie, en provenance d'ENR**. Ce mécanisme a donc un potentiel positif pour le développement des filières économiques liées aux ENR, mais n'est pas un levier effectif de réduction d'empreinte environnementale liée à l'électricité consommée en centres de données.

Deuxièmement, il existe un **risque de comptabiliser deux fois un GO**, donc que deux entreprises affirment avoir consommé le même MWh d'énergie provenant de renouvelables, comme le démontre le cas de l'Islande en 2023.⁹¹

Enfin, l'achat d'un certificat de provenance de l'énergie de type GO, dans sa version "unbundled", **ne mène pas forcément à la mise en place de nouveaux moyens de production d'électricité de provenance ENR**, puisque des certificats sont émis pour des moyens de production déjà existants⁹².

Les RECs, ou Renewable Energy Certificate, sont un autre mécanisme de marché similaire aux GoOs, que l'on trouve dans les documentations de centres de données basés aux USA, ou dont l'opérateur est une entreprise états-unienne.

Dernier mécanisme, de plus en plus utilisé dans le secteur : **les PPA, ou Power Purchase Agreement**. Un Power Purchase Agreement (PPA) est un contrat d'achat d'électricité sur le long terme entre un

⁹⁰ Directive EU 2018/2001 on the promotion of the use of energy from renewable sources

⁹¹ How Iceland sold the same Green Electricity twice, Industry Decarbonization Newsletter, 2023

⁹² Renewable Energy Purchasing and the Market-based (Scope 2) Method, Université d'Edimbourg

producteur d'électricité et un acheteur. Ce contrat définit un prix fixe de l'électricité sur une durée généralement longue (10-15 ans). L'avantage du PPA sur le GO est donc qu'il aide le producteur d'électricité contre la volatilité des prix du marché de l'électricité et encourage donc davantage à l'installation d'énergies renouvelables.

Il existe en fait plusieurs types de PPAs⁹³:

Types de PPA	Caractéristiques
PPA sur site (physical, on-site)	L'électricité est produite et livrée sur site de façon physique et directe.
PPA hors site (physical, off-site)	Le producteur livre l'énergie renouvelable au consommateur via le réseau de distribution public, contrairement au PPA local.
Accord d'achat de PPA hors site Ou Portefeuille de PPA hors site (sleeved)	C'est un PPA hors-site mais dans lequel un fournisseur de services énergétiques joue l'intermédiaire entre les producteurs et le consommateur. Il se charge de différents processus (gestion du groupe d'équilibrage, rassemblement de différents producteurs d'électricité, etc.)
PPA synthétiques Ou virtuels (virtual)	Il s'agit de contrat qui dissocie le flux physique de l'énergie renouvelable du flux financier qui permet des arrangements contractuels encore plus flexibles.

Les PPA sur site, comme l'installation locale en propre de moyens de production ENR, dans le cas des centres de données, semblent être plus l'exception que la règle. On peut tout de même noter l'usage de panneaux solaires sur le toit, ou à proximité, du bâtiment par certains acteurs, pour auto-produire une petite partie de l'énergie nécessaire, mais aussi pour réduire la chaleur subie par le centre de données en été⁹⁴.

Il semblerait de plus que l'adoption des PPA soit encore minoritaire sur le sol français, les opérateurs voyant moins l'intérêt de compenser l'électricité consommée, qui est plutôt bas carbone en France par rapport à la moyenne européenne⁹⁵. Tendance pourtant notable, Digital Realty⁹⁶, Equinix⁹⁷ et Data4⁹⁸, semblent s'engager dans cette voie en 2024.

Les certificats de garantie d'origine, qu'ils soient de type GoOs, REC ou PPA, sont souvent un levier utile pour les opérateurs de centres de données et de services cloud. La comptabilisation des impacts environnementaux associées à la consommation d'électricité peut être ainsi orientée de manière à présenter une empreinte carbone allégée. Les données d'empreinte carbone de l'électricité consommée sont alors présentées dans leur version "market-based", c'est-à-dire que les mécanismes de marché, dont les certificats de provenance de l'énergie, sont pris en compte dans le calcul. Cette pratique est

⁹³ [FAQ Operat, sur les PPA](#)
⁹⁴ [Visite #1 et #2 d'un centre de données d'Infomaniak, cocadmin](#)
⁹⁵ [Intensité carbone de l'électricité Française, ElectricityMap](#)
⁹⁶ [Digital Realty signe deux contrats d'achat d'énergie renouvelable en France avec wpd, Wpd, 2024](#)
⁹⁷ [Equinix et wpd concluent l'un des plus importants contrats d'achat d'électricité verte en France \(PPA\), finançant la création de sept nouveaux parcs éoliens et la décarbonation du réseau, Equinix, 2024](#)
⁹⁸ [Data4 signe deux « Power Purchase Agreement » \(PPA\) avec des leaders français de l'énergie renouvelable, Data4, 2024](#)

omniprésente dans l'industrie, même si certains acteurs, encouragés par des organismes d'“autorégulation” comme la Science-Based Targets initiative⁹⁹, tendent à proposer systématiquement un affichage “location-based” (en excluant les certificats de provenance énergétique du calcul d'impact) en plus du “market-based”. Il faut également noter que les recommandations méthodologiques des régulateurs¹⁰⁰ et des spécialistes de l'évaluation d'impacts environnementaux¹⁰¹ vont dans le sens d'un affichage de données d'impacts environnementaux de type “location-based”, sans condition.

Hydrogène et pile à hydrogène

Le marché s'intéresse depuis plus d'une dizaine d'années aux **piles à hydrogène, qui permettraient théoriquement de se passer des énergies fossiles utilisées dans les générateurs de secours**, particulièrement polluantes tant d'un point de vue du potentiel de réchauffement global que de la qualité de l'air, mais aussi potentiellement des batteries nécessaires aux onduleurs¹⁰².

À ce titre, les piles à hydrogène sont **un sujet de recherche actif**. Microsoft et Caterpillar ont par exemple fait la démonstration d'un contrôleur de micro-réseau permettant de faire fonctionner deux systèmes de stockage d'énergie par batterie Cat Power Grid Stabilization PGS 1260, ainsi qu'une pile à combustible à hydrogène de 1,5 MW. Un événement de coupure d'alimentation, de 48 heures a ainsi été simulé, au centre de données de Microsoft à Cheyenne, où une pile à combustible à hydrogène a été intégrée à l'installation électrique pour prendre le relais¹⁰³.

Sean James, directeur principal de la recherche sur les centres de données chez Microsoft en 2021 déclarait « *Les recherches et les résultats de la démonstration de la pile à combustible à hydrogène nous aideront à atteindre notre objectif de devenir neutre carbone d'ici 2030* ». Il est aujourd'hui intéressant de mettre en regard de ces espérances, la trajectoire pour le moins inquiétante de l'entreprise en matière d'émissions de gaz à effet de serre, en grande partie due à la course à l'intelligence artificielle¹⁰⁴.

De manière plus mesurée, certains projets mettent en avant **l'utilisation “d'hydrogène vert”** (comprendre : produit à base de moyens de production d'électricité bas-carbone, puisque l'hydrogène est fabriqué par électrolyse), produit localement en se basant sur des moyens de production à base d'énergies renouvelables. Ceci permettrait de stocker de l'énergie sous forme d'hydrogène lorsque les conditions sont optimales pour les moyens de production d'électricité à base de renouvelables (solaire et éolien), puis de se servir de cette réserve lorsque les conditions sont défavorables pour cette production (la nuit, par exemple). **Atos et HDF Energy annonçaient ainsi travailler sur des piles à hydrogène à forte puissance de 1.5 MW, en 2021.**

Le centre de données EcobioH2, à Avignon, est un exemple de centre de données en production, s'appuyant sur l'hydrogène¹⁰⁵. Il faut cependant souligner la **taille extrêmement réduite du projet**, en comparaison des centres de données existants. En effet, la pile à hydrogène permet de fournir 4kW de puissance, dont 3 kW sont en fait utilisés pour l'alimentation électrique du tiers-lieu LaScierie, présent sur

⁹⁹ [The evolution of scope 2 accounting, target setting and monitoring, SBTi, 2022](#)

¹⁰⁰ Voir le RCP Cloud & Datacenters de l'ADEME

¹⁰¹ [Voir le position-paper d'Etienne Lees-Perasso, expert de l'Analyse de Cycle de Vie de services, de produits numériques et des bâtiments, sur la prise en compte des certificats de garantie d'origine](#)

¹⁰² [Des piles à hydrogène pour alimenter les datacenters, Techniques de l'ingénieur, 2021](#)

¹⁰³ [Caterpillar démontre le potentiel des piles à combustible à hydrogène dans un datacenter de Microsoft, DCMag, 2024](#)

¹⁰⁴ [Microsoft's AI obsession is jeopardizing its climate ambitions, The Verge, 2024](#)

¹⁰⁵ [EcobioH2 : premier datacenter au monde fonctionnant à l'hydrogène, DCMag, 2023](#)

le même site¹⁰⁶. Si cette initiative est sans nul doute socialement importante et à encourager, la démonstration de l’usage d’une pile à hydrogène pour un centre de données d’1 kW n’est pas de nature suffisante pour entrevoir le remplacement généralisé des générateurs de secours dans des centres de données avoisinant au minimum 1 MW de puissance IT. Ce modèle soulève cependant la question de la décentralisation et de la réduction de la demande en données, en puissance de calcul et donc en énergie, vis-à-vis des centres de données.

Exemple plus significatif, Microsoft a démontré la possibilité du remplacement de générateurs de secours au diesel, pour une puissance de 3 MW, ce qui équivaldrait à environ à 10000 serveurs (pour des usages moins intensifs que l’IA générative)¹⁰⁷. Il faut cependant souligner la taille de l’installation nécessaire : 12 m de long et plusieurs mètres de large et de haut, pour stocker les 18 piles à combustibles et les deux conteneurs qui les compartimentent. Notons également que la moyenne de puissance électrique parmi les centres de données français en activité, collectés dans la base de données de la présente étude, est de 7,5 MW.

La technologie basée sur des piles à hydrogène présente un potentiel d’amélioration pour l’alimentation électrique de secours des centres de données. Son applicabilité à grande échelle semble pour le moment se faire attendre. Qui plus est, l’appellation “hydrogène vert” cache un autre sujet : avec quelle production électrique sera produit l’hydrogène nécessaire à cet usage ? La veille effectuée pour cette étude n’a pas permis de mettre en évidence un cas de production d’hydrogène s’appuyant sur des énergies renouvelables, pour un centre de données en production dépassant quelques kW de puissance électrique. L’application effective de cette technologie semble lointaine.

Efficacité des onduleurs

Les onduleurs sont des équipements **essentiels** au sein d’un centre de données. Ils sont le **premier relais des arrivées électriques principales en cas de coupure**, permettant de disposer du temps nécessaire pour le lancement des moyens de production de secours (groupes électrogènes) et ainsi d’éviter toute interruption de service. **Ces équipements sont également réputés relativement peu efficaces.** Une étude de Moura et al., de 2016¹⁰⁸, indique que le choix d’UPS présentant une plus grande efficacité énergétique est un levier significatif de réduction de consommation d’énergie. La Commission Européenne définit un Code de Conduite pour les fabricants, qui donne des objectifs d’efficacité énergétique en fonction de la catégorie et de la capacité de l’équipement¹⁰⁹.

Table 1 – Standard weighted UPS efficiency requirements (%) or “CoC for UPS standard requirements”

Power Range (kW)	Performance Classification		
	VFD	VI	VFI
≥ 0,05 to ≤ 0,3	89,0%	88,0%	84,0%
> 0,3 to ≤ 3,5	92,0%	91,0%	86,0%
> 3,5 to ≤ 10	93,7%	92,4%	87,5%
> 10 to ≤ 200	96,0%	93,0%	90,0%
> 200	97,0%	94,0%	92,0%

¹⁰⁶ Ibid.

¹⁰⁷ [Microsoft montre qu’on peut alimenter un datacenter avec des piles à hydrogène, ICT Journal, 2022](#)

¹⁰⁸ [Energy savings potential of uninterruptible power supplies in European Union, Moura et al., 2016](#)

¹⁰⁹ [European Code Of Conduct for UPS, 2021](#)

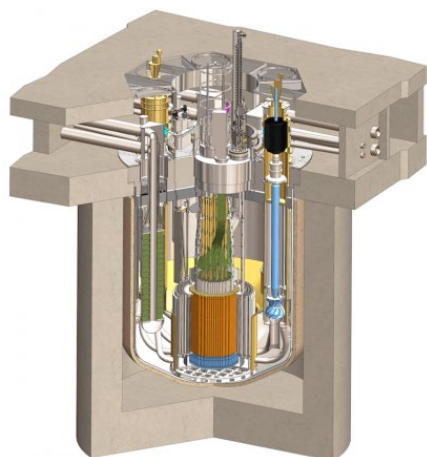
Table 2 – Elite weighted UPS efficiency requirements (%) or “CoC for UPS elite requirements”

Power Range (kW)	Performance Classification		
	VFD	VI	VFI
≥ 0,05 to ≤ 0,3	91,0%	90,0%	85,5%
> 0,3 to ≤ 3,5	94,0%	93,0%	87,5%
> 3,5 to ≤ 10	95,7%	94,4%	90,0%
> 10 to ≤ 200	97,0%	95,0%	91,5%
> 200	98,0%	96,0%	93,5%

Certains modèles proposent un mode de fonctionnement optionnel, permettant d’augmenter leur efficacité énergétique, au prix d’une fiabilité amoindrie, souvent appelé “Eco-mode”, mais les fabricants mettent en garde leurs clients sur la balance bénéfice-risque de ce type d’usage¹¹⁰, ce qui ne semble pas plaider en la faveur d’une adoption généralisée.

¹¹⁰ [Eco-mode: Benefits and Risks of Energy-saving Modes of UPS Operation, Schneider Electric](#)

Petit Réacteur Modulaire (PRM) ou Small Modular Reactor (SMR)



Source: [Le Monde informatique](#)

Face à la demande électrique croissante et à une anticipation des difficultés des producteurs d'électricité à suivre cette demande - et/ou à prioriser les centres de données sur d'autres usages - de nombreux opérateurs, comme OVH¹¹¹, Microsoft¹¹², Equinix¹¹³, envisagent d'investir dans des SMR sur leurs sites de centres de données pour assurer leur autonomie énergétique. Ces réflexions semblent être au stade de la recherche et du développement de prototypes. Equinix envisage leur possible mise en place d'ici 5 à 7 ans, et souligne l'intérêt de pouvoir remettre le delta d'électricité produite mais non consommée dans le réseau pour d'autres utilisateurs.

Selon Thomas A. Hemphill¹¹⁴, dans un article dédié aux défis du déploiement des SMR pour des centres de données dédiés à l'IA, les avantages des SMR sont la capacité :

- "à répondre à la croissance de la demande d'électricité (une considération importante pour les centres de données à très grande échelle - et potentiellement d'informatique quantique - de l'IA au cours de la prochaine décennie),
- à réaffecter les sites de combustibles fossiles désaffectés
- et à s'intégrer de manière transparente aux sources d'énergie renouvelables intermittentes et aux systèmes énergétiques, aux micro-réseaux résilients et hautement fiables, autosuffisants et, sur le plan environnemental, les moins générateurs de tonnes d'équivalent CO₂ par gigawatt-heure (13 tonnes). (selon l'OCDE et l'AIEA)"

Les SMR apparaissent avant tout comme une stratégie sectorielle sérieusement à l'étude pour prendre la main sur **une production importante et autonome** d'électricité pour les centres de données. Cependant, les débats publics éludent la problématique de la production de nouveaux déchets nucléaires. De même, le deuxième avantage listé par Thomas A. Hemphill, à savoir la réaffectation des sites de combustibles fossiles désaffectés, pose la question de la dépollution de ces sites avant réaffectation. **La faisabilité à une échéance courte est également une question en suspens**, si l'on en croit les récentes informations

¹¹¹ Entretien, octobre 2024

¹¹² <https://www.lemondeinformatique.fr/actualites/lire-des-mini-reacteurs-nucleaires-pour-alimenter-les-datacenters-de-microsoft-91710.html>

¹¹³ Entretien, novembre 2024

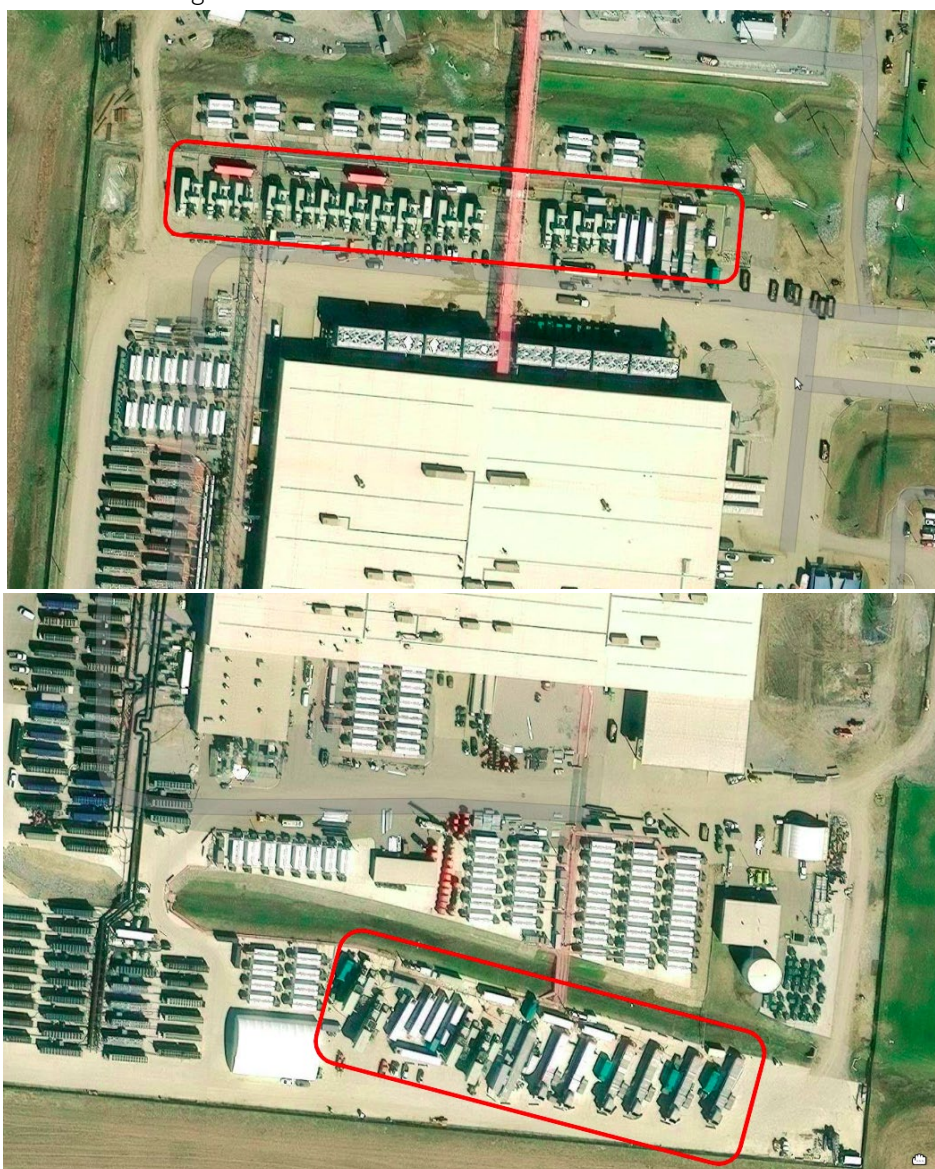
¹¹⁴ Thomas A Hemphill, US AI data centers and deployment challenges for small modular reactors: proposed regulatory policy recommendations, *Science and Public Policy*, Volume 51, Issue 5, October 2024, Pages 999–1003

concernant le rapport d'audit piloté par le haut-commissaire à l'énergie atomique, visant les startups françaises lancées dans la course aux micro-réacteurs¹¹⁵, telle que la startup Naarea¹¹⁶.

Retour à la production locale d'électricité à partir d'énergies fossiles

Pour les mêmes raisons que les investissements dans les SMR augmentent, les grands acteurs de l'IA générative et les filières économiques associées, ont opéré un changement de cap pour permettre le développement des grands centres de données malgré les délais de raccordement au réseau électrique et les contraintes de production et d'approvisionnement de l'électricité.

Les centres de données emblématiques Stargate I d'OpenAI et Colossus I d'xAI, en sont un exemple, comme le montrent les images ci-dessous.



Images agrandies du centre de données Colossus I d'xAI (> 250 MW), à Memphis dans le Tennessee. En rouge encadré : les génératrices d'électricité au méthane (CH₄)¹¹⁷, en densité et avec des capacités bien plus

¹¹⁵ EXCLUSIF. L'audit secret qui crucifie la plupart des start-up françaises du nucléaire, *Le point*, 2024

¹¹⁶ Naarea "Nuclear Abundant Affordable Resourceful Energy for AI"

¹¹⁷ SolarTurbines, filiale de Caterpillar

importantes que les génératrices au diesel classiquement présentes sur les sites de centres de données.
(source: "image captured on Mar 10, 2025, as shown in the 2020-25 version of the World Imagery map")



Image agrandie d'une partie des centres de données du site Stargate I d'Open AI (4,5GW à terme), à Abilene, sous le curseur rouge, les génératrices au gaz, donc les cheminées sont apparentes.¹¹⁹

Très récemment, les fabricants de génératrices ont commencé à s'orienter vers le reconditionnement de moteurs d'avion, pour fournir les opérateurs de centres de données dans des délais raccourcis, par rapport à des turbines neuves¹²⁰.

1.4.2. Évolutions matériel IT

Évolutions actuelles

Les évolutions du matériel IT présent en centre de données sont sujet à des tendances inverses et en apparence contradictoires. La loi de Moore est une projection tirée d'une hypothèse de Gordon Moore, cofondateur d'Intel, concernant une hypothèse de doublement du nombre de transistors par circuit intégré, sur une même surface, tous les dix-huit mois, à coût unitaire réduit¹²¹. Ce doublement récurrent du nombre de semi-conducteurs des microprocesseurs est de plus en plus débattu,¹²² bien qu'il reste relativement perceptible.

¹¹⁸ [Esri, World Imagery Wayback](#)

¹¹⁹ [Stargate I, OpenAI, via Google Maps](#)

¹²⁰ [ProEnergy repurposes jet engines to power data centers amid gas turbine shortages - report](#)

¹²¹ [La phase G : les GPU et les IA génératives comme nouvelle phase de l'histoire environnementale de la numérisation ? \(Partie 1\), Gauthier Roussilhe](#)

¹²² [Jensen Huang says Moore's law is dead. Not quite yet](#)

Our World
in Data

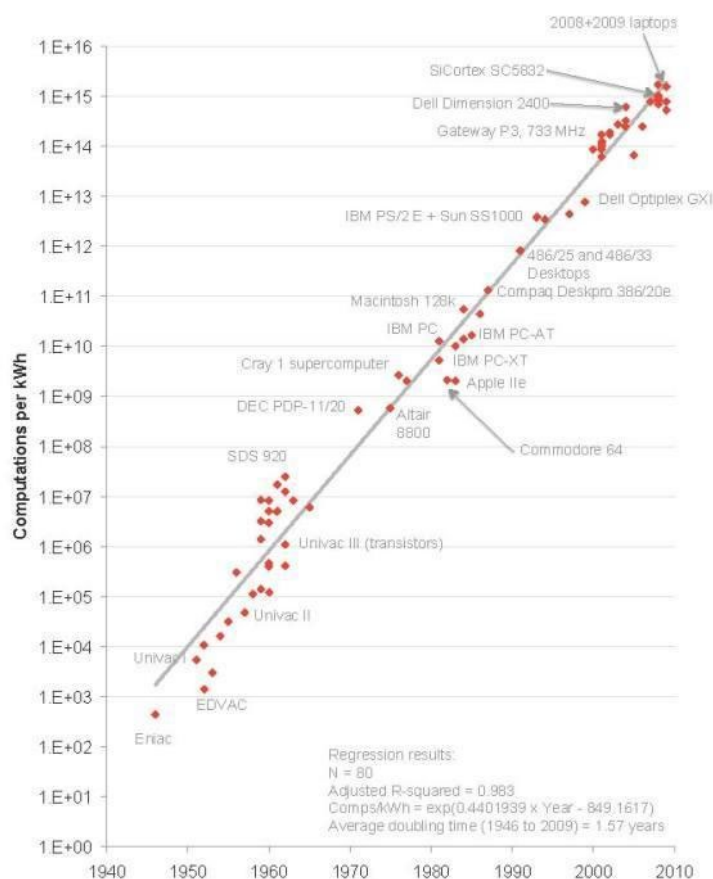
in Data



Représentation graphique de la loi de Moore par Our World in Data

D'un point de vue énergétique et environnemental, l'utilisation qui est couramment faite de la loi de Moore pose question, puisqu'elle semble légitimer la seule innovation comme axe de réduction des impacts environnementaux, par le mythe de la dématérialisation.

La loi de Koomey, énonce quant à elle que le nombre de calculs effectués par joule (quantité d'énergie) dépensé, double tous les 1.57 ans. Cette loi a également œuvré pour une perception généralisée du secteur, que l'innovation permet une moindre consommation énergétique.



Graphique illustrant la loi de Koomey, de 1940 à 2010

Cette loi a été énoncée pour la dernière fois en 2009¹²³/2011¹²⁴ et a été réévaluée comme chutant à un doublement du nombre de calculs par joule tous les 2.29 ans, pour la période 2008-2023 (Prieto and al., 2024¹²⁵). De plus, les projections effectuées par Penney et Zhang, de la Carnegie Mellon University, en 2022, plaident pour un **ralentissement drastique de la loi de Koomey jusqu'à 2050, qui passerait à un doublement du nombre de calculs par joule tous les 12 ans.**¹²⁶

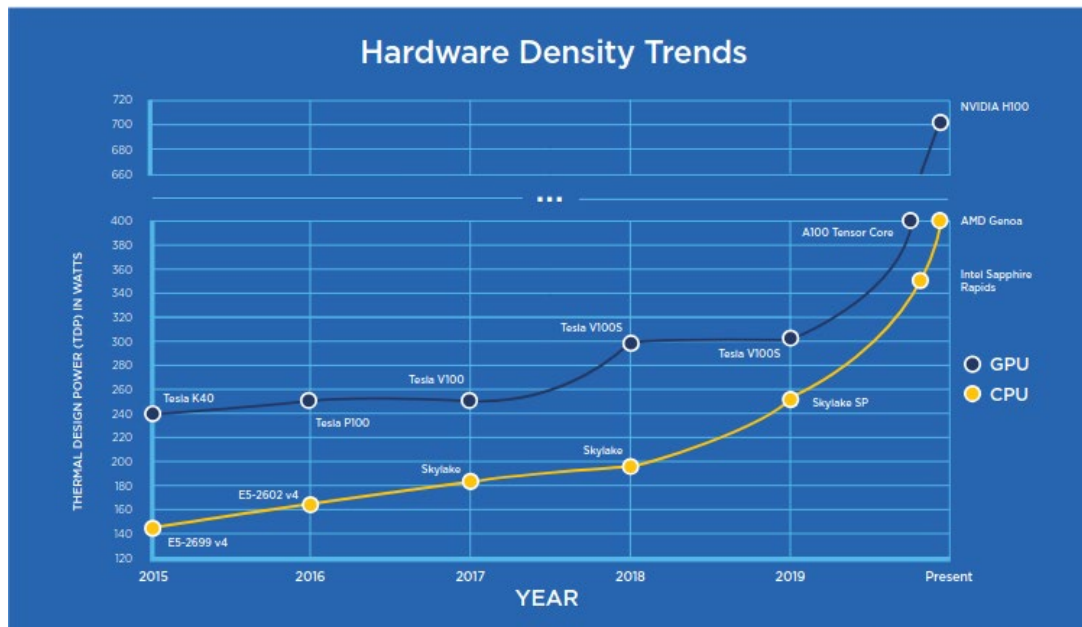
En parallèle, la puissance électrique demandée par une carte de calcul, qu'elle soit un CPU (Central Processing Unit) ou un GPU ([Graphical Processing Unit](#)), augmente drastiquement avec les années, avec une accélération en apparence durable à partir de 2019.

¹²³ [Assessing trends in the electrical efficiency of computation over time, Koomey and al., 2009](#)

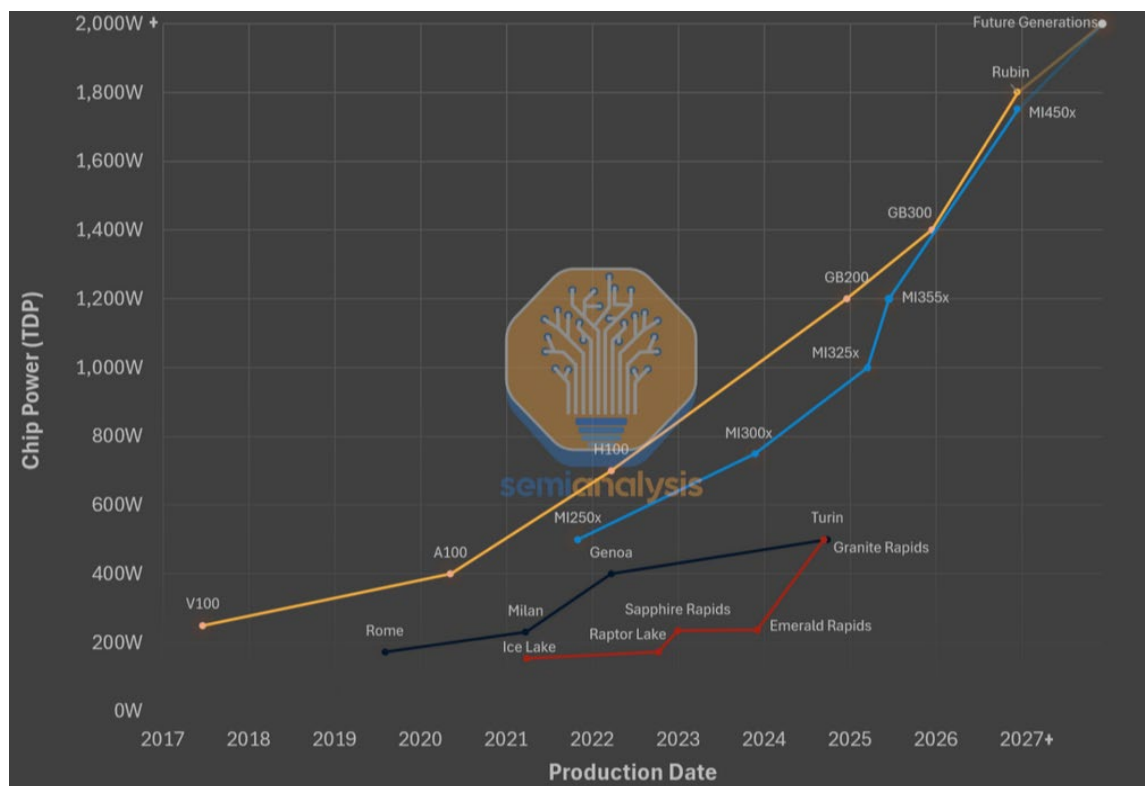
¹²⁴ [Implications of Historical Trends in the Electrical Efficiency of Computing, Koomey and al., 2011](#)

¹²⁵ [Evolution of computing energy efficiency: Koomey's law revisited, Prieto and al., 2024](#)

¹²⁶ [Future Data Center Energy Efficiency and Consumption, Penney, Zhang, 2022](#)



Graphique illustrant l'évolution des TDP (Thermal Design Power), à travers les générations de CPU et de GPU, par GRC ¹²⁷



Courbes des TDP des GPUs fabriqués par Nvidia (en jaune), des GPUs fabriqués par AMD (en bleu), des CPUs fabriqués par AMD (en noir) et par Intel (en rouge), par SemiAnalysis ¹²⁸

¹²⁷ [The Future of Immersion Cooling: The Path to Cooling 1000W Chips, and Beyond!](#), GRC, 2023

¹²⁸ [Datacenter Anatomy Part 2 – Cooling Systems, Semi Analysis](#)

Si une augmentation du nombre de calculs par Joule est certainement toujours un élément constatable aujourd'hui, **cet indicateur relatif à l'échelle du semi-conducteur ne permet pas de rendre compte de l'augmentation bien réelle et soutenue de la consommation d'énergie des composants, ainsi que celle des serveurs informatiques** (dont la configuration moyenne à l'installation évolue chaque année vers plus de capacité de calcul, mémoire ou de stockage).

Le paradoxe de Jevons¹²⁹ ou effet rebond, s'applique de manière manifeste, de la même manière aux semi-conducteurs dans les matériels informatiques. En effet, le gain en efficacité énergétique des capacités de calculs semble compensé par l'augmentation des capacités des composants et leur volume de production.

CPU : Les architectures RISC de la société ARM

Les CPU les plus répandus en centres de données sont basés sur des architectures x86 (aussi appelée x86_64 depuis que le 64 bit a remplacé le 32 bit sur le marché), dont le processeur Intel 8086 était le premier exemple.

Les processeurs x86 sont initialement issus d'une conception CISC, pour *Complex Instruction Set Computer*, c'est à dire qu'ils présentent un grand nombre d'instructions, ce qui les rend plus complexes que les processeurs issus d'une conception RISC, pour *Reduced Instruction Set Computer*, dont sont issus les processeurs ARM. Ces processeurs présenteraient une consommation d'énergie moindre pour une même capacité de calcul.¹³⁰

Les architectures ARM gagnent des parts de marché. Les fournisseurs de cloud américains (AWS, GCP et Azure), ainsi que français (Scaleway), proposent déjà des charges de travail basées sur ARM à leurs clients. **Si l'on considère qu'elle pourrait également avoir un rôle pour l'IA générative,¹³¹ cette adoption est à surveiller pour estimer les tendances futures de consommation d'énergie des centres de données.**

GPU : Graphical Processing Unit

Un GPU, pour *Graphics Processing Unit*, est un processeur graphique, présentant classiquement un plus grand nombre de cœurs de calcul qu'un CPU, optimisé pour permettre un grand nombre de traitements en parallèle (ce qui est adapté aux traitements d'images, notamment). Bien que ces composants soient principalement utilisés pour le multimédia et le jeux vidéo sur le marché grand public, leur usage a progressivement évolué pour le minage de crypto-monnaies et pour l'entraînement de modèles d'IA.

L'évolution matérielle récente la plus significative est celle des GPUs destinés à soutenir le boom de l'IA générative. La puissance de calcul et le nombre de semi-conducteurs continuent de croître¹³², la puissance électrique maximale demandée par GPU augmente de manière considérable et ce à chaque nouvelle itération¹³³ et la densité énergétique par unité de surface gravée s'est fortement accrue depuis le passage au noeud technologique de 7 nm, ce qui explique en partie, l'augmentation significative de la densité énergétique des GPUs récents¹³⁴.

¹²⁹ [Paradoxe de Jevons](#)

¹³⁰ [ARM vs x86, Versus, 2024](#)

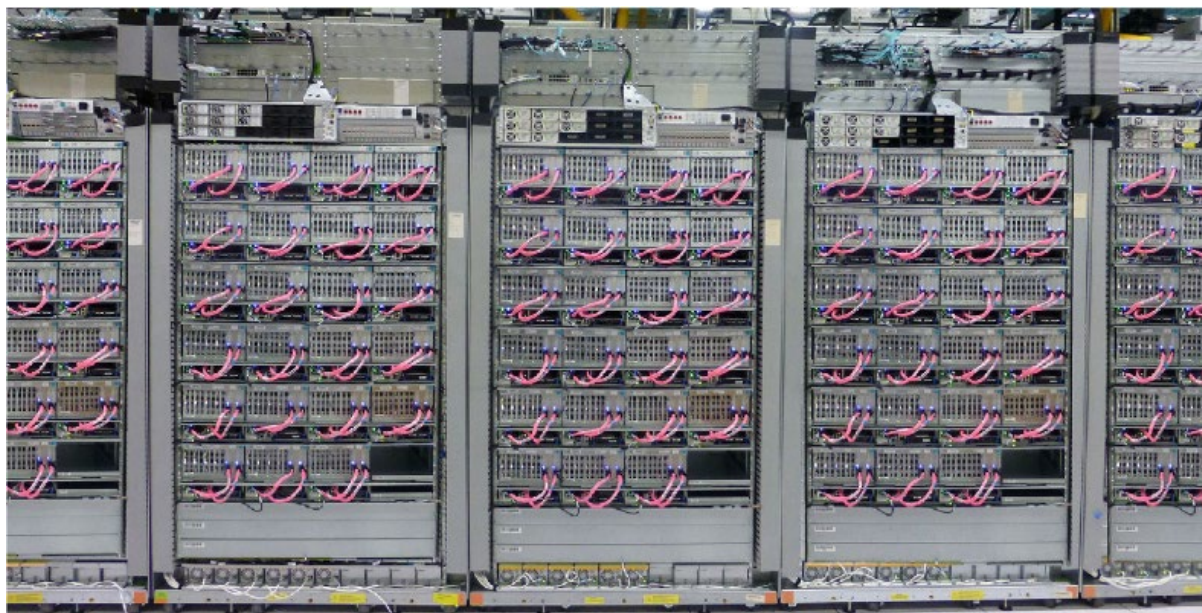
¹³¹ [Performance Per Watt is the New Moore's Law, Aitken, Arm.com, 2021](#)

¹³² <https://doi.org/10.48550/arXiv.1911.11313>

¹³³ <http://arxiv.org/pdf/2311.02651>

¹³⁴ <https://arxiv.org/pdf/1911.11313>

Le nombre de GPUs vendus et mis en production est également à considérer pour comprendre les tendances actuelles de consommation d'énergie induites par l'IA. Dans les centres de données spécialisés et adaptés à un usage intensif de l'IA, la densité énergétique par baie augmente considérablement par rapport aux moyennes du marché (voir [Augmentation de l'intensité énergétique par Rack](#)).



Des baies hébergeant de nombreux GPUs, interconnectés par le réseau, dans un centre de données de Google.

Source: The datacenter as a computer, Third edition, Barroso, Hölzle, Ranganathan¹³⁵

Sur l'image ci-dessus, au sein d'un centre de données de Google, on peut voir des baies pleines de serveurs spécialisés pour la mise à disposition de puissance de calcul par GPU et plus spécifiquement l'entraînement de modèles d'IA. Chaque serveur est composé logiquement d'un CPU et d'une carte mère, ainsi que d'un ou plusieurs GPUs attachés à la carte en PCI-express. Chaque GPU au sein d'un même serveur est connecté aux autres avec un réseau haute performance de type NVlink, ce qui permet de répondre aux contraintes de performance et de latence lors de l'entraînement de modèles de *deep-learning*.¹³⁶

Au moment de notre étude, la carte faisant référence, dans le secteur du machine learning et du deep learning, est la H100 d'Nvidia. Le Thermal Design Power ou TDP, valeur de puissance souvent prise comme référence pour un matériel¹³⁷, est de 700 W. Cependant, la carte ne fonctionne pas seule : si l'on considère le serveur nécessaire à faire fonctionner cette carte, comprenant également des CPUs, ainsi qu'une côte part de la puissance nécessaire pour l'environnement technique, la puissance maximum demandée pour le fonctionnement d'une seule H100 peut atteindre jusqu'à 1700 W¹³⁸, ce qui est proche de la puissance demandée par un four à chaleur tournante domestique de gamme intermédiaire. Pour illustrer la taille des déploiements nécessaires aux grands modèles de LLM, Meta a utilisé un cluster de 16000 cartes H100 pour entraîner le modèle Llama 3.1 405B¹³⁹, qui n'était qu'une itération de cette famille de modèles. Si

¹³⁵ [The datacenter as a computer, Third Edition, 2022 - Barroso, Hölzle, Ranganathan](#)

¹³⁶ Voir à ce titre [les prévisions de croissance annuelle des revenus de Nvidia](#)

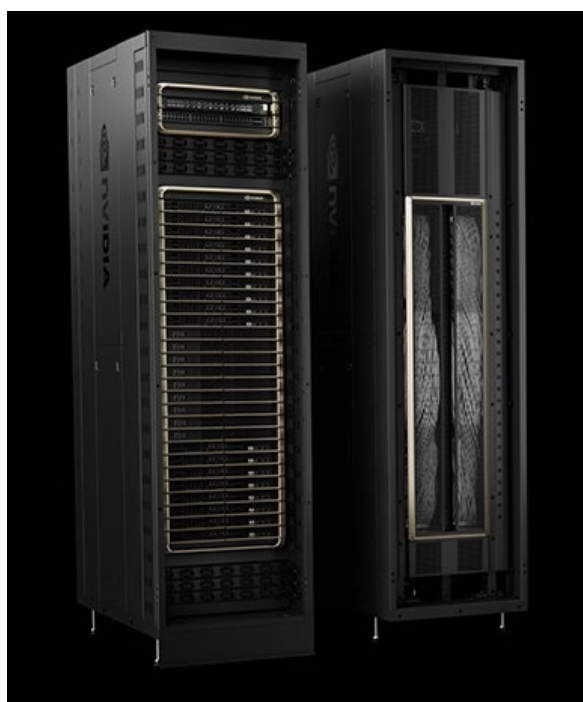
¹³⁷ Bien qu'il s'agisse en théorie d'une puissance thermique que la carte peut supporter/dissiper, voir [Hard du Hard • TDP et consommation, entre technologie et entourage, Comptoir Hardware, 2020](#)

¹³⁸ [Can AI scaling continue through, Epoch AI, 2024](#)

¹³⁹ [Introducing Llama 3.1: Our most capable models to date, Meta, 2024](#). Nous n'avons cependant pas d'information quant à la durée de cet entraînement. Par ailleurs, ceci ne représente qu'une itération parmi de nombreuses, le modèle ayant eu de nouvelles versions depuis.

L'on reprend l'estimation précédente de la puissance maximale environnée par carte H100, on obtient une puissance raccordée de 27 MW pour ce cluster.

Le catalogue d'Nvidia évolue continuellement. Au moment de cette étude, la dernière annonce en date concerne les cartes GB200. Celles-ci posent cependant des problèmes plus structurels à leur installation, en particulier pour des raisons de puissance électrique nécessaire et de refroidissement. Pour simplifier le déploiement, Nvidia propose maintenant des offres couvrant une ou plusieurs baies entières, comprenant un certain nombre de cartes, des serveurs pour les héberger, le réseau et les équipements nécessaires à l'alimentation électrique de l'ensemble. L'offre GB200 NVL72 consiste en une baie complète de ce type, pour une puissance de 120 kW par baie. La GB200 NVL36x2 est un ensemble de deux baies avec une puissance par baie moindre, mais qui cumule une puissance maximale de 132 kW. En comparaison, les baies refroidies par air, déployées avec une forte densité de cartes H100, ne dépassent pas 40 kW par baie¹⁴⁰. Modèle plus récent, la NVL576 demande jusqu'à 600 kW de puissance électrique.



Une baie GB200 NVL72, hébergeant 72 GB200 Grace Blackwell Superchip, chacune comprenant 1 CPU Grace et 2 GPU Blackwell, Source : [Nvidia](#)

Ces nouveaux prérequis de puissance électrique s'accompagnent de pré-requis concernant le refroidissement. Une puissance de plus de 40 kW par baie rend le refroidissement par air insuffisant. La stratégie la plus couramment employée est celle du Direct Liquid to Chip Cooling, que nous détaillons dans la section [Techniques de refroidissement](#).

Le rythme d'apparition de nouvelles architectures de GPU et le volume de déploiement de ces cartes est important, et il n'est certainement qu'une partie des volumes totaux de déploiement de GPUs pour des usages de machine learning, d'IAgen ou d'autres usages.

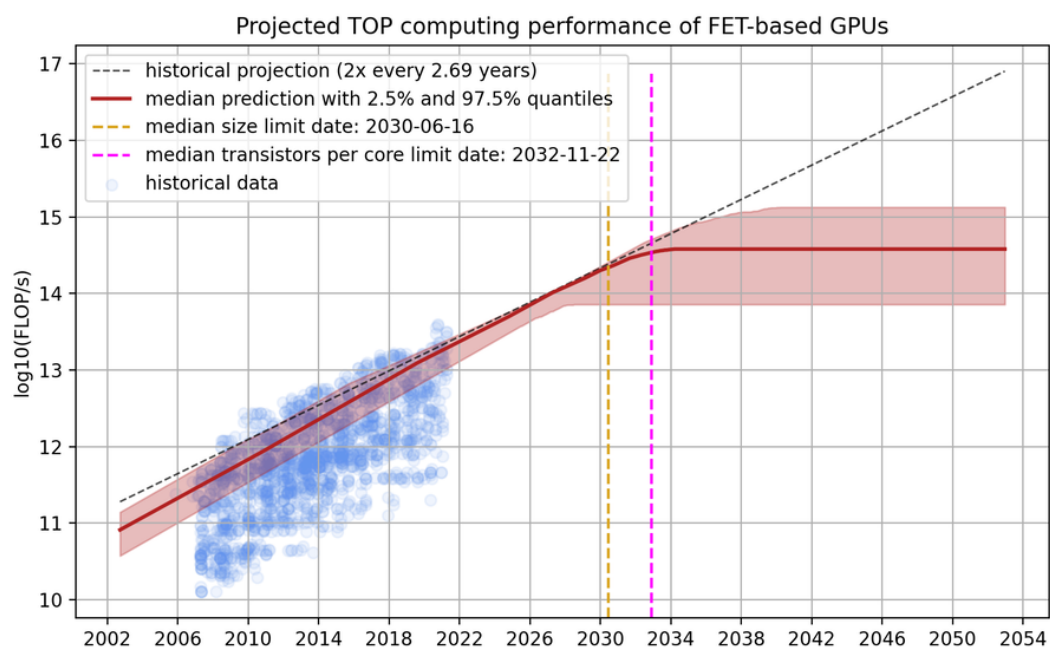
Certains acteurs proposent d'autres configurations, conçues en interne. Ainsi, OVHCloud propose parmi ses serveurs des modules conçus en interne (appelés "sandwich de GPU") composés de deux GPU Nvidia

¹⁴⁰ [GB200 Hardware Architecture – Component Supply Chain & BOM, Semi Analysis, 2024](#)

RTX 4500 (non spécialisés pour l'IA), chaque serveur pouvant accueillir quatre modules, soit 8 GPUs en tout par machine.¹⁴¹ Cette stratégie s'insère dans la volonté d'OVHCloud de proposer des offres à moindre coût pour des projets d'IA plus modestes ne nécessitant pas les dernières architectures de GPU en date.¹⁴²

L'utilisation intensive de GPUs s'accompagne de problématiques énergétiques et de refroidissement. Ces problématiques peuvent également remettre en question l'architecture du bâtiment si l'usage de ces cartes est majoritaire dans le centre de données. Le dimensionnement de leur usage est variable, entre les acteurs possédant des centres de données hyperscale, qui conçoivent des bâtiments exclusivement pour cet usage et des acteurs de la colocation, pour qui l'IA ne représente encore qu'une part relativement faible des demandes client (5 à 10% d'après nos interviews¹⁴³). Dans le premier cas, toute l'infrastructure électrique et de refroidissement doivent être dimensionnées en conséquence¹⁴⁴, dans l'autre cas, des solutions existent pour refroidir une sous-partie du centre de donnée ou d'une salle IT, pour répondre à un besoin localisé d'IA. Nous détaillons ces mécanismes dans la section [Techniques de refroidissement](#).

La continuité de l'augmentation des performances des nouveaux modèles de GPU et de la consommation d'énergie associée est un point crucial. **EpochAI estime par exemple que l'augmentation des performances des nouvelles cartes atteindra un plateau, entre 2027 et 2035¹⁴⁵.**



Projection des performances des nouvelles générations de GPU, par EpochAI

¹⁴¹ [Immersion dans un Datacenter avec OVHCloud, chaîne Youtube Développeur Libre](#)

¹⁴² Interview de l'acteur OVHCloud, pour l'étude

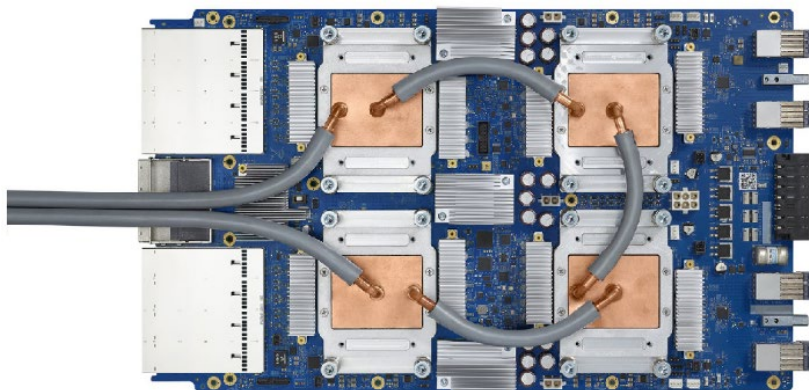
¹⁴³ Interview de l'acteur Equinix

¹⁴⁴ [Schneider Electric annonce une nouvelle conception de référence de datacenter d'IA avec Nvidia, DCMag, 2024](#)

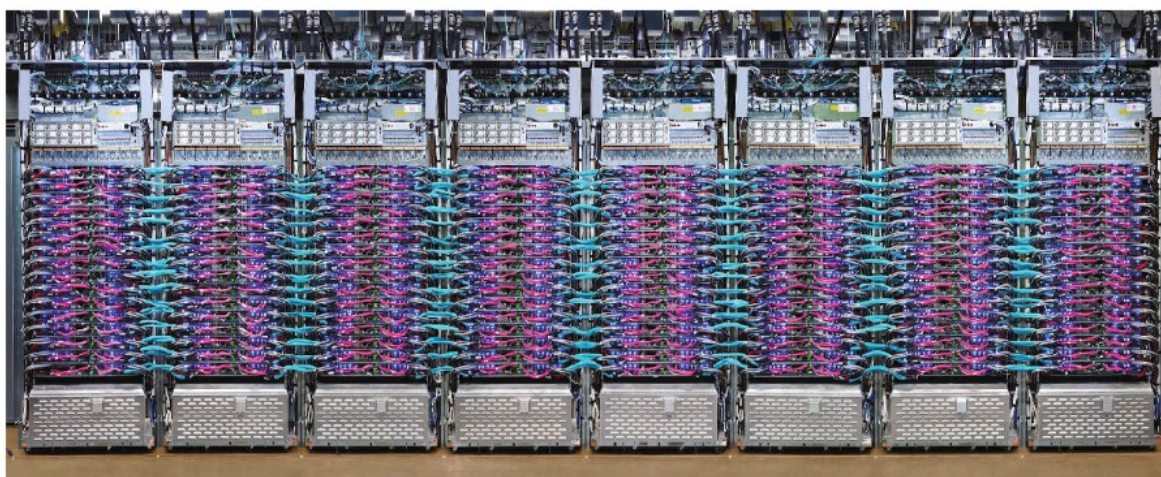
¹⁴⁵ [Predicting GPU Performance, EpochAI, 2022](#)

TPU

Les GPUs sont performants pour les usages de type machine learning mais restent du matériel générique. Google utilise également des TPUs, pour Tensor Processing Unit, à la fois pour l'entraînement des modèles et pour l'inférence (leur usage après entraînement)¹⁴⁶. Ces puces sont plus proches des ASICs (Application Specific Integrated Circuit) que des GPUs (Graphical Processing Unit), c'est à dire qu'elles sont fabriquées avec pour objectif un usage plus précis, sont moins complexes, moins polyvalentes, mais plus performantes pour cet usage¹⁴⁷.



Une carte portant des TPUv3, dans un centre de donnée de Google, Source : [The datacenter as a computer, 3.2.2 TPUs, Barroso, Hölzle, Ranganathan](#)



8 baies hébergeant des TPUv3, dans un centre de données de Google, Source : [The datacenter as a computer, 3.2.2 TPUs, Barroso, Hölzle, Ranganathan](#)

Au moment de la rédaction de cette étude, Google en est à sa troisième génération de TPUs. Si l'un des objectifs affichés par Google pour l'usage des TPUs est l'efficacité énergétique, il est pour le moment difficile de parler des gains effectifs de ce type de technologie, déployées à l'échelle, par rapport aux GPUs, d'un point de vue énergétique, du fait du manque de données et d'études sur le sujet.

Évolutions potentielles d'avant-garde

La spintronique, l'informatique quantique, l'ordinateur photonique ou encore la mémoire biologique restent à ce jour des champs de recherche plus ou moins avancés et encore peu répliqués sur des

¹⁴⁶ [The datacenter as a computer, 3.2.2 TPUs, Barroso, Hölzle, Ranganathan](#)

¹⁴⁷ [Understanding ASICs for Network Engineers, Pete Lumbis, Packet Pushers, 2018](#)

volumes importants. Cependant, les deux premiers sont annoncés comme prometteurs^{148 149} pour augmenter drastiquement l'efficacité énergétique¹⁵⁰ des systèmes de calcul et de stockage.

Évolutions de l'environnement technique

Évolution bâtiment : les centres de données modulaires¹⁵¹

Contrairement aux centres de données traditionnels, construits sur site, les **centres de données modulaires sont construits hors-site et livrés par pièces**, voire presque prêts à l'emploi dans le cas de conteneurs. On retrouve plusieurs types de modules dans la presse spécialisée¹⁵² :

- **Data Hall préfabriqué** : une salle entière livrée en pièces, qui peut être reliée à un centre de données existant sur place, permettant son extension. Ce type de module est généralement destiné à être installé en intérieur et permet de déployer de nouvelles salles IT plus rapidement et plus simplement.



Une "modulo box" de Celeste une fois installée¹⁵³

¹⁴⁸ <https://www.cairn.info/revue-responsabilite-et-environnement-2023-2-page-30.htm?ref=doi>

¹⁴⁹ Nellow veut provoquer "une révolution en microélectronique", Le Dauphiné, 2024

¹⁵⁰ L'avantage de l'ordinateur quantique est-il à chercher du côté de sa consommation énergétique ? Fellous-Asiani, 2022

¹⁵¹ Schneider Electric agrandit son usine de Data centres modulaires, DCMag, 2024

¹⁵² Qu'est-ce qu'un datacenter modulaire ? Pure Storage

¹⁵³ Modulo box, Celeste



Une "modulo box" de Celeste, à la livraison

- **Conteneur** : tout le nécessaire à l'hébergement d'équipements (alimentation électrique, refroidissement, ...) est fourni dans un conteneur ressemblant aux conteneurs de transport de marchandises.¹⁵⁴ On retrouve notamment ces configurations pour le mining de Blockchain.¹⁵⁵



Un centre de données modulaire, dans un conteneur, par Huawei, Source : Datacenter Dynamics

- **Enceinte personnalisée** : destiné à être installé en extérieur, permet une extension d'un centre de données existant sans étendre le bâtiment.

Certaines sources évoquent également des centres de données modulaires en béton.¹⁵⁶

Ce type d'approche semble être en essor, du fait des demandes croissantes d'hébergement et des délais nécessaires à la construction d'un centre de données traditionnel. Schneider Electric a par exemple agrandi son plus grand centre de fabrication de centres de données modulaires à Barcelone, fin 2024, faisant passer le bâtiment de 7000m² à 12000m².¹⁵⁷

¹⁵⁴ [Huawei launches 40ft and 20ft data center containers, Datacenter Dynamics, 2014](#)

¹⁵⁵ https://www.sciencefriday.com/wp-content/uploads/2024/07/Bitcoin_PR.jpg

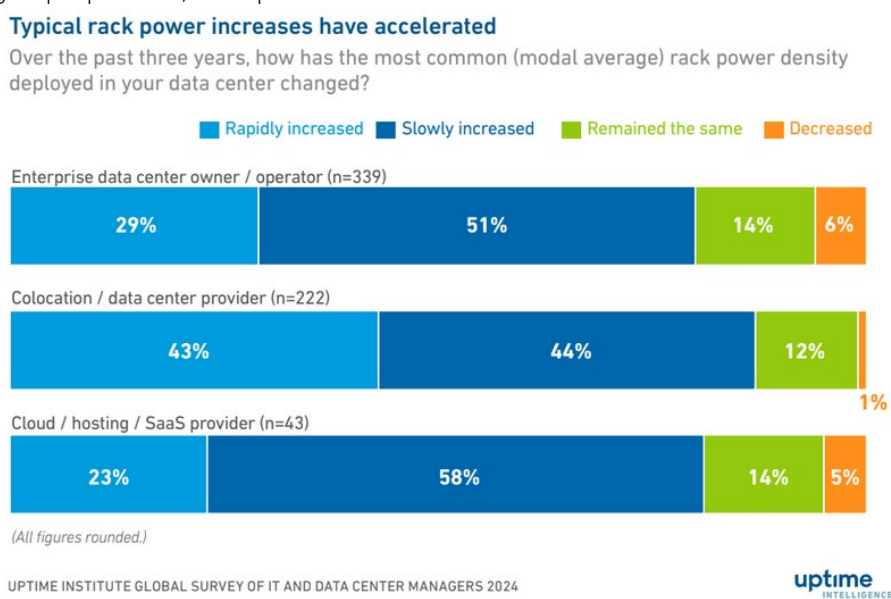
¹⁵⁶ [Datacenters modulaires et en conteneurs, APL Datacenter](#)

¹⁵⁷ [Schneider Electric : 5 000 m² supplémentaires d'espace de production dédié, DCMag, 2024](#)

Augmentation de l'intensité énergétique par Rack

L'augmentation significative des usages de l'IA et en particulier de l'IA générative redéfinit les besoins de puissance électrique disponible dans une même baie informatique (rack). Ainsi on peut trouver des estimations de la puissance demandée par les futures baies orientées IA de Nvidia, *Blackwell Ultra* et *Rubin AI*, jusqu'à 900 kW par baie,¹⁵⁸ là où les modèles les plus demandeurs actuels atteignent 130 kW comme décrit dans la section [GPU](#), soit une multiplication par près de 7 de la puissance par baie entre les configurations les plus puissantes actuelles et celles des futures baies prévues.

La puissance électrique disponible dans une baie complète pleine hauteur (42 unités) démarre classiquement autour de 2 ou 3 kW. Cette puissance disponible peut augmenter de manière significative en fonction du type de centre de données et de l'offre d'hébergement choisie. L'Uptime Institute a tenté de dégager des moyennes de la densité énergétique par baie à travers son sondage Global Data Center Survey 2024. On y trouve d'abord un rapport d'opinion sur la vitesse d'augmentation des besoins en densité énergétique par baie, de la part des acteurs du secteur :



Opinion des acteurs du centre de données sur la vitesse d'augmentation des besoins en densité énergétique par baie, Uptime Institute, 2024

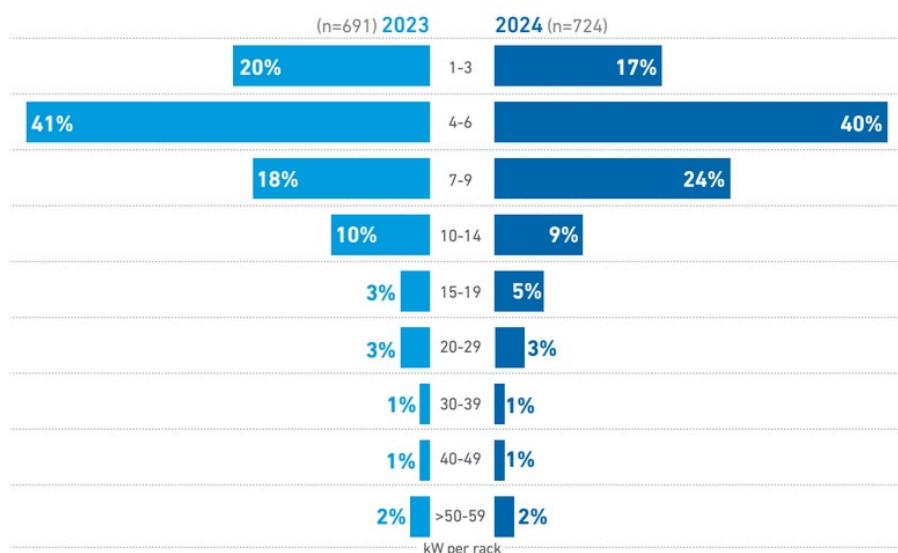
Malgré quelques nuances dans les réponses positives, le constat d'une augmentation significative apparaît comme étant majoritaire, la question portant sur la période 2021-2024.

La répartition des densités comparées entre 2023 et 2024, semble le confirmer :

¹⁵⁸ <https://www.gamingdeputy.com/fr/densite-de-rack-de-plus-de-1-000-kw-avec-les-gpu-rubin-ultra-ai-de-nouvelle-generation-de-nvidia/>

Rack densities of 7 kW to 9 kW have become more common

What is the most common (modal average) server rack density deployed in your data center?



UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2024

uptime
INTELLIGENCE

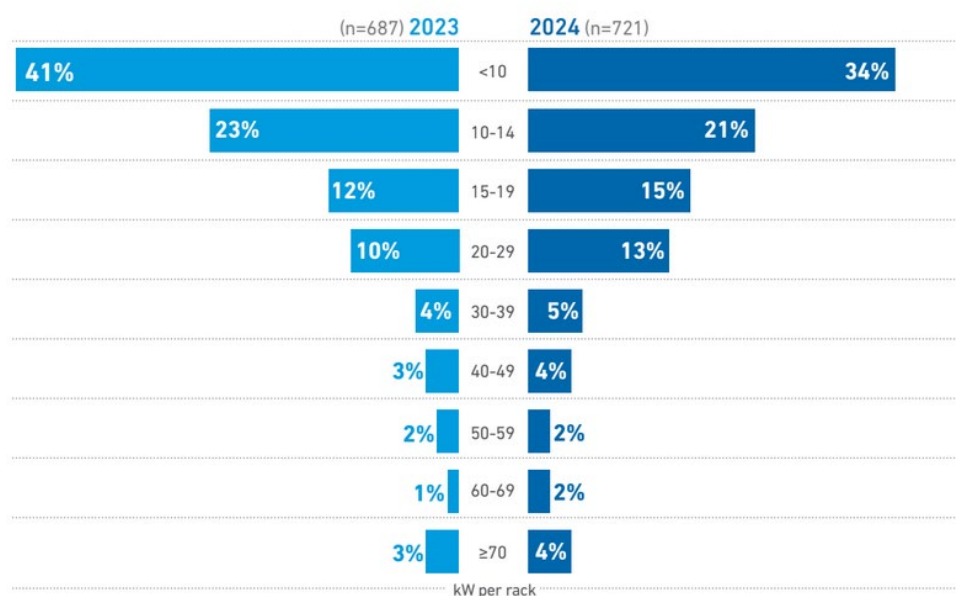
Comparatif des densités énergétiques par baie, entre 2023 et 2024, issues du sondage Global Data Centre Survey, Uptime Institute, 2024

On constate en effet une augmentation significative des réponses dans la catégorie 7 à 9 kW par baie, passant de 18 à 24 %, une baisse légère de la part des réponses dans la catégorie 1 à 3 kW par baie (de 20 à 17%) et une augmentation légère de la part des réponses dans la catégorie 15 à 19 kW par baie (de 3 à 5 %).

Il s'agit ici de valeurs moyennes, l'Uptime Institute ayant également fait l'exercice de demander les densités énergétiques maximales par baie à ses répondants :

Highest densities in the 15 kW to 29 kW range have become more common

What is the highest server rack density deployed in your data center?



UPTIME INSTITUTE GLOBAL DATA CENTER SURVEY 2024

uptime
INTELLIGENCE

On constate ici une augmentation sur presque toutes les catégories au-delà de 15 kW par baie jusqu'à 70 kW par baie, avec une stagnation seulement pour la catégorie 50 à 59 kW par baie.

L'augmentation moyenne des densités énergétiques par baie semble donc bien réelle, même si l'utilisation des baies orientées IA comme la NVL72 de Nvidia atteignant 120 kW, ou plus récemment la NVL576 (également Nvidia) atteignant 600 kW, par baie, reste l'exception plutôt que la règle. **L'utilisation de GPUs dans une moindre densité** (notamment des modèles de type A100 ou H100) et **dans des baies traditionnelles est en revanche certainement l'une des composantes de cette augmentation**.¹⁵⁹

L'augmentation de la densité énergétique par baie s'accompagne forcément d'un besoin de refroidissement accru. L'une des tendances envisagées jusque-là pour répondre à ce besoin semble avoir été d'espacer davantage les baies entre elles, ce qui a un effet sur la taille des bâtiments.¹⁶⁰

De manière plus marquée, le Direct Liquid Cooling dans ses versions To Chip et Door Cooling, semblent les tendances les plus actuelles pour répondre à ce besoin de refroidissement plus intense (voir [Techniques de refroidissement](#)).

Techniques de refroidissement

L'eau a quatre fonctions, deux spécifiquement pour les salles informatiques :

1. **En amont**, l'eau est utilisée pour produire l'énergie et l'électricité utilisée par les centres de données.
2. **Pendant la phase usage**, l'eau pour **refroidir** les serveurs et recharger les circuits fermés.
3. **Pendant la phase usage**, l'eau pour **humidifier** l'air des salles informatiques.
4. Enfin, l'eau pour les sanitaires, mais aussi pour le nettoyage et l'arrosage des **équipements techniques**.

On se concentre ici sur l'eau consommée ([voir définition eau consommée](#)) pour faire fonctionner les salles IT (humidification et refroidissement), et sur l'eau consommée pour produire l'énergie nécessaire aux centres de données.

Les nécessités de refroidissement sont intimement liées à la température visée au sein des salles. Celle-ci a augmenté au fil des ans mais dépend de différents paramètres, c'est notamment une condition pour assurer le matériel informatique.

Selon l'étude en cours de Critical Building pour l'ADEME :

- Plus de 80% de la puissance du parc étudié est refroidie avec des températures supérieures ou égales à 24°C¹⁶¹. Après recoupement des données il s'agit pour la plupart de centres de données récents.

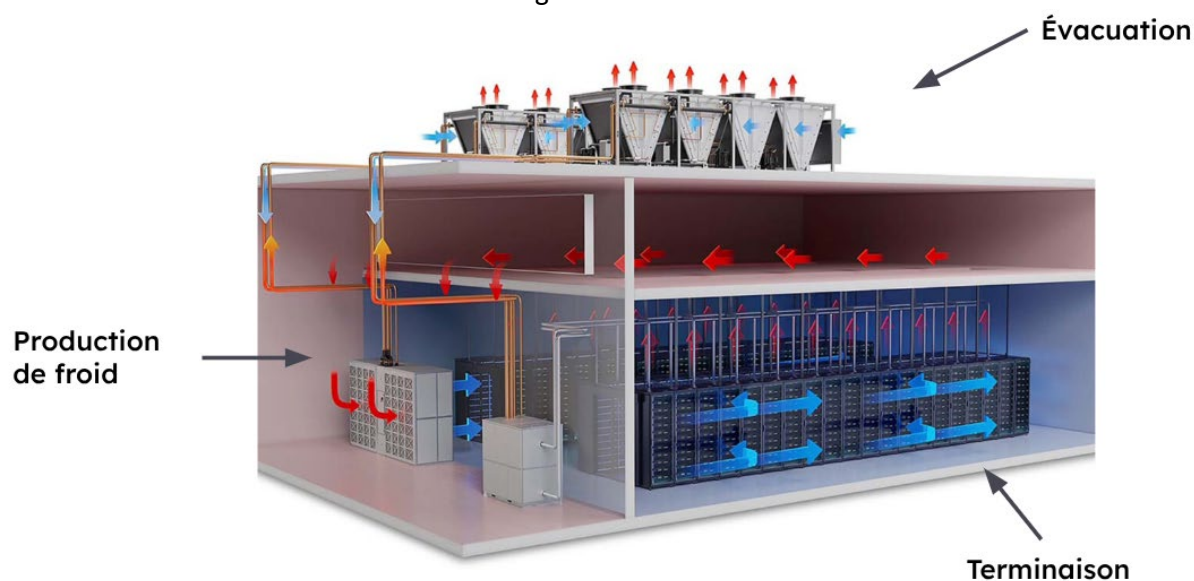
¹⁵⁹ Les interviews d'HPE et Equinix ont confirmé cette impression

¹⁶⁰ [Les data centers menacés d'obsolescence à cause de l'IA, La tribune, 2024](#)

¹⁶¹ Ces 24°C correspondent à la température de consigne ("soufflage en salle" selon le terme employé dans l'étude). Le choix de cette température reste une sécurisation relativement arbitraire, certains fournisseurs (comme Scaleway) faisant le choix d'augmenter cette température de consigne à 27°C par exemple.

- Certains sites ne peuvent pas augmenter leur température de soufflage en raison d'une urbanisation inadaptée.

Les consommations et prélèvements d'eau diffèrent en fonction des techniques de refroidissement utilisées et de la combinaison de ces technologies entre elles.



Distinction entre trois grandes parties d'un système de refroidissement : terminaison, production de froid et évacuation¹⁶²

Le tableau ci-dessous tente une synthèse (non exhaustive¹⁶³) des technologies de refroidissement présentes sur le marché, selon si leur rôle concerne la production de froid ou bien la terminaison du refroidissement (refroidissement des équipements IT).

Technologies impliquées dans le refroidissement	
Production de froid	
Catégorie	Déclinaisons
Prod. mutualisée (avec boucle d'eau glacée)	Groupes frigorifiques à condensation à air
	Groupes frigorifiques à condensation à eau
Prod intégrée/autonome (avec fluide frigorigène)	Détente directe dans l' air
	Détente indirecte dans l' eau (chaude)

¹⁶² Image venant du [site d'Accuspec](https://www.accuspec.com/)

¹⁶³ Ce tableau est inspiré de l'étude de Critical Building, pour le compte de l'ADEME, concernant les technologies de refroidissement en centre de données et leur évolution.

"Sans production" de froid (à coupler avec une autre technologie de production de froid)		
Catégorie	Déclinaisons	
Free-cooling	à air direct	
	à air indirect (avec échangeur thermique)	
Free-chilling	sur l'air extérieur (échange d'énergie eau/air extérieur)	
Geo-cooling	Par eau de nappe phréatique	
	Par boucle fermée	
River-cooling		
Sea-cooling		
Adiabatique (à coupler avec free-cooling ou free-chilling)		
Catégorie	Déclinaisons	
Adiabatique	Couplé au free-cooling	
	Couplé au free-chilling	
Terminaison		
Catégorie	Déclinaisons	Sous-déclinaisons
Refroidissement terminal par air	Centralisé en recyclage	Armoires de climatisation
		Fan Wall
		Centrales de Traitement d'Air indirectes
	Air décentralisé ou individuel	In-Rows
		Portes froides (door-cooling)
Refroidissement terminal par eau	Direct Liquid Cooling	DLC (direct to chip)
Refroidissement terminal par bain d'huile	Immersion Cooling	Liquide Phase Unique
		Liquide Double Phase

Le choix de la bonne solution de refroidissement pour un centre de données dépend de plusieurs facteurs :

- la taille des salles,
- la situation géographique,
- les variations de la température extérieure durant l'année (pour le free cooling)
- les usages développés au sein du centres de données
- le modèle économique de l'opérateur

Le refroidissement par air ou "Air-cooling"

Historiquement et de manière toujours majoritaire, le refroidissement, en terminaison dans les salles IT, s'effectue par air.

On peut distinguer plusieurs types de technologies pour acheminer l'air froid et extraire l'air chaud :

- Le CRAC ("Computer Room Air Conditioning") fait référence à une infrastructure de climatisation, courante dans les centres de données en raison de ses coûts modérés. Ce système fonctionne à la manière d'un climatiseur traditionnel, avec un compresseur qui assure le refroidissement par soufflage d'air sur un échangeur rempli de réfrigérant. Il est toutefois impossible de varier les flux d'air, du moins sur les systèmes les plus courants. Ce système consomme beaucoup d'énergie.
- Le CRAH ("Computer Room Air Handler") renvoie à un système de traitement d'air qui fonctionne avec une centrale d'eau glacée, comme on en trouve dans de nombreux immeubles de bureaux. L'eau glacée passe par un serpentín afin de refroidir l'air transporté par les unités. C'est un système adapté aux endroits où il fait naturellement froid.

Les premiers centres de données étaient conçus de manière simpliste, des armoires informatiques (baies/racks), dans une salle munie de systèmes de climatisation. Sans agencement particulier de la salle, refroidir un espace non confiné rempli d'équipements qui dissipent la chaleur est très inefficace. Pour augmenter l'efficacité énergétique du refroidissement par air, plusieurs principes d'agencement des salles IT ont vu le jour et sont aujourd'hui la norme.

Confinement des allées, allées froides et allées chaudes, faux plancher

Pour permettre une circulation efficace des flux d'air froid provenant des systèmes de refroidissement par air, les salles machines sont surélevées et l'espace entre le faux plancher et le sol est utilisé pour faire circuler l'air froid sous les rangées de baies.

Les rangées de baies sont positionnées pour former des allées. Les équipements sont positionnés dans les baies de manière à ce que l'aspiration d'air ait toujours lieu du même côté et que l'air rejeté s'effectue de l'autre côté. Entre deux rangées orientée face avant (du côté où les ventilateurs des équipements aspirent l'air), l'allée est confinée (portes coulissantes et fermeture de l'espace au-dessus des baies, dans le meilleur des cas). Dans cette allée appelée "allée froide", l'air froid circulant sous le faux plancher est diffusé pour refroidir les équipements. A l'extérieur de l'allée froide, la température est plus importante puisque l'air provient des équipements et s'est réchauffé. Les allées non confinées, entre deux rangées de baies sont appelées "allées chaudes". L'air ambiant, plus chaud, est aspiré par les CRAHs, refroidit, puis réinjecté sous le faux plancher.

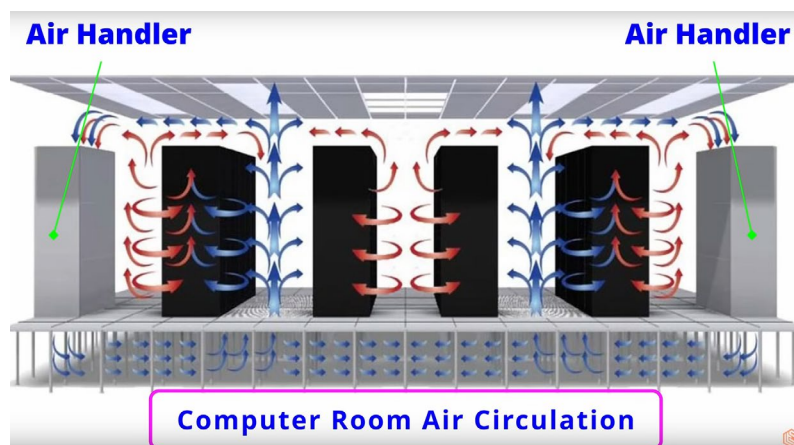


Schéma représentant les allées froides et chaudes¹⁶⁴

Refroidissement par air et évaporation de l'eau

Comme mentionné précédemment, certains centres de données avec terminaison de refroidissement par air sont également équipés de tours à évaporation externes. C'est ce type d'installation qui est généralement la plus consommatrice d'eau.

Selon l'article Making AI less thirsty,¹⁶⁵ par ces tours de refroidissement, l'eau s'évapore à hauteur de 80% des eaux prélevées. En moyenne, et en fonction des conditions climatiques et de la configuration du centre de données, ces centres de données peuvent faire s'évaporer entre 1 et 9 litres par kWh d'énergie des serveurs : 1L/kWh IT concerne la moyenne globale pour le parc de Google, et 9 litres pour un grand centre de données commercial pendant un été en Arizona.



Centres de données de Google, à The Dalles, dans l'Oregon (Crédit : Google¹⁶⁶)

¹⁶⁴ [Visite d'un Datacenter avec tours de refroidissement, Lawrence Systems](#)

¹⁶⁵ Li, Pengfei & Yang, Jianyi & Islam, Mohammad & Ren, Shaolei. (2023). Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models. 10.48550/arXiv.2304.0327

¹⁶⁶ [Galerie de photos de Google](#)

Selon David Mytton, dans un article de 2021¹⁶⁷, les consommations d'eau des centres de données concernés varient entre 1800 et 2900 litres par MWh d'électricité consommée. Toutefois, ce chiffre dépend fortement de l'efficacité du système de refroidissement, ainsi que des variations saisonnières et de l'emplacement du centre de données.

Il faut noter que cette configuration de refroidissement, particulièrement intensive en consommation d'eau, n'existe pas pour le moment en France. Il existe toutefois des centres de données utilisant du refroidissement adiabatique (voir [la section en rapport](#)).

Free-cooling

Le Free cooling consiste en le refroidissement par l'air provenant de l'extérieur du bâtiment, de manière directe ou indirecte. Il nécessite de monitorer et filtrer l'air avec précaution (pollution, taux d'humidité...). Un système complémentaire de refroidissement est toujours prévu car les conditions climatiques ne sont pas toujours adéquates pour le free cooling.

Ce type de refroidissement est dépendant de la température de l'air extérieur et est donc toujours utilisé en couplage avec un autre système, souvent un refroidissement par air classique. La phase d'interviews de cette étude a fait ressortir qu'un système de free cooling peut fournir entre 20 et 90 % du besoin en refroidissement de l'air intérieur, en fonction de la température de l'air extérieur, donc de manière variable au fil des saisons, le reste étant effectué par un autre système de production de froid ou un système [adiabatique](#), le plus souvent.

Selon l'article Making AI less thirsty,¹⁶⁸ quand les systèmes "outside air cooling" avec une assistance de refroidissement par eau sont utilisés, ce sont 70% des eaux prélevées qui sont consommées ou évaporées (données sur le parc de Meta). En langage moins technique, cela veut dire qu'un "free cooling" qui utilise un complément de refroidissement par eau résulte en un non-retour de 70% de l'eau prélevée dans son circuit d'origine.

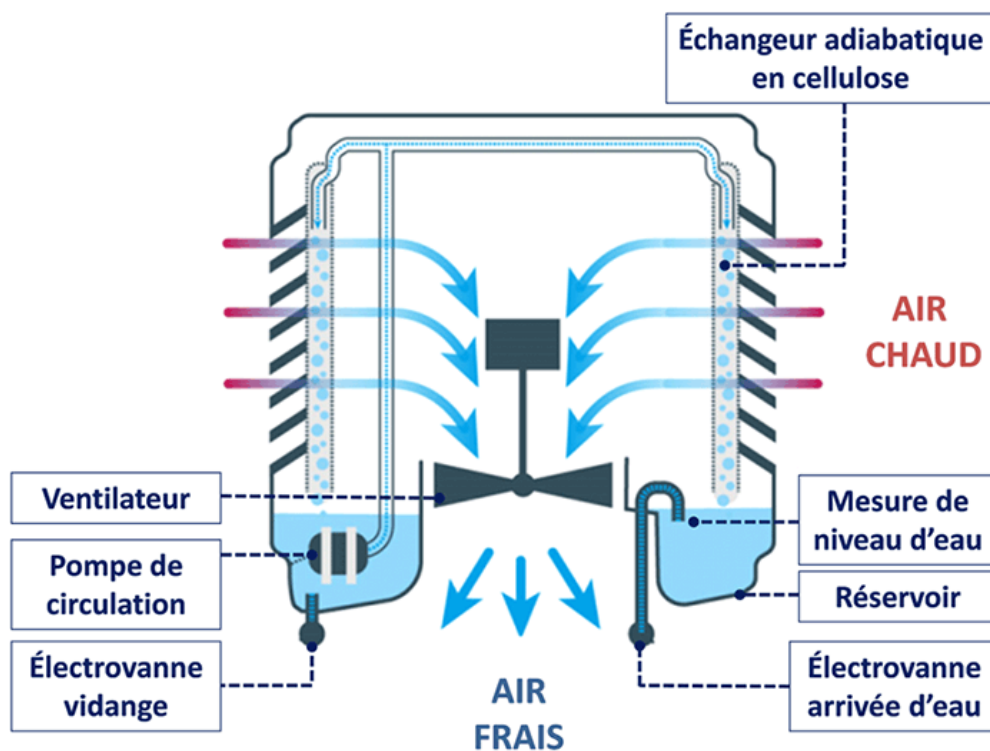
Ce type de couplage permet de réduire la consommation d'eau dans le cas de systèmes de cooling towers, puisqu'elle est estimée à 0,2 L/kWh IT (même source Meta), mais avec tout de même une variabilité selon le climat et l'équilibre entre l'utilisation de l'air extérieur et les compléments eau.

Refroidissement adiabatique

Le refroidissement adiabatique consiste à faire passer l'air chaud que l'on souhaite refroidir, par des filtres humides. Au contact de l'air, l'eau s'évapore et l'air se refroidit (puisque l'eau utilise l'énergie pour passer de l'état liquide à l'état gazeux).

¹⁶⁷ Mytton, D. Data centre water consumption. *npj Clean Water* **4**, 11 (2021). <https://doi.org/10.1038/s41545-021-00101-w>

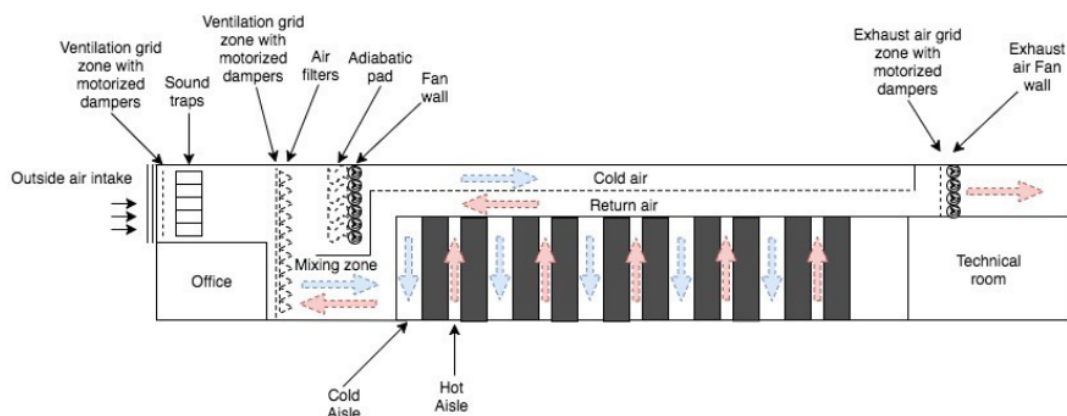
¹⁶⁸ *idem* 137



Principe du refroidissement adiabatique, source : xpair¹⁶⁹

On trouve dans les centres de données plusieurs versions de ce type de refroidissement, soit en cycle ouvert, soit en cycle fermé, ce qui implique une consommation d'eau (par évaporation) plus ou moins importante.

Le bon fonctionnement d'un refroidissement adiabatique est dépendant de la température de l'air entrant, mais aussi et surtout de son hygrométrie. Un air trop humide ne peut pas être refroidi correctement. Pour ces raisons, le refroidissement adiabatique est le plus souvent utilisé en combinaison avec une autre technique de refroidissement, soit des systèmes HVAC, soit du free-cooling, comme dans le centre de données DC5 de Scaleway.¹⁷⁰



¹⁶⁹ [Le refroidissement adiabatique, le futur de la climatisation, Xpair](#)

¹⁷⁰ [Visite du data center Scaleway DC5 par Vivien Guéant – LaFibre.info](#)

Immersion cooling

L'Immersion Cooling consiste à refroidir les équipements par contact direct avec un liquide diélectrique et ignifuge. Ce type de liquide est soit de l'eau déionisée, de l'huile minérale, des fluides à base de fluorocarbures ou des fluides synthétiques. Contrairement à l'eau utilisée pour le [DLC / Water Cooling](#), qui peut avoir un impact négatif sur le matériel en cas de fuite et de contact direct, ces liquides sont inoffensifs vis-à-vis du matériel. Ainsi, les équipements sont plongés directement dans le liquide. On distingue deux grandes familles de refroidissement par immersion :

Le refroidissement par immersion avec liquide à phase unique

Le refroidissement par immersion peut utiliser un liquide de refroidissement à phase unique, c'est-à-dire que ce liquide ne change jamais d'état : il ne bout pas, ne s'évapore pas et ne gèle pas. Le liquide une fois réchauffé est transféré au contact d'un circuit de refroidissement à eau, puis réintroduit dans le bassin de refroidissement une fois à bonne température.

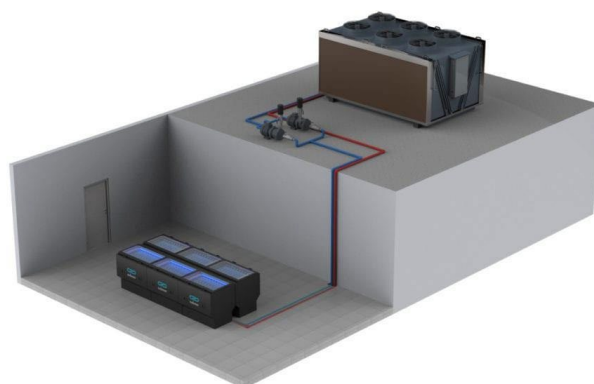


Illustration d'un bassin de refroidissement par immersion et du circuit de refroidissement à eau associé, source : Submer¹⁷²

L'avantage de ce type de liquide est que les bassins peuvent être accessibles librement, sans couverture nécessaire, puisqu'aucune évaporation du liquide est à prévoir.

Le refroidissement par immersion à double phase

Le refroidissement par immersion peut également utiliser du liquide à double phase, c'est-à-dire qu'il peut être à l'état liquide, ou bien à l'état gazeux. Dans ce cas, le système tire avantage du changement de phase de ce liquide, puisqu'il requiert de l'énergie (comme c'est le cas pour le [refroidissement adiabatique](#)).

¹⁷¹ Visite du data center Scaleway DC5 par Vivien Guéant – LaFibre.info

¹⁷² [What Is Immersion Cooling?](#), Submer

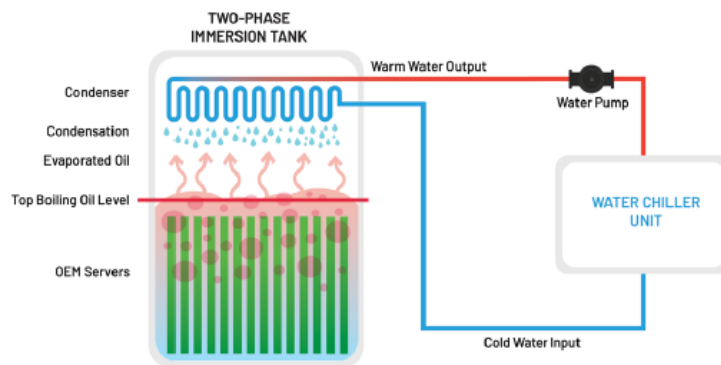


Schéma du refroidissement par immersion avec liquide à double-phase, source : Submer

Le liquide est refroidi lorsqu'il atteint le point d'ébullition et reste ainsi à une température intermédiaire. Il semble que ce second mode de fonctionnement soit moins plébiscité que le premier, car l'évaporation du liquide pose plus de problèmes logistiques : il faut notamment que les bacs de refroidissement soient fermés autant que possible. De plus, les liquides diélectriques à double phase sont, semble-t-il, encore à un prix prohibitif, bien que permettant une efficacité énergétique encore plus importante que les liquides à phase unique.¹⁷³

Les avantages du refroidissement par immersion

Qu'il s'agisse de refroidissement par immersion à simple ou à double phase, toutes les sources s'accordent pour dire que le refroidissement par immersion permet de consommer sensiblement moins d'énergie qu'un système de refroidissement classique. Certains centres de données pilotes affichent ainsi des PUE aussi bas que 1.02 grâce à l'immersion cooling (avec toutes les limites [que présentent l'indicateur PUE](#)).

L'avantage de cette technologie n'est pas une disparition de tout système de refroidissement (puisqu'il faut tout de même refroidir le liquide lui-même, comme vu précédemment). Cependant l'efficacité d'un refroidissement avec contact direct du liquide sur les composants est encore plus importante que le refroidissement par eau type [DLC](#) et bien plus importante qu'un refroidissement par air, ce qui réduit les volumes à refroidir et la déperdition de chaleur.

Il semble aussi que ce type de refroidissement permette une occupation des sols moindre, notamment car l'espace nécessaire aux systèmes de refroidissement est moins important.

Les limites du refroidissement par immersion

Cette étude a permis de relever, à travers les interviews menées et la veille effectuée, les points suivants pouvant expliquer la lente adoption du refroidissement par immersion à ce jour.

D'abord, le refroidissement par immersion requiert des précautions supplémentaires à la manipulation, à l'extraction et à l'installation des serveurs, ce qui nécessite un minimum de formation des personnes intervenantes, un peu de matériel supplémentaire (chiffon, gants, ...), mais surtout ce qui rend moins flexible les changements de matériel. Chez un fournisseur de services cloud par exemple, le rythme de désinstallation/installation des machines et de changement de composants, pourrait être gêné par ces précautions de manipulation supplémentaires.¹⁷⁴

¹⁷³ [Comparison of server liquid-cooling technologies](#), Syska, Hennessy Group, Inc.

¹⁷⁴ [Immersion cooling : une technologie de refroidissement prometteuse pour les serveurs, à condition de la maîtriser](#), APL Datacenter

Ensuite, la disposition des machines et donc l'urbanisme des salles IT doit être radicalement différent.



Un caisson de refroidissement par immersion A103-CC0 de chez Gigabyte¹⁷⁵

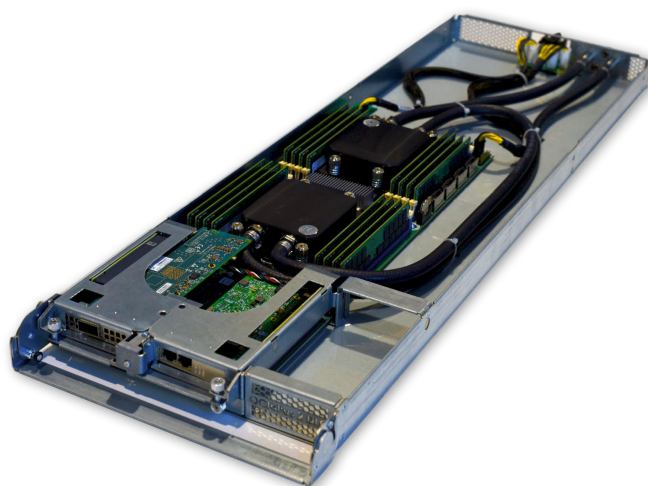
Dans cette configuration, les équipements doivent être installés à la verticale et non à l'horizontale, comme c'est le cas dans des baies informatiques classiques. La forme et la visserie nécessaire à leur installation diffère également, ce qui implique d'acheter des équipements livrés dans le bon format, donc que le fournisseur possède le nécessaire dans son catalogue.

Les interviews menées pour cette étude ont également soulevé la question de la corrosion, liées aux liquides de refroidissement, mais aussi du recyclage de ces liquides. Nous n'avons pas été en mesure de sourcer ces problématiques.

DLC et DLcC

Direct liquid cooling ou DLC : le Direct Liquid Cooling, consiste à refroidir de l'eau (deminéralisée ou bien un mélange d'eau et de glycol) qui est apportée directement au niveau des baies informatiques, voire au contact des composants eux-mêmes, au sein des équipements, l'eau étant un médium bien plus caloporteur que l'air. Le DLcC ou Direct Liquid Cooling to Chip, désigne généralement la version la plus aboutie de cette technologie de refroidissement, dans laquelle le liquide est apporté jusqu'au contact de la puce elle-même.

¹⁷⁵ [Caisson de refroidissement par immersion A103-CC0 de Gigabyte](#)



Un serveur "DLC" OctoPus 2GG de chez 2CRSi, les tuyaux noirs visibles sur la partie droite apportent l'eau au plus proche des composants les plus responsables d'émission de chaleur¹⁷⁶

Ce système est historiquement très utilisé sur les systèmes HPC, qui présentent classiquement une densité énergétique bien supérieure aux serveurs traditionnels. Son usage est aussi de plus en plus répandu en ce qui concerne les infrastructures dédiées à l'IA. Son utilisation dans la version "direct-to-chip", soit avec un passage de l'eau au plus proche du composant, semble systématique dès que la puissance demandée dépasse 40 kW pour une baie.

Ainsi les produits Nvidia récents, comme les cartes GB200¹⁷⁷ ont pour pré-requis l'utilisation du DLC pour leur refroidissement.

Selon le "Code of Conduct 2021", ce système permettrait de récupérer de la chaleur à plus haute température. On pourrait donc en conclure qu'il présente un avantage pour la récupération de chaleur fatale.

Door cooling

La mise en place du DLC "to-chip" nécessite de concevoir les salles IT en ce sens et d'installer des serveurs prévus pour cet usage. C'est un investissement qui est difficilement fait dans le cas de salles IT existantes. Pour répondre à la problématique de refroidissement des systèmes d'IA comprenant une certaine densité de [GPUs](#) dédiés à cet usage, mais aussi pour les baies comprenant des infrastructures de stockage utilisées par ces systèmes d'IA, dans les salles IT existantes, une autre version du Direct Liquid Cooling est de plus en plus utilisée : le Door Cooling.

Il s'agit d'une variante de l'application du DLC : il faut disposer d'un système de DLC, mais le raccordement jusqu'aux machines ou aux composants n'est pas nécessaire, ce qui simplifie drastiquement son installation. Le circuit d'eau froide est ici connecté à la porte arrière de la baie ("rear door cooling"), qui dispose d'un échangeur eau vers air, permettant ainsi de projeter de l'air refroidi

¹⁷⁶ [Serveur OctoPus 2GG, 2CRSi](#)

¹⁷⁷ [GB200 hardware architecture and components, Semi Analysis, 2024](#)

directement dans la baie. Ce type de système permettrait de fournir le refroidissement nécessaire à une baie d'une capacité de 75 kW¹⁷⁸.



RDHX with 42u or 48u Rack Options

Des portes de baies équipées du système de door cooling de Motivair

River cooling

Selon l'emplacement géographique du centre de données, il est possible d'utiliser les eaux de surface pour son refroidissement (rivière, lac...). L'eau prélevée est rejetée sans être stockée. Il n'y a donc a priori pas de phénomène d'appauvrissement des nappes ou des rivières. Il est toutefois nécessaire que l'eau rejetée soit à la même température que le milieu dans lequel elle est rejetée. Une autorisation est nécessaire (exemple de Marseille, Digital Realty : MRS 2 et 3).

Il faut noter justement que l'eau rejetée à Marseille serait plus chaude que prévue (27 degrés au lieu de 23,4), et pourrait causer des dégâts sur la biodiversité marine : les risques d'eutrophisation (déséquilibre du milieu provoqué par l'augmentation de la concentration d'azote et de phosphore)¹⁷⁹.

Etude de cas : la technologie propriétaire d'OVH, Hybrid Immersion Liquid Cooling »

Dans son dernier rapport, OVH met en avant un système de refroidissement qui lui permet d'atteindre un WUE de 0,2L/kWh, contre, selon l'US Dep. of Energy, un WUE moyen des industries du cloud à 1,8L/kWh. OVH a notamment mis en place ce système sur son nouveau centre de données SBG5 à Strasbourg.

L'« Hybrid Immersion Liquid Cooling » combine le watercooling (refroidissement par eau) et l'immersion cooling (immersion complète du serveur dans un fluide diélectrique). Elle est composée d'un système de

¹⁷⁸ [Motivair, Liquid Door Cooling](#)

¹⁷⁹ La Quadrature du net, 2024

refroidissement par eau directement sur la puce et d'un système de refroidissement par immersion naturel passif monophasé :

- Watercooling : refroidissement d'un dissipateur thermique par des blocs d'eau disposés sur les processeurs (CPU, GPU) avec la même solution que celle utilisée dans tous les serveurs OVHcloud et un serpentin de convection propriétaire relié à une sous-station de pompage (PSS) et à un dry cooler pour évacuer la chaleur à l'extérieur du DC ;
- ImmersionCooling : le fluide est contenu dans un réservoir et refroidit l'ensemble des équipements informatiques dans le serveur, et pas uniquement les processeurs ; ce fluide remplace l'air circulant dans les serveurs OVHcloud et améliore ainsi l'efficacité de tout composant non refroidi par les systèmes de refroidissement à eau d'OVHcloud.

Pistes d'amélioration

Le Code of conduct 2021 recommande plusieurs pistes pour diminuer l'utilisation d'énergie et d'eau pour le refroidissement :

- Récupération des eaux de pluie : récupération et stockage des eaux de pluie pour refroidissement par évaporation ou autres sources non potables qui réduisent la consommation énergétique
- Utilisation d'autres sources d'eau : utilisation d'eaux grises pour refroidissement par évaporation
- Mesure de la consommation d'eau : toutes les sources d'eau doivent être monitorées, pour trouver les leviers de réduction de consommation d'eau. Le WUE, défini depuis 2022, par la norme ISO/IEC 30134-9 et l'EN 50600-4-9 doit aider en ce sens.

Selon Karimi and al¹⁸⁰, avec deux cas d'étude à Phoenix, les PUE et WUE des techniques basées sur l'eau ou sur l'air ont des résultats distincts :

<u>Karimi 2022</u>		<u>PUE</u>	<u>WUE</u>	
DC1 Phoenix	54	1,84	3,1	air-cooled chillers
DC2 Phoenix	34,5	1,63	9,11	water-cooled chillers and evaporative cooling

Comme expliqué dans la section concernant le [PUE](#), la consommation d'énergie d'un centre de données évolue fortement en fonction de la température extérieure, mais également en fonction de l'humidité de l'air extérieur dans le cas du refroidissement adiabatique et du free-cooling (puisque'il faut employer une autre technique de refroidissement dans ce cas).

La variabilité de la consommation électrique des centres de données en fonction des conditions météorologiques extérieures (température, rayonnement solaire...) dépend fortement du type de centre

¹⁸⁰ Leila Karimi, Leeann Yacuel, Joseph Degraft- Johnson, Jamie Ashby, Michael Green, Matt Renner, Aryn Bergman, Robert Norwood, Kerri L. Hickenbottom, *Water-energy tradeoffs in data centers: A case study in hot-arid climates*, Resources, Conservation and Recycling, Volume 181, 2022

de données, de la technique de refroidissement utilisée et de son efficacité. L'évaluation de cette thermosensibilité semble difficile en France par manque de sites anciens permettant un retour d'expérience détaillé sur la question. L'effet sur la consommation électrique du changement climatique, associé à une augmentation des températures moyennes mais aussi à une hausse de la fréquence des événements climatiques extrêmes comme les canicules, reste donc difficile à estimer.

Évolutions des pratiques et choix d'architecture IT

Virtualisation

La virtualisation désigne un ensemble de technologies permettant d'abstraire un système logiciel du système d'exploitation principal et du matériel sous-jacent. Cette abstraction permet un contrôle fin de l'allocation des ressources matérielles, mais aussi d'isoler ce système du reste des logiciels exécutés sur le serveur, notamment pour plus de sécurité.

Cette pratique est à la base des plateformes des services Cloud, puisqu'elle permet de mettre à disposition des utilisateurs de la plateforme, un sous ensemble de ressources, de programmes et de données, qui seront inaccessibles aux autres utilisateurs, puis de les facturer pour cet usage. La virtualisation est ainsi un mécanisme essentiel pour l'hébergement de services numériques, qui est devenu la norme depuis les années 2000. Ce mécanisme est aussi décrit par l'industrie comme central dans la "consolidation" des ressources informatiques, c'est-à-dire dans la capacité d'une entreprise à rassembler un maximum d'usages et de logiciels sur une même machine, de manière à employer le moins de machines possibles, principalement pour des raisons économiques (tout en permettant de consommer moins d'énergie et de ressources).

Le fait d'abstraire le système et les applications qui en dépendent, du matériel sous-jacent, permet également la migration de ce système, aussi appelé "machine virtuelle", d'une machine physique à une autre. Cette migration dite "à froid" lorsqu'il est nécessaire d'éteindre la machine virtuelle pour effectuer l'opération, ou "à chaud" lorsque la machine virtuelle peut rester en activité durant l'opération, permet ainsi aux acteurs de l'hébergement et du Cloud de remplacer des composants matériels et d'effectuer des opérations de maintenance simplement tout en permettant la continuité de service.

La grande question liée à la virtualisation, lorsque l'on s'intéresse aux impacts environnementaux d'une infrastructure d'hébergement de type Cloud, est le taux de remplissage des machines physiques. Autrement dit, la part des ressources informatiques disponibles sur les serveurs physiques (CPU, mémoire, GPU, stockage), effectivement allouées et utilisées par les clients, détermine la part de ressources inutilisées qui vont venir aggraver les impacts environnementaux globaux de la plateforme, d'un point de vue comptable.

Ces taux de remplissage sont une information bien gardée par les fournisseurs de services cloud et semblent être au centre de leur stratégie concurrentielle. En parallèle, la virtualisation est toujours présentée par le secteur comme un levier important de la consolidation des ressources, donc de l'efficacité énergétique et matérielle des centres de données.

L'autre mécanisme expliquant cette analyse réside dans l'élasticité permise par les plateformes et les technologies de Cloud. Parfois appelée "auto-scaling", cette élasticité permet de déclencher

automatiquement le lancement ou l'arrêt d'une machine virtuelle, occupant ou libérant ainsi des ressources, en fonction de la demande du moment, soit par exemple lorsque le nombre de requêtes provenant des utilisateurs augmente.

Puisqu'il ne s'agit que d'un ensemble de technologies et de pratiques, qui peuvent être utilisées de diverses manières, rien n'empêche une entreprise de pratiquer massivement la virtualisation tout en gâchant des ressources (en laissant allumés beaucoup plus de serveurs physiques que nécessaire, en réservant trop de ressources pour chaque applicatif, etc.). Le taux de ressources ainsi gâchées dans les infrastructures on-premise est un des arguments massue des fournisseurs de cloud pour encourager une migration vers leurs plateformes.

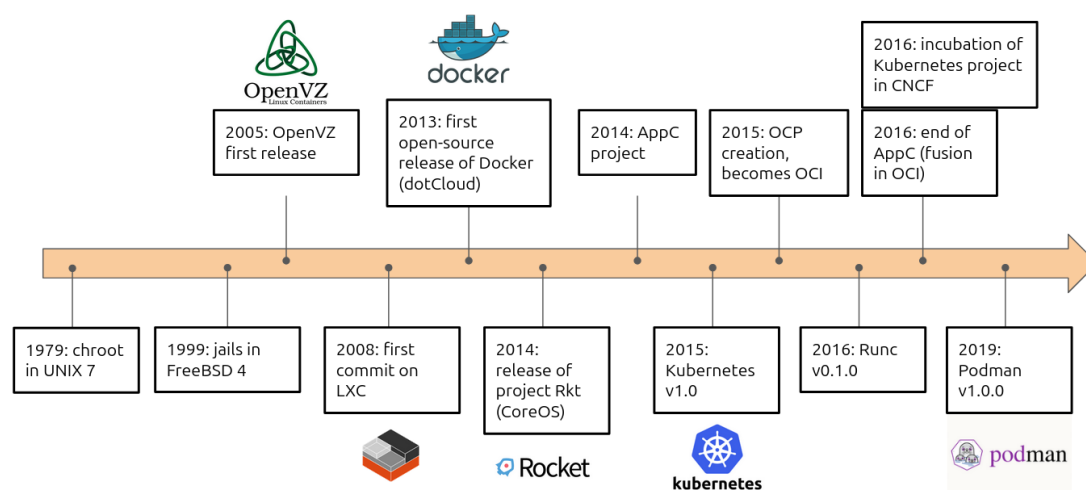
Pourtant, les projections de réduction d'impact permises par ce type de migration vers le cloud, partiellement atteignables sous certaines conditions, ne semble pas entièrement vérifiables ni résistantes à une analyse plus poussée¹⁸¹.

Conteneurisation

Peu après la généralisation massive de la virtualisation, un autre type de technologie et de pratiques a connu un regain d'intérêt, qui s'est converti en tendance durable jusqu'à aujourd'hui dans les infrastructures d'hébergement et de cloud : les conteneurs.

Un "conteneur" est basé sur un ensemble de technologies permettant également une forme d'abstraction d'un ensemble de logiciels, vis-à-vis du système d'exploitation et du matériel sous-jacent. Réputés comme consommant moins de ressources informatiques qu'une machine virtuelle à usage égal et permettant un démarrage plus rapide (facteurs clés des arguments pour adoption), les conteneurs ont été généralisés suite à plusieurs événements au tournant de la décennie 2010 : d'abord la sortie de la première version d'OpenVZ en 2005, puis la première version open-source de Docker en 2013, édité alors par l'entreprise DotCloud.

Ce second événement coïncide avec l'avènement des méthodes DevOps à partir de 2008 et la volonté des entreprises de fluidifier les processus de déploiement de leurs applications, ce qui comprend une meilleure collaboration entre les équipes de développement et les équipes responsables des opérations d'infrastructures.



¹⁸¹ [Les réductions d'émissions de CO2 promises par les Cloud providers sont-elles réalistes ? Boavizta, 2022](#)

Historique des événements technologiques à l'origine de la démocratisation des conteneurs dans les plateformes d'hébergement et de cloud¹⁸²

La démocratisation des conteneurs a fondamentalement changé les pratiques au sein des équipes de développement et des infrastructures IT. Les applications sont maintenant empaquetées dans une image de conteneur (version statique, décrite, de ce que doit contenir un conteneur lorsqu'il sera déployé), ce qui permet aux développeurs de fournir un outil prêt à l'emploi pour les personnes responsables du déploiement, ou bien d'automatiser, bien plus simplement qu'auparavant, le déploiement de l'application en production. Si l'on prend en compte l'effet systémique de l'accélération des déploiements permise par cette évolution, la balance entre le gain et le coût énergétique et environnemental de l'apport de cette technologie est difficile à estimer.

Dans un sens, le conteneur permet une allocation plus fine des ressources matérielles par application, que ne le permet une machine virtuelle : le conteneur ne contient que le logiciel qui fournit le service et ses dépendances, là où une machine virtuelle embarque un système d'exploitation complet. Grâce à cette légèreté, les plateformes et technologies de Cloud permettent de déployer au besoin de nouveaux conteneurs, en fonction du trafic, d'une manière encore plus fine que ce n'était permis avec les machines virtuelles (voir section précédente), ce qui semble bon pour un moindre impact environnemental.

Dans l'autre sens, la fluidification et la plus grande performance de déploiement de nouvelles applications permises par le mouvement DevOps et les conteneurs, ont certainement accéléré le nombre de déploiements et ainsi permis d'augmenter les ressources consommées en valeur absolue. La quantification de ces deux phénomènes et du probable effet rebond associé est à ce jour difficile à effectuer, par manque de données fiables relatives aux usages réels des entreprises de ces technologies.

Les catalogues des fournisseurs de services Cloud sont bien plus larges que la seule fourniture de machines virtuelles ou de conteneurs et les services proposés influencent massivement la manière dont les projets informatiques sont à présent gérés et déployés en production. Une analyse de ces offres dépasse le cadre de la présente étude.

IA Générative

Les modèles d'IA générative et en particulier de LLMs sont l'un des facteurs clef de l'augmentation des consommations des centres de données de manière récente. Il faut cependant noter que ce secteur est en constante évolution. Si l'augmentation importante des volumes de calculs et de données traitées, ainsi que de la consommation d'énergie associée, des grands modèles de LLMs est pour le moment soutenue, plusieurs facteurs sont à prendre en compte pour tenter de projeter ces consommations à l'avenir.

Du point de vue des modèles, du logiciel et des données, plusieurs facteurs et pratiques sont susceptibles de limiter l'occupation de ressources matérielles et la consommation d'énergie.

L'Afnor Spec IA Frugale¹⁸³, publiée en 2024, fournit 91 bonnes pratiques pour réduire l'empreinte d'un système (modèles) ou d'un service d'IA.

¹⁸² [Virtualization & containers course, Benoit Petit, 2021](#)

¹⁸³ [Afnor Spec IA Frugale, 2024](#)

Nombre de bonnes pratiques selon les domaines et les étapes du cycle de vie

Étape du cycle de vie	Service	Données	Infrastructures	Total
0 – Transverse	9			9
1 – Initialisation	4	1	5	10
2 – Conception et Développement	10	10	6	26
3 – Vérification et validation	2	3	2	7
4 – Déploiement	2	1	3	6
5 – Exploitation et suivi	1	4	2	7
6 – Validation continue	3	4	1	8
7 – Réévaluation	4	6	1	11
8 – Mise hors service		3	4	7
Total	35	41	33	

Nombre de bonnes pratiques issues de l'Afnor Spec IA Frugale, par catégorie

Les bonnes pratiques ayant été estimées comme les plus impactantes sont : Instruire la frugalité dans chaque projet IA (BP12), Acculturer et former les parties prenantes (BP14), Choisir la solution pour répondre au besoin en considérant les alternatives à l'IA (BP02), Optimiser l'usage de l'équipement existant (BP20), Utiliser des méthodes d'analyse de besoin pour mettre en œuvre la frugalité (BP01).

On note la prépondérance des pratiques consistant à n'utiliser l'IA qu'en adéquation stricte avec un besoin bien identifié, ainsi qu'à l'utiliser dans le cadre d'une démarche de frugalité.

A l'inverse de ces recommandations pour un moindre impact, le laboratoire de recherche Limites Numériques montre que l'usage de l'IA générative est en partie forcé à travers le design des interfaces des outils numériques¹⁸⁴.

Durée de vie des équipements

L'augmentation des durées de vie du matériel est un levier de réduction d'empreinte environnementale important, qu'il s'agisse de terminaux utilisateurs ou d'équipements IT présents en centre de données.

Il faut noter une évolution sur ce sujet chez les fournisseurs de services Cloud, principalement pour des raisons économiques. Google¹⁸⁵, Azure¹⁸⁶ et AWS ont annoncé augmenter la durée de vie de ses serveurs à 6 ans¹⁸⁷ sur les trois dernières années.

Certains acteurs plus modestes, comme Infomaniak, maximisent la récupération de composants matériels et la réutilisation des châssis des serveurs pour atteindre 10 à 15 ans de durée de vie par machine (pour la partie fonctionnelle).

¹⁸⁴ [Le forcing de l'IA, Limites Numériques](#)

¹⁸⁵ [Google increases server life to six years, will save billions of dollars, DatacenterDynamics, 2023](#)

¹⁸⁶ [Microsoft extends Azure server lifetimes by 50%, Network World, 2022](#)

¹⁸⁷ [Amazon extends the life of its servers to six years, expects \\$900m benefit in 90 days, The Register, 2024](#)

Cependant, la course à l'IA générative semble avoir contrebalancé cette tendance, si l'on en croit les bénéfices records d'Nvidia (voir section [fournisseurs de matériel](#)).

1.5. Évolutions environnementales et sociales

Les impacts sur l'environnement des nouvelles technologies, et plus spécifiquement ici, des centres de données, sont multiples. Un raccourci fréquent consiste à considérer que l'énergie serait le seul indicateur de leur impact environnemental, ce qui est réducteur. Ce raccourci a pour conséquence d'invisibiliser d'autres effets sur l'environnement intervenant en amont et aval de la phase d'utilisation (fabrication, transport, fin de vie) et les effets multiples sur l'environnement, sur toutes les phases : potentiel de réchauffement global, épuisement des ressources en minéraux et métaux, changement d'utilisation des sols, potentiel de privation d'eau de l'utilisateur, eutrophisation, ...

Bien que cette étude traite des consommations des centres de données, nous tenons à rappeler que l'énergie est avant tout un indicateur de flux :

- L'indicateur de flux rend compte d'une quantité d'un flux donné, transitant pour un processus sur une période donnée. Par exemple, la quantité d'énergie finale consommée est une indication quantitative d'un flux mais ne permet pas de rendre compte de ses effets sur l'environnement (dépendante non seulement de sa quantité mais d'autres paramètres tels que le mix électrique utilisé pour la production d'énergie primaire, le taux de déperdition entre énergie primaire et finale, les localisations de la production et la consommation...)
- L'indicateur d'impact environnemental rend compte de conséquences environnementales constatées ou potentielles, issues d'un ou plusieurs indicateurs de flux. Par exemple : les émissions de Gaz à Effet de Serre permettent de déterminer un potentiel de réchauffement climatique.

Nous traiterons dans ce chapitre des principales consommations des centres de données (énergie, eau, foncier) et leurs conséquences dynamiques environnementales et sociales.

1.5.1. Énergie

Les centres de données consomment de l'énergie :

- Électricité pour le bâtiment, le matériel IT et les équipements techniques en particulier de refroidissement,
- Énergies fossiles pour les générateurs de secours (plus rarement huiles végétales, mais en cours de développement)

Les émissions eq.CO2 et les impacts environnementaux de la production électrique impliquée dépendent du mix énergétique du pays ou de la région (tout comme la consommation d'eau utilisée dans cette production).

A l'échelle des bâtiments, on parle de :

Définitions des puissances électriques

Puissance
raccordée
totale

Puissance électrique en MW mise à disposition par RTE ou Enedis

Puissance
constatée

Puissance électrique correspondant à la consommation électrique effective du centre de données

Puissance IT

Puissance électrique constatée de la partie informatique

Puissance
non-IT

Puissance électrique constatée de la partie hors IT (refroidissement, onduleurs, ...)

Les consommations d'énergie par les centres de données

Données Monde

Les études internationales indiquent une consommation actuelle et prospective en forte croissance pour les centres de données.

A l'échelle mondiale, l'AIE a revu en 2023 ses estimations à la hausse. Alors que l'évolution de la consommation électrique des centres de données était présentée comme stable (à environ 200 TWh) depuis plus de 10 ans, l'estimation mise à jour pour 2024¹⁸⁸ présentait une consommation de **460 TWh**.¹⁸⁹ D'un point de vue prospectif, l'AIE table sur **une hausse de plus de 75% de la consommation par rapport à 2024 pour atteindre 800 TWh en 2026**. En 2021, Schneider Electric a publié une étude prévisionnelle à l'échelle mondiale qui évalue une forte croissance des centres de données de 341 TWh en 2020 à **719 TWh en 2030**.¹⁹⁰

De son côté, JLL projette une croissance de la puissance installée de 15% par an jusqu'en 2027¹⁹¹.

Données Etats-Unis

Dans son rapport 2024 "Powering Intelligence : Analyzing Artificial Intelligence and Data Center Energy Consumption" ¹⁹² l'EPRI (Electric Power Research Institute) estime à **152 TWh** la consommation électrique

¹⁸⁸ <https://www.iea.org/reports/world-energy-outlook-2024>

¹⁸⁹ <https://iea.blob.core.windows.net/assets/6b2fd954-2017-408e-bf08-952fdd62118a/Electricity2024-Analysisandforecastto2026.pdf>

¹⁹⁰ https://download.schneider-electric.com/files?p_Doc_Ref=digitalecoandclimate&p_enDocType=EDMS

¹⁹¹ 2025 Global Data Center Outlook, JLL, 2024

¹⁹² <https://www.epri.com/research/products/000000003002028905>

des centres de données aux Etats-Unis en **2023**, soit environ 4% de la consommation globale. L'EPRI note une forte croissance depuis 2020.

L'EPRI a aussi travaillé sur quatre scénarios prospectifs. Le scénario le plus prudent se base sur les estimations de croissance des acteurs du centre de données et a été produit avant l'avènement de ChatGPT et le boom de l'IA. Le plus agressif, plus récent, se base sur une hypothèse de déploiement rapide de l'IA. **Ils aboutissent à des consommations en 2030 qui varient de 196 à 404 TWh, soit des augmentations de 28% à 166% par rapport à 2023.**

Données Europe

Dans la "Directive on energy efficiency and amending Regulation" la Commission Européenne estime que la demande en électricité des centres de données pourrait passer de 77 TWh en 2018 à **99 TWh en 2030**¹⁹³ soit une croissance de près de 30% de la consommation en 12 ans. De son côté, le rapport "Energy-efficient cloud computing technologies and policies for an eco-friendly cloud market" du Borderstep Institute qui est à la base de la directive indique que par rapport à 2018, la consommation énergétique des centres de données devrait augmenter de 21% pour atteindre **92,6 TWh/an en 2025**.¹⁹⁴

Données France

Les évolutions tendancielles à 2030 ou 2050 montrent que les centres de données vont représenter dans les années à venir une part de plus en plus importante de l'impact environnemental du numérique.

L'étude ADEME/ARCEP de **2021** tablait sur une consommation de **16 TWh en 2030**.¹⁹⁵ Dans le volet 3 de son rapport "Evaluation de l'impact environnemental du numérique en France et Analyse Prospective" l'ADEME indique que *"La part des centres de données devrait passer d'environ 23% de la consommation électrique du numérique à plus de 42% sur cette période ce qui correspondrait à une multiplication par 3,6 entre leur consommation de 2020 et celle de 2050"*¹⁹⁶.

De manière plus globale, en 2020, l'ADEME a inclus les centres de données dans ses scénarios prospectifs de "Transitions 2050". Un des scénarios estimait une consommation électrique d'environ 35 TWh en 2050.

Dans son rapport sur le bilan prévisionnel de l'équilibre offre-demande 2023-2035, le gestionnaire du réseau de transport d'électricité RTE tablait, dans sa consultation, sur une consommation entre 20 et 35 TWh en 2035, en notant un haut niveau d'incertitude, qui intègre des prudenances sur leur raccordement effectif.¹⁹⁷ Ces chiffres ont été revus dans une version ultérieure du rapport, annonçant une trajectoire entre 23 et 28 TWh de consommation des centres de données en 2035.¹⁹⁸

¹⁹³ https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ%3AJOL_2023_231_R_0001&qid=1695186598766

¹⁹⁴ <https://digital-strategy.ec.europa.eu/en/library/energy-efficient-cloud-computing-technologies-and-policies-eco-friendly-cloud-market>

¹⁹⁵ <https://www.arcep.fr/la-regulation/grands-dossiers-thematiques-transverses/lempreinte-environnementale-du-numerique/etude-ademe-arcep-empreinte-environnemental-numerique-2020-2030-2050.html>

¹⁹⁶ <https://librairie.ademe.fr/consommer-autrement/5226-evaluation-de-l-impact-environnemental-du-numerique-en-france-et-analyse-prospective.html>

¹⁹⁷ <https://assets.rte-france.com/prod/public/2023-06/Bilan-pr%C3%A9visionnel-2023-point-d-%C3%A9tape.pdf> page 26

¹⁹⁸ <https://assets.rte-france.com/prod/public/2024-07/2024-06-07-chap2-consommation.pdf>

Dans l'édition 2025 de son observatoire (enquête "Pour un numérique soutenable"), l'ARCEP enquête sur 21 opérateurs de colocation, représentant 150 sites en France en 2023.¹⁹⁹

La consommation d'électricité des centres de données étudiés progresse de 8% entre 2021 et 2023, en passant de 1,9 à 2,4 TWh. L'augmentation de la consommation électrique des centres de données en 2023 est principalement portée par les centres de données mis en service entre 2021 et 2023, qui représentent moins d'un cinquième des centres de données étudiés.

La consommation spécifique de l'IA

Répartition apprentissage et inférence

La phase d'[apprentissage](#) des modèles d'IA est souvent considérée comme la plus énergivore. HuggingFace a indiqué que son projet BLOOM a consommé 433 MWh d'électricité pendant la phase d'apprentissage.²⁰⁰ Le chercheur Alex De Vries a documenté la consommation de l'apprentissage d'autres LLM (Large Language Models), y compris GPT-3, Gopher et OPT, qui ont consommé respectivement 1 287, 1 066 et 324 MWh.²⁰¹ Un article de SemiAnalysis²⁰² indique qu'OpenAI avait besoin de 3 617 serveurs HGX A100 et donc de 28 936 GPU pour opérer le moteur d'inférence de **ChatGPT**, ce qui implique une consommation d'énergie de **564 MWh par jour**. Cela semble considérablement plus important comparé à la consommation de 1 287 MWh estimée pour la phase d'apprentissage de GPT-3. Dans l'article cité précédemment,²⁰³ De Vries indique que **Google estime que 60 % de la consommation d'énergie liée à l'IA provenait de la phase d'inférence** (pour la période de 2019 à 2021). À contrario, HuggingFace indique que **le modèle BLOOM consomme beaucoup moins d'énergie pendant l'inférence que pendant la phase d'apprentissage**.²⁰⁴ Cette répartition relative entre énergie consommée par la phase d'apprentissage et phase d'inférence est potentiellement liée à un différentiel quant à la quantité d'inférences réalisées par l'ensemble des utilisateurs pour chacun des modèles.

Il existe donc une incertitude sur la répartition précise de la consommation d'énergie entre les phases d'apprentissage et d'inférence, bien qu'il y ait consensus sur le fait que les deux étapes sont très consommatrices d'énergie.

L'enseignement est qu'il faut bien considérer l'ensemble du cycle dans les estimations et tenir compte des évolutions des usages des utilisateurs de ces services.

Les estimations de consommation

L'AIE, dans son rapport "Electricity 2024 - Analysis and forecast to 2026", effectue quelques extrapolations pour illustrer l'impact potentiel de l'IA en termes de consommation électrique et conclut que d'ici 2026, l'industrie de l'IA devrait connaître une croissance exponentielle et consommer au moins dix fois sa demande de 2023.²⁰⁵

Gartner a publié en novembre 2024 un article assez alarmiste sur les besoins en énergie électrique de l'IA.²⁰⁶ Le cabinet de conseil estime que l'énergie nécessaire aux centres de données pour faire

¹⁹⁹ https://www.arcep.fr/fileadmin/user_upload/observatoire/enquete-pns/edition-2025/enquete-annuelle-pour-un-numerique-soutenable_edition2025.pdf

²⁰⁰ <https://arxiv.org/abs/2211.02001>

²⁰¹ https://www.researchgate.net/publication/374598219_The_growing_energy_footprint_of_artificial_intelligence

²⁰² <https://semianalysis.com/2023/02/09/the-inference-cost-of-search-disruption/>

²⁰³ https://www.researchgate.net/publication/374598219_The_growing_energy_footprint_of_artificial_intelligence

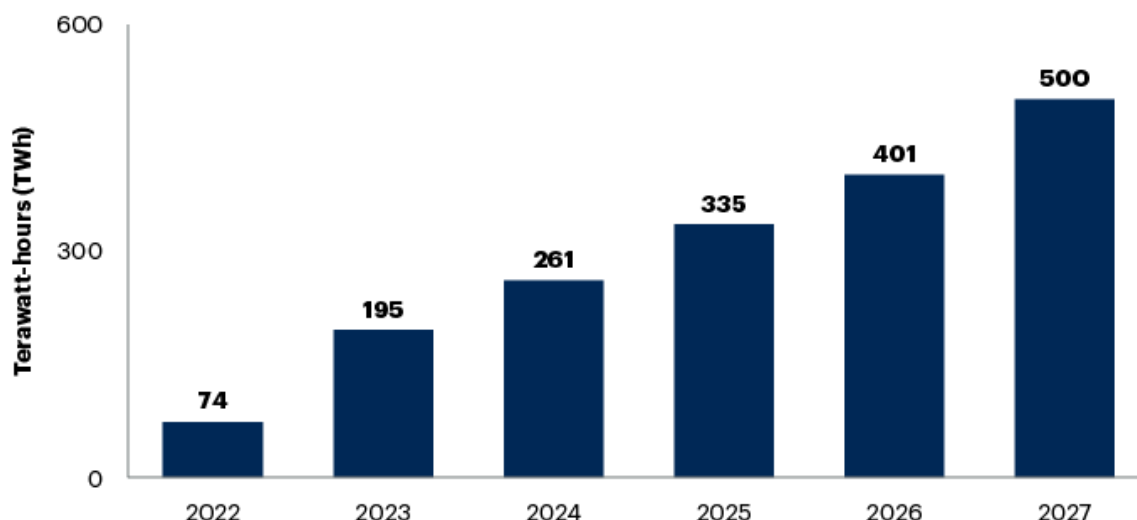
²⁰⁴ <https://arxiv.org/abs/2211.02001>

²⁰⁵ <https://iea.blob.core.windows.net/assets/6b2fd954-2017-408e-bf08-952fdd62118a/Electricity2024-Analysisandforecastto2026.pdf>

²⁰⁶ <https://www.gartner.com/en/newsroom/press-releases/2024-11-12-gartner-predicts-power-shortages-will-restrict-40-percent-of-ai-data-centers-by-20270>

fonctionner des serveurs optimisés pour l'IA atteindra **500 térawattheures (TWh) par an en 2027**, soit 2,6 fois plus qu'en 2023. En conséquence, Gartner prévoit que 40 % des centres de données d'IA existants seront soumis à des contraintes opérationnelles liées à la disponibilité de l'électricité d'ici 2027.

Power required for AI data centers to run newly added AI servers per year



Indicateur : le Power Usage Effectiveness, dominant mais contesté


Le Power Usage effectiveness, ou PUE, est l'indicateur d'efficacité énergétique le plus utilisé dans l'industrie. Initialement proposé par le consortium The Green Grid²⁰⁷, il est devenu une norme de facto dans le secteur. Le décret tertiaire s'appuie également sur cette métrique pour demander une meilleure efficacité énergétique aux centres de données. Derrière un indicateur aux apparences simples se cache des pratiques de calculs diverses et peu représentatives, ce qui n'aide en rien à la transparence du secteur.

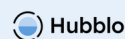
Ce qui dit la norme EN 50600-4-2

La norme EN 50600-4²⁰⁸ (et son équivalent ISO/IEC 30134) définit les KPI d'efficacité utiles au niveau d'un centre de données. Parmi ces indicateurs se trouve le Power Usage Effectiveness ou PUE, qui est l'indicateur le plus communiqué par les centres de données lorsque sont abordées les thématiques énergétiques et environnementales. La norme EN 50600-4-2 définit le PUE et les méthodes de calcul associées, les autres normes composant l'EN 50600-4 définissent les autres indicateurs d'efficacité concernant les centres de données et les méthodes de calcul de chacun d'eux : Water Usage Effectiveness, Carbon Usage Effectiveness, IT Energy Efficiency for servers, IT Energy Utilization for servers, Renewable Energy Factor, Energy Reuse Factor et Cooling Efficiency Ratio.

²⁰⁷ The Green Grid

²⁰⁸ <https://www.boutique.afnor.org/fr-fr/norme/nf-en-5060041/technologie-de-linformation-installation-et-infrastructures-de-centres-de-t/fa186315/58682>

Water & GWP	EN 50600-4-8 ISO/IEC 30134-8 : Carbon Usage Effectiveness / CUE	EN 50600-4-9 ISO/IEC 30134-9 : Water Usage Effectiveness / WUE	 <div> Évaluation Norme World </div>	
IT Efficiency	EN 50600-4-4 ISO/IEC 30134-4 : IT Energy Efficiency for servers / ITEEsv	EN 50600-4-5 ISO/IEC 30134-5 : IT Energy Utilization for servers / ITEUsv		
Energy	EN 50600-4-2 ISO/IEC 30134-2 : Power Usage Effectiveness / PUE	EN 50600-4-3 ISO/IEC 30134-3 : Renewable Energy Factor / REF	EN 50600-4-6 30134-6 : Energy Reuse Factor / ERF	EN 50600-4-7 : Cooling Efficiency Ratio / CER
Definitions	EN 50600-4-1 ISO/IEC 30134-1 : Overview of & general requirements of KPIs			



Récapitulatif des composantes de la norme EN 50600-4²⁰⁹

En ce qui concerne le PUE, la formule de base est la suivante :

$$\text{PUE} = \frac{\text{Total Facility Energy}}{\text{IT Equipment Energy}}$$

Le dénominateur "IT Equipment Energy" comprend : la consommation d'énergie des équipements IT présents dans le centre de données, mais aussi les équipements nécessaires à la maintenance comme les KVM (Keyboard-Video-Mouse), les écrans utilisés par les responsables de maintenance présents sur place, les ordinateurs fixes et portables, ainsi que tout autre équipement numérique impliqué dans le contrôle et le pilotage du centre de données.

Le numérateur "Total Facility Energy" comprend tout ce qui est comptabilisé dans "IT Equipment Energy", mais aussi la consommation :

- des équipements destinés à délivrer la puissance électrique : UPS (batteries de secours), commutateurs électriques, générateurs (groupes électrogènes), les PDUs (power distribution unit), batteries, ainsi que la perte énergétique induite
- des équipements de refroidissement: chillers, cooling towers, pumps, computer room air-handling units (CRAHs), computer room air-conditioning units (CRACs), and direct expansion air-handler (DX) units
- du reste, comme l'éclairage, la sécurité incendie

Au-delà de cette formule générique, l'EN 50600-4-2 indique 3 niveaux de calcul pour le PUE :

1. Un PUE calculé est de **niveau 1**, lorsque la consommation d'énergie IT est **mesurée directement à la sortie des UPS** (Uninterruptible Power Supply). Ce niveau de calcul est moins représentatif de la réalité puisque la perte énergétique induite, entre les UPS et les PDU finaux (dans les baies), est comptée comme de la consommation d'énergie IT.
2. Un PUE calculé est de **niveau 2**, lorsque la consommation d'énergie IT est **mesurée entre les PDU principaux (en amont des baies) et les PDU finaux (dans les baies)**. Ce niveau de calcul donne un PUE **plus réaliste** que le niveau 1.
3. Un PUE calculé est de **niveau 3**, lorsque la consommation d'énergie IT est **mesurée entre les PDU finaux (dans les baies) et les équipements informatiques**. Ce niveau est le **plus fin** proposé dans

²⁰⁹ [Datacenters, Cloud et environnement : S'orienter dans la jungle des référentiels, Hubblo](#)

la norme et permet une plus grande transparence sur la part de consommation d'énergie effectivement due aux équipements IT.

Il est également précisé dans la norme, qu'un PUE annoncé inférieur à 1.5, se doit d'être calculé avec un niveau de calcul PUE 2 et qu'un PUE annoncé inférieur à 1.2, se doit d'être calculé avec un niveau calcul PUE 3.

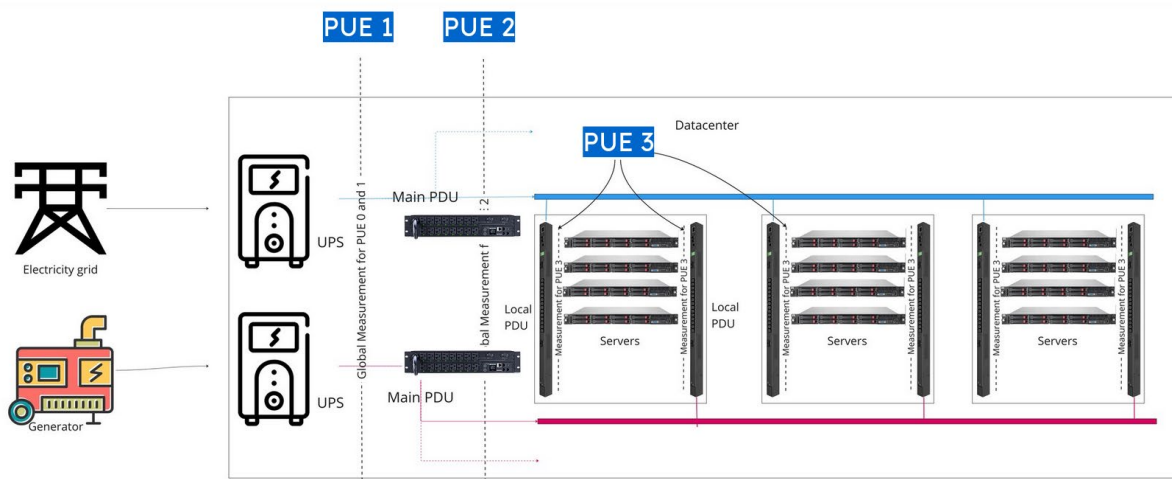


Schéma illustrant les trois positions de mesure de la consommation d'énergie IT (dénominateur du PUE), selon le niveau de précision demandé par la norme : 1, 2 ou 3.²¹⁰

La norme précise également que les relevés de puissance doivent être **effectués au moins plusieurs fois dans l'année**, pour que la moyenne annuelle de PUE annoncée soit la plus représentative possible de la variabilité du calcul au cours d'une année.

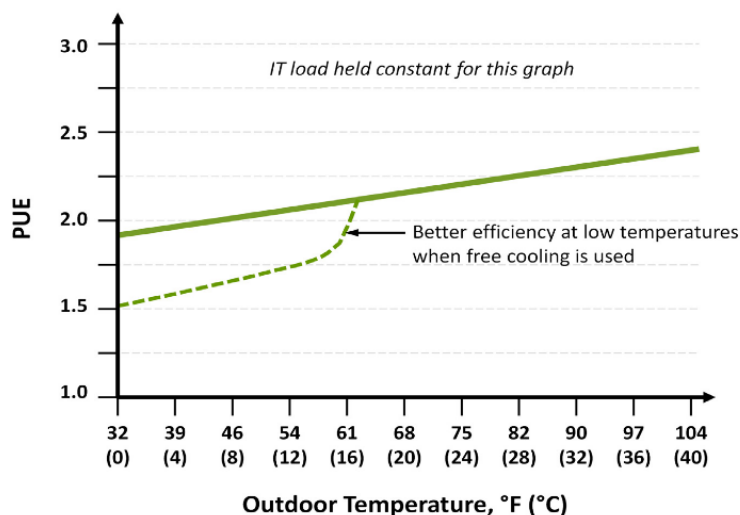
La variabilité du PUE

Comme tout indicateur d'efficacité et puisque le dénominateur du PUE est la consommation d'énergie de l'ensemble des équipements IT, le **PUE est variable** pour plusieurs raisons.

D'abord **la charge des équipements IT est variable dans le temps**, classiquement plus importante le jour que la nuit, avec des pics de charge difficilement prévisibles en fonction des usages et des services déployés. Cette variabilité impacte donc la variable "IT Equipment Energy" dans le calcul présenté précédemment. Dit autrement, **il est possible et courant** pour un centre de données, **de consommer momentanément plus d'énergie IT** qu'à la normale et **ainsi de pouvoir revendiquer un meilleur PUE**.

En fonction du type de refroidissement, la température extérieure peut influencer largement le numérateur "Total Facility Energy", le PUE est généralement moins bon en été qu'en hiver, surtout si le centre de données s'appuie en partie sur du [Free Cooling](#).

²¹⁰ [Datacenters, Cloud et environnement : S'orienter dans la jungle des référentiels, Hubblo](#)



Graphique représentant la variabilité du PUE en fonction de la température extérieure²¹¹

Un indicateur à communication variable

Malgré les lignes de conduites définies dans la norme EN 50600-4-2 et la variabilité du PUE, les pratiques constatées dans le secteur des centres de données concernant le calcul et l’affichage de cet indicateur semblent souvent alimenter la confusion.

En premier lieu, la périodicité des relevés puissances électriques ne semble pas toujours respectée. **Un relevé unique annuel de la puissance IT et de la puissance demandée par le centre de données dans son ensemble ne permet pas de refléter la variabilité de ces valeurs.** Cette pratique, appelée parfois “PUE 0”, ne semble pourtant pas exclue explicitement. De plus, il est fréquent de lire des PUE annoncés de manière permanente, qui sont des PUE calculés en amont par le constructeur du centre de données. **Ces “PUE constructeur” sont des estimations du PUE atteignable au cours du fonctionnement du centre de données de la part du fabricant, ce qui ne reflète en rien la réalité constatée une fois le centre de données en activité.** Pourtant, ces valeurs sont parfois considérées suffisantes pour être affichées, parfois de manière pérenne sans réévaluation au-delà de la première année d’exploitation.

La divergence avec la norme EN 50600-4 et les largesses d’utilisation du PUE rendent complexe une interprétation fine de l’évolution de l’efficacité énergétique du secteur. De nombreux centres de données affichent un PUE de 1.2, ce qui signifie en théorie qu’un calcul de type PUE 3 a été effectué, mais **seuls de rares acteurs font certifier la méthode de calcul de leur PUE**²¹². Il faut également noter que dans le contexte de la colocation, relever la puissance électrique au plus proche des serveurs (pour calculer un PUE de niveau 3) semble **impossible** du point de vue de l’opérateur du centre de données, puisque ce sont **les clients qui ont l’accès** exclusif aux baies contenant leur matériel (hors demande explicite de télémaintenance).

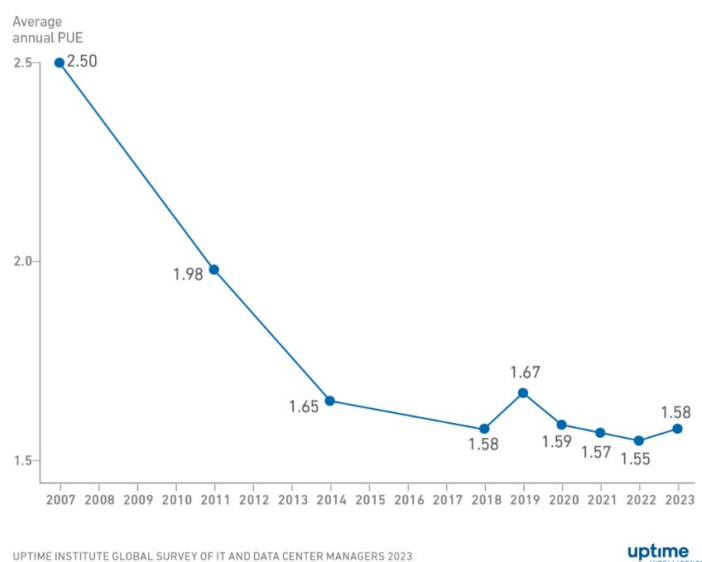
D’un côté les promesses d’une efficacité énergétique en constante amélioration dans le secteur sont nombreuses. De l’autre, l’Uptime Institute indique qu’un plateau semble avoir été atteint en ce qui concerne le PUE moyen, autour de 1.5 (cf. graphique ci-dessous).

²¹¹ PUE, a comprehensive examination of the metric, ASHRAE, The Green Grid

²¹² OVHCloud fournit sur son site web un rapport d’audit de la méthode de calcul de leurs indicateurs d’efficacité

Figure 1. Industry-wide PUE improvement slows, then stops

What is the average annual PUE for your data center? (n=567)



Graphique des résultats de sondage de l'Uptime Institut sur le PUE moyen de ses répondants²¹³

Vers une plus grande transparence ?

Le PUE est bien souvent mis en avant pour justifier le moindre impact environnemental d'un centre de données, d'après son constructeur ou son exploitant. Outre le fait que le PUE ne représente que l'efficacité énergétique et non tous les impacts environnementaux, il semble problématique qu'autant de calculs, projections et articles de presse se contentent de cette métrique. Sa variabilité et les disparités de son utilisation plaident pour une transparence bien plus grande du secteur, par exemple en affichant des métriques absolues de consommation d'énergie sur les périmètres IT et bâtiment. C'est en ce sens que s'oriente l'acte délégué de mars 2024 de **l'Energy Efficiency Directive de la Commission Européenne**.²¹⁴ Ce texte demande, entre autres, aux États membres, de collecter un certain nombre d'informations auprès des centres de données présents sur leur territoire, comprenant notamment la puissance IT demandée et la consommation d'énergie globale du centre de données, même si les données destinées à être affichées publiquement ne concernent *a priori*, une nouvelle fois, que le PUE (moyen et agrégé sur un ensemble de centres de données d'une même taille).

État du réseau électrique

Le SDDR 2025

Un réseau électrique sous-utilisé

En 2024, RTE annonçait que la puissance de l'ensemble des projets de centres de données construits, en cours et à l'étude en Ile-de-France atteint 7,5 GW, soit le niveau de puissance nécessaire à l'alimentation de l'ensemble de la Métropole du Grand Paris. Dans le SDDR 2025, RTE note cependant que de nombreux

²¹³ Uptime Institute, "Global PUEs: Are they going anywhere ?"

²¹⁴ https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=OJ:L_202401364

projets tardent à se réaliser, ces chiffres ont donc évolués : le SDDR 2025 cite "40 nouveaux projets pour une puissance complémentaire de 5,3 GW qui disposent déjà de droits d'accès au réseau."

Enfin, le SDDR met également en avant de nouveaux principes concernant les industries et les centres de données :

- "Publication des zones favorables au développement de projets de forte puissance de consommation (250 MW/750 MW). Dans ces zones, le développement du réseau se limitera aux seules infrastructures de raccordement sur le réseau 400 kV (leur construction sera d'autant plus rapide que les sites industriels seront situés à proximité des nœuds électriques concernés).
- Identification de zones prioritaires car elles bénéficient d'un régime d'investissement spécifique et partagé entre RTE et les industriels"
- Introduction d'une procédure spécifique pour des raccordements de sites de très forte puissance, sous réserve de rapidité dans la délivrance des autorisations administratives (3 ans maximum pour garantir une pleine puissance à 1 GW)
- Modification du cadre de raccordement pour empêcher les phénomènes de spéculation sur le réseau (réservation de puissance pour des projets fictifs).
- Augmentation du taux de préfinancement par RTE des équipements et matériels électriques nécessaires au raccordement des projets industriels, sous réserve qu'il s'agisse de matériels standards.

Des effets de saturation

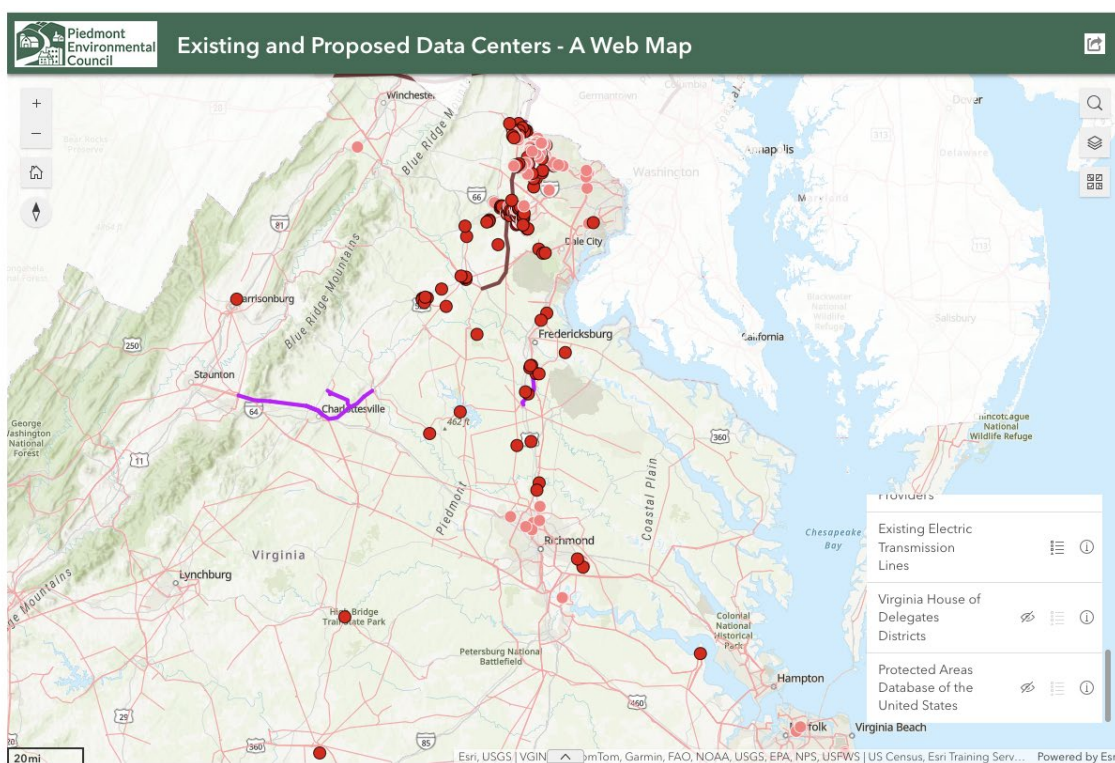
Les projets de centres de données ont en moyenne une puissance variant de 50 à 200 MW, contre 10 MW pour un projet industriel classique. Le projet le plus puissant en Île-de-France atteint à ce jour 1,4 GW (Fouju, 77). La logique qui consiste à traiter les demandes de raccordement au fur et à mesure de leur arrivée selon les règles en vigueur jusque-là pose question en accordant à un seul projet une capacité aussi importante. L'augmentation du besoin de puissance anticipée en Île-de-France dans le cadre de la transition écologique nécessite un renforcement du réseau, notamment au sud de la région. Les demandes de raccordement de centres de données accélèrent ce phénomène : le renforcement du réseau devra ainsi intervenir plus tôt qu'anticipé. En se concentrant dans certaines zones (autour de Villejust et sur un arc allant de Saint Denis au Bourget), **ces projets participent fortement à saturer des poches électriques déjà en tension**. Le montant global des coûts de raccordement sera très élevé (de l'ordre de plusieurs centaines de millions d'euros). Enfin, les besoins liés à la réindustrialisation (exemple : site de Villaroche, 77) doivent également être anticipés dans le cadre d'une réponse globale pour que les adaptations du réseau ne retardent pas le développement des activités. C'est ce que RTE propose dans son dernier SDDR.

À titre de comparaison, on peut noter que ce besoin de renforcement des capacités de transmission électrique²¹⁵ se retrouve ailleurs dans le monde. Un exemple notable est celui de la Virginie du Nord aux États-Unis - plus grande concentration mondiale de centres de données - où cette énorme croissance suppose de nouveaux investissements, localement contestés ([voir plus bas](#)). Dominion Energy déclare avoir reçu des commandes de clients qui pourraient doubler la capacité des centres de données en Virginie d'ici 2028, avec une taille de marché prévue de 10 gigawatts d'ici 2035. Cela représente une

²¹⁵ <https://www.datacenterfrontier.com/energy/article/33013010/dominion-virginias-data-center-cluster-could-double-in-size>

augmentation considérable par rapport à la consommation électrique actuelle des centres de données, qui a atteint 2,67 gigawatts en 2022, et 4 gigawatts en 2024²¹⁶. Aujourd'hui, ce secteur de la Virginie du Nord (comtés de Loudoun, Prince Williams, Fairfax) compte environ 300 centres de données. Il faut cependant rappeler que le réseau électrique étasunien est à la fois moins interconnecté et plus fragile que le réseau français.

Sur la carte suivante, les renforcements de réseaux sont indiqués en marron et violet. Les points rouge vif sont des projets de centres de données, les rouges pâles existent déjà. Cette carte a été réalisée par le Piedmont Environmental Council (PEC) qui a été fondé en 1972 et a pour mission "de protéger et de restaurer les terres et les eaux du Piémont de Virginie, tout en construisant des communautés plus fortes et plus durables."



<https://www.pecva.org/region/culpeper/existing-and-proposed-data-centers-a-web-map/>

Cette situation se retrouve aussi en Irlande²¹⁷ où les générateurs de secours prennent le relais d'un réseau en peine, et dans d'autres États des Etats-Unis (voir tableau ci-dessous). Cette saturation est emblématique de la façon dont la concentration des centres de données vient bousculer la robustesse des réseaux électriques.

²¹⁶ <https://wtop.com/business-finance/2024/03/northern-virginia-again-ranks-no-1-data-center-market-but-challenges-are-mounting/>

²¹⁷ <https://www.thejournal.ie/investigates-data-centres-6554698-Nov2024/>

Table 1. Opportunities and challenges for states ranked in the Top 15 for data center growth [29, 31, 32, 33, 34, 35, 36, 37]

STATE	OPPORTUNITIES	CHALLENGES
Virginia	Unparalleled network infrastructure; proximity to federal government agencies	Community pushback; regulatory scrutiny, particularly concerning environmental impact; transmission
Texas	Business-friendly; ample land availability	Electric grid reliability and pace of expansion
California	Robust technological ecosystem	High real estate and power costs; stringent environmental regulations
Illinois	Strategic location; significant tax incentives; nuclear generation and increasing renewable energy investments to address sustainability	Transmission constraints and rapidity of development
Oregon	Low electricity rates, low carbon emissions, moderate climate, tax incentives, and skilled workforce	Complex environmental regulations; demand for green energy solutions, and pace of growth
Arizona	Solar electricity, low risk of natural disasters; recent market growth	Water scarcity; need for sustainable cooling solutions
Iowa	Low electric rates; renewable energy availability	Geographic limitation (distant from major U.S. data hubs)
Georgia	Availability of land and power; friendly business environment	Balancing rapid expansion with local resource impacts
Washington	Abundant renewable energy resources	High energy costs; strict regulatory measures
Pennsylvania	Strategic location near major East Coast markets; relatively low energy costs	Aging infrastructure
New York	Hub for financial services; high connectivity demand	Space limitations; high energy costs
New Jersey	Close to major metropolitan areas; robust fiber-optic infrastructure ; transmission capacity from recent build out	High property and energy costs
Nebraska	Low energy costs; generous tax incentives	Remote location might limit connectivity options
North Dakota	Significant tax incentives; low cost of operations	Limited connectivity; need for more robust infrastructure
Nevada	Tax abatements; low electricity prices	Water scarcity; need for sustainable cooling solutions

Dans ce tableau du rapport de l'EPRI "Powering intelligence. Analyzing Artificial Intelligence and Data Center Energy Consumption", plusieurs défis ont trait au réseau : fiabilité du réseau, contraintes de transport de l'électricité, notamment pour la Virginie, l'Illinois, ou le Texas. La carte ci-dessous, issue du même rapport, complète le propos.

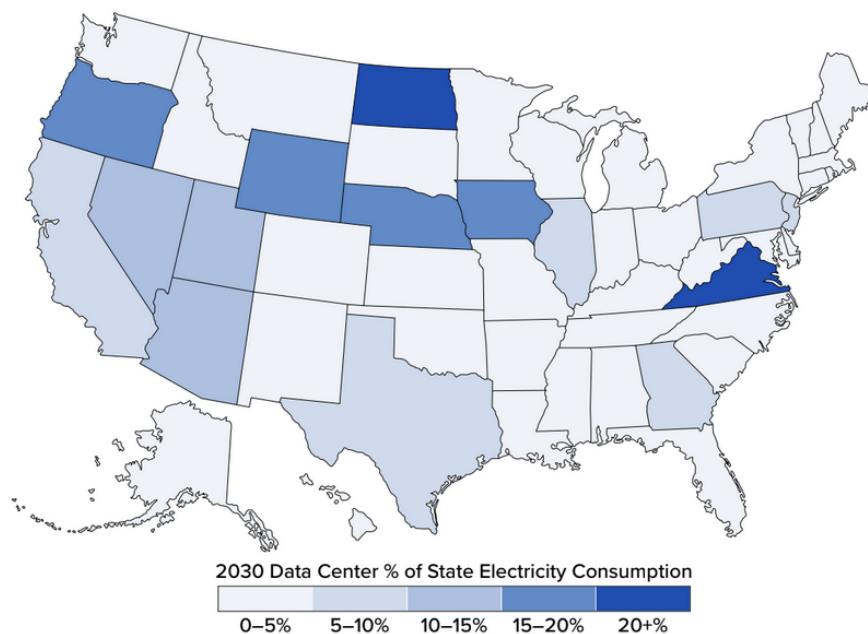


Figure 8. 2030 projected data center share of electricity consumption (assumes average of the four growth scenarios and that non-data center loads grow at 1% annually) [4, 8, 9]

Un réseau électrique sous-utilisé

La capacité réservée par un centre de données n'est en général pas utilisée en totalité, ou tout du moins, pendant de longues années, avant la mise en œuvre à pleine puissance du site. Pendant cette période cependant, cette puissance "réservée" ne peut bénéficier à d'autres utilisateurs, RTE devant maintenir disponible cette capacité pour le réservataire.

L'implantation rapide et non concertée des centres de données en Île-de-France a des impacts négatifs sur le réseau électrique :

- **Une désoptimisation du réseau** : l'implantation concentrée de centres de données dans des zones restreintes nécessite des renforcements de réseau tandis que des capacités sont disponibles ailleurs. Le manque d'anticipation et de planification conduit aussi à augmenter le linéaire des raccordements, coûteux et générateur de nuisances ;
- **Un frein à l'installation de nouveaux clients consommateurs** : outre la question du foncier, en accaparant une partie de plus en plus importante du réseau électrique, les centres de données rendent plus difficile l'implantation d'autres activités consommatrices d'électricité en augmentant les coûts et les délais de raccordement (centres de données versus installations stratégiques telles que des électrolyseurs, site de batteries de stockage, des gigafactories, etc.).

Cette situation de saturation et de congestion se retrouve dans plusieurs pays européens, et pousse à des investissements très importants qui ne sont pas toujours mutualisés, rationnels ou optimisés, tout en pesant en partie sur les contribuables locaux.

RTE envisage donc de mettre un système appelé par le monde anglo-saxon "use it or loose it", où la puissance de raccordement du centre de données sera ajustée à sa consommation réelle.

RTE a aussi annoncé début 2025 la mise en place d'un dispositif dit "fast track", destiné à accélérer le raccordement des infrastructures stratégiques, notamment les centres de données. Cet accord vise à raccourcir significativement les délais pour les projets considérés comme prioritaires pour la souveraineté numérique ou la transition énergétique, en allégeant certaines étapes procédurales et en mobilisant des moyens techniques spécifiques.

Les enjeux en matière énergétique

Le pilotage énergétique

Au-delà du bouclage en énergie annuelle, l'équilibre offre-demande repose sur une adéquation entre production et consommation à tout instant. Ainsi, même s'il y a des surplus en énergie, ils ne sont pas forcément localisés temporellement au même moment que la consommation et un déficit de marge en puissance peut apparaître. C'est ce déficit de marge à 2030 dont parle le Bilan Prévisionnel 2023 de RTE, en précisant qu'il est évitable en activant plusieurs leviers (cf. synthèse : [Bilan prévisionnel - Édition 2023](#)) : niveau du nucléaire, maintien de thermique, développement de moyens de flexibilité de la demande... C'est à ce titre que la demande des centres de données peut poser problème si elle est non flexible.

Il est intéressant de noter qu'il existe des projets comme [DCFlex | EPRI Micro Sites](#) auquel participe RTE pour encourager la flexibilité des centres de données. Cette flexibilité paraît compliquée pour certains acteurs type colocation et reste aujourd'hui non démontrée : c'est à un point de débat.

Par ailleurs, les charges de travail d'IA générative, ont un profil de charge particulier, présentant des creux et des pics de demande de puissance dans des délais très courts, jusqu'à quelques secondes.

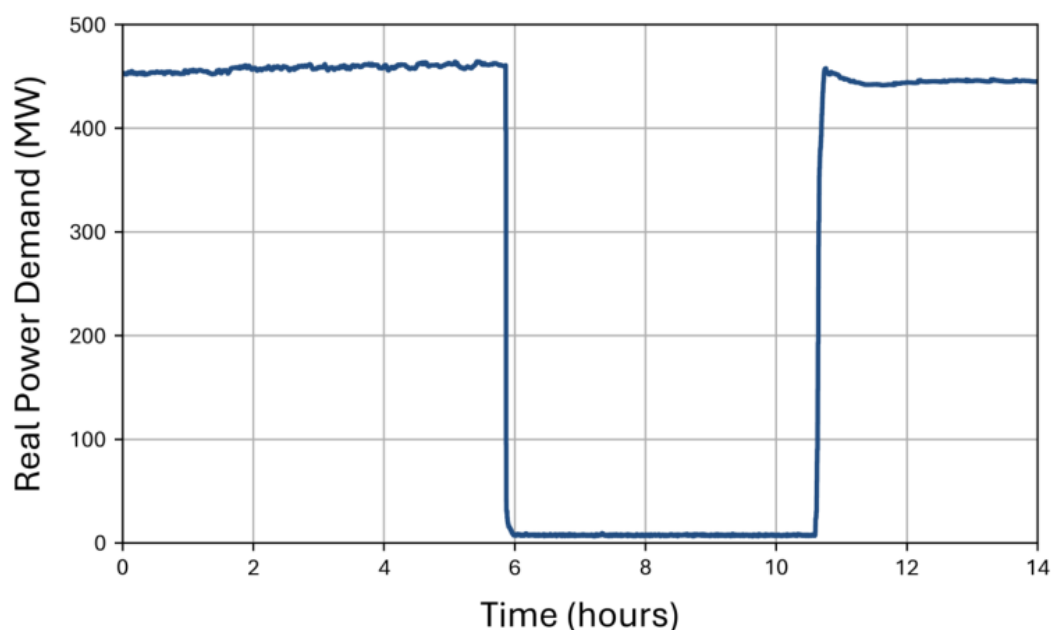


Figure 3.1: Data Center Load Ramp Down and Up

Graphique issu du livre blanc de la North American Electric Reliability Corporation, montrant la courbe de puissance électrique demandée par un centre de données, présentant des variations de plus 400 MW en quelques dizaines de secondes²¹⁸

Le réseau électrique est opéré de manière à assurer un relatif équilibre entre production et demande. Un écart trop important et trop soudain entre les deux, peut avoir un effet à la fois sur la tension et sur la fréquence du réseau, ce qui peut entraîner des perturbations sur les équipements qui le composent, menant dans certains cas à l'arrêt total du service. Pour des raisons distinctes du sujet de ce rapport, le Texas, suite à une tempête de froid en 2021²¹⁹ et l'Espagne en 2025 après des surtensions en cascade²²⁰ sont les cas, récents, les plus emblématiques d'interruption de service d'un réseau électrique.

Si les variations citées dans le livre blanc du NERC ne sont pas à ce jour associées à des coupures réelles de service sur des grilles électriques, l'augmentation des capacités et des puissances électriques demandées par les nouveaux centres de données conçus pour l'IA gen (Colossus 1 de xAI, Stargate I de OpenAI) et les projets de futurs centres de données (Colossus II d'xAI, 5 nouveaux sites pour le projet Stargate d'OpenAI cumulant à terme 10 GW, mais aussi en France le centre de données de Fouju avec 1,4 GW), soulève la question de la capacité des acteurs du réseau électrique à encaisser des variations rapides de demande de l'ordre de plusieurs centaines de MW à plus d'1 GW.

Dans un article en rapport, SemiAnalysis expose le cas d'un centre de données de Meta ayant hébergé l'entraînement du modèle Llama 3.²²¹ Avec "seulement" 30 MW de puissance installée pour l'IT (24000

²¹⁸ [Livre Blanc du NERC sur les Large Loads](#)

²¹⁹ [What Really Happened During the Texas Power Grid Outage?, Practical Engineering](#)

²²⁰ [Black-out en Espagne et au Portugal : ce que dit l'enquête des experts européens](#), Connaissance des énergies

²²¹ [AI Training Load Fluctuations at Gigawatt-scale - Risk of Power Grid Blackout?, SemiAnalysis](#)

GPUs H100), l'entraînement de ce modèle et son profil de consommation était la source de perturbations sur le réseau électrique. Pour pallier ce problème sans solution spécifique dédiée, les ingénieurs de Meta ont développé une option dans le programme d'entraînement, intitulée "pytorch-no-powerplant-blowup=1". Ce qui signifie littéralement "demander à Pytorch (le cadriciel de développement du modèle), de ne pas exploser la centrale de production d'électricité". Cette option activée avait pour effet, lors de la fin de la phase d'entraînement, de lancer des charges de travail inutiles, de manière à réduire progressivement la puissance électrique demandée, plus lentement que ce n'était le cas sans cette option.

D'autres solutions ont vu le jour pour mitiger la problématique des variations fortes et rapides de puissance demandée dans les centres de données d'IA Gen. Tesla propose des batteries à grande capacité : les Tesla Megapacks. Cette solution permet notamment d'assurer la continuité de service du centre de données en cas de coupure, comme dans un centre de données classique, mais également de réguler la demande électrique en combinant électricité produite (soit localement soit sur le réseau) et énergie stockée.



Centre de données Colossus I d'xAI, les Tesla Megapack sont encadrées en rouge

Cette nouvelle contrainte et les solutions techniques associées posent également la question de la dépendance aux acteurs concernés à la fois par l'essor de l'IA Générative et par les solutions techniques permettant de mitiger les risques court terme d'un point de vue social et sociétal.

Les évolutions de la production électrique

La croissance des consommations d'électricité des centres de données pose la question de la production d'électricité en amont.

La demande électrique annuelle des centres de données est importante, et elle s'ajoute à un contexte d'électrification avec des ordres de grandeur encore plus vertigineux, cf. chapitre 2 du Bilan Prévisionnel 2023 de RTE, [Chapitre 2 - La consommation](#). Pour autant, si l'électrification est accompagnée par le développement des renouvelables, une réhausse du productible nucléaire existant (étant donné qu'il n'y aura pas de nouvelles centrales nucléaires à horizon 2035) et des mesures d'efficacité énergétique, les acteurs énergétiques estiment qu'il n'y aura pas de problèmes de bouclage en énergie (surtout étant donnée la nature interconnectée du système électrique européen) et même en électricité décarbonée (cf. chapitre Environnement). La France reste d'ailleurs largement exportatrice dans le scénario A du BP23 (cf. chapitre 6). Cette analyse se retrouve dans la synthèse du BP23, par exemple p38 : [Bilan prévisionnel - Édition 2023](#).

Par ailleurs, si les consommations s'emballaient très fortement, de nouvelles orientations pourraient être prises côté privé comme public :

- Développement de SMR sur le site centres de données (voir plus haut, partie [technologies](#))
- La réouverture de centrales au charbon : ce n'est pas une option souhaitable mais la crise énergétique liée à la guerre en Ukraine a par exemple conduit à la réouverture de la centrale charbon de St Avold (Moselle), et les retards de livraison de l'EPR Flamanville a repoussé la fermeture de la centrale de Cordemais.
- La prolongation au-delà de 50 ans des réacteurs des centrales nucléaires dont EDF avait prévu la fermeture : *"Les orientations affichées par le Président de la République dans le discours de Belfort et confirmées depuis intègrent la volonté de prolonger l'exploitation des réacteurs actuels dès lors que ceux-ci répondent aux normes de sûreté définies par l'ASN. La loi d'accélération du nucléaire inclut en complément des dispositions visant à simplifier la procédure de réexamen périodique des réacteurs âgés de plus de 35 ans. D'ici 2035, le défi sera notamment de s'assurer de la prolongation des réacteurs les plus anciens (palier 900 MW) au-delà de 50 ans."*²²²

La Programmation Pluriannuelle 2025 (PPE) fixe quant à elle des objectifs de production d'énergies renouvelables et de récupération plus ambitieux que la précédente²²³.

Aux Etats-Unis, dont le système électrique est évidemment très différent sur plusieurs points, l'annonce de la réouverture de la centrale de Three Mile Island en Pennsylvanie pour un usage largement dédié à Microsoft a ainsi pu faire débat. Cette centrale est d'une part connue pour un des plus gros accidents nucléaires qui a eu lieu aux Etats-Unis, mais elle souligne également les tensions entre pro-nucléaires qui mettent en avant les faibles émissions carbone tandis que les anti-nucléaires/pro-ENR indique que les premiers ne prennent pas en compte les autres impacts environnementaux du nucléaire notamment les déchets (le fameux "tunnel carbone").

Les enjeux l'architecture du réseau électrique

Le fonctionnement du réseau électrique français a été pensé pour transporter et distribuer l'électricité produite par des centrales électriques très puissantes, dans une pensée technique très centralisatrice. Le développement des ENR et de productions décentralisées vient donc fortement questionner l'architecture-même du réseau.

²²² Bilan Prévisionnel RTE édition 2025

²²³ https://concertation-strategie-energie-climat.gouv.fr/sites/default/files/2024-11/241104_Projet%20de%20Programmation%20pluriannuelle%20de%20l%27%C3%A9nergie%203%20VFF.pdf

Des réflexions similaires sont entamées aux Etats-Unis. Le think tank EPRI prône ainsi dans son rapport *Powering intelligence. Analyzing Artificial Intelligence and Data Center Energy Consumption*, une coordination rapprochée entre centres de données et compagnies électriques, pour mieux intégrer les productions d'énergie décentralisées que ces opérateurs pourraient être amenés à introduire dans le système. Si les productions décentralisées sur site ou à proximité sont plus développées aux Etats-Unis (fermes solaires, barrages hydrauliques, éoliens, voire SMR à venir) la tendance reste plus discrète en France, sauf pour les SMR semblerait-il.

En France, les opérateurs de centres de données peuvent investir dans des PPA pour verdir leur mix énergétique mais ces PPA sont la plupart du temps non localisés. Les productions d'ENR se développent mais pas de leur fait, plutôt d'opérateurs énergétiques classiques.

Cette coopération rapprochée citée par l'EPRI pourrait, selon le think tank, se décliner sur l'utilisation des générateurs de secours pour soutenir le réseau, ou les productions résiduelles de production locales notamment de SMR.

Sur le premier sujet, dans l'ouvrage *Sous le feu numérique* (Diguët, Lopez), était décrit le système mis en place par la compagnie Portland General Electric (Oregon), s'assurant sa réserve électrique de secours grâce aux générateurs de secours des DC, mais aussi d'entrepôts alimentaires, d'hôpitaux et d'autres acteurs industriels. Ces démarches d'effacement, également évoquées en France au plus haut de la crise énergétique en 2022, a cependant le grave inconvénient de polluer l'air par la combustion de fuel. En revanche, cela pourrait être envisagé avec les batteries électriques servant également de secours.

Sur le second sujet, sur les principes de l'autoconsommation électrique, on peut imaginer qu'un centre de données qui aurait une production locale d'énergie pourrait remettre un surplus dans le réseau. Cela dépendra évidemment de ses propres besoins, besoins qui sont actuellement en croissance permanente.

Tout cela amène à penser aussi une gestion différente du réseau, qui peut se déployer sur plusieurs axes :

- De meilleurs outils de modélisation de la grille électrique ;
- Une réflexion plus radicale sur la transformation de l'architecture du réseau national, en lien avec la grille européenne.

Les enjeux de mutualisation

RTE doit traiter les demandes de raccordement dans l'ordre d'arrivée des opérateurs, ce qui aboutit régulièrement à un manque d'optimisation des investissements à réaliser pour renforcer le réseau. Pour régler ce problème, RTE a donc proposé à la Commission de Régulation de l'Energie (CRE) de déployer un nouveau dispositif mis en place par la loi APER : l'ORAM, ou Offre de RAccordement Mutualisée, testé sur le secteur de Ciroliers, au sud de l'Essonne. Ce dispositif permet à RTE de déployer des zones d'accueil mutualisées pour les installations de consommation, et, lorsque cela est pertinent sur le plan technico-économique, d'anticiper le développement de l'infrastructure électrique dans un secteur donné et **d'en répercuter équitablement les coûts aux consommateurs qui bénéficient de la capacité de raccordement au réseau ainsi créée.**

Dans le secteur de Ciroliers, 900 MW de centres de données sont en cours de raccordement, et 900 MW de demandes complémentaires ont été reçues pour des DC. RTE souligne : "Cette ORAM permettra de créer 1200 MW de capacité en investissant 250 M€ dans le réseau pour créer deux postes électriques à

225 kV (Bondoufle et Quartier Saint Louis), modifier le poste 225 kV de Sénart, ajouter des autotransformateurs 400 / 225 kV dans les postes de Cirolliers et du Chesnoy et faire évoluer le réseau 225 kV dans la zone. **Les clients consommateurs qui bénéficieront de cette capacité créée devront s'acquitter d'une quote part de 147,71 k€/MW (également approuvée par la CRE) correspondant à ces investissements, en sus du coût de leurs ouvrages propres de raccordement."**

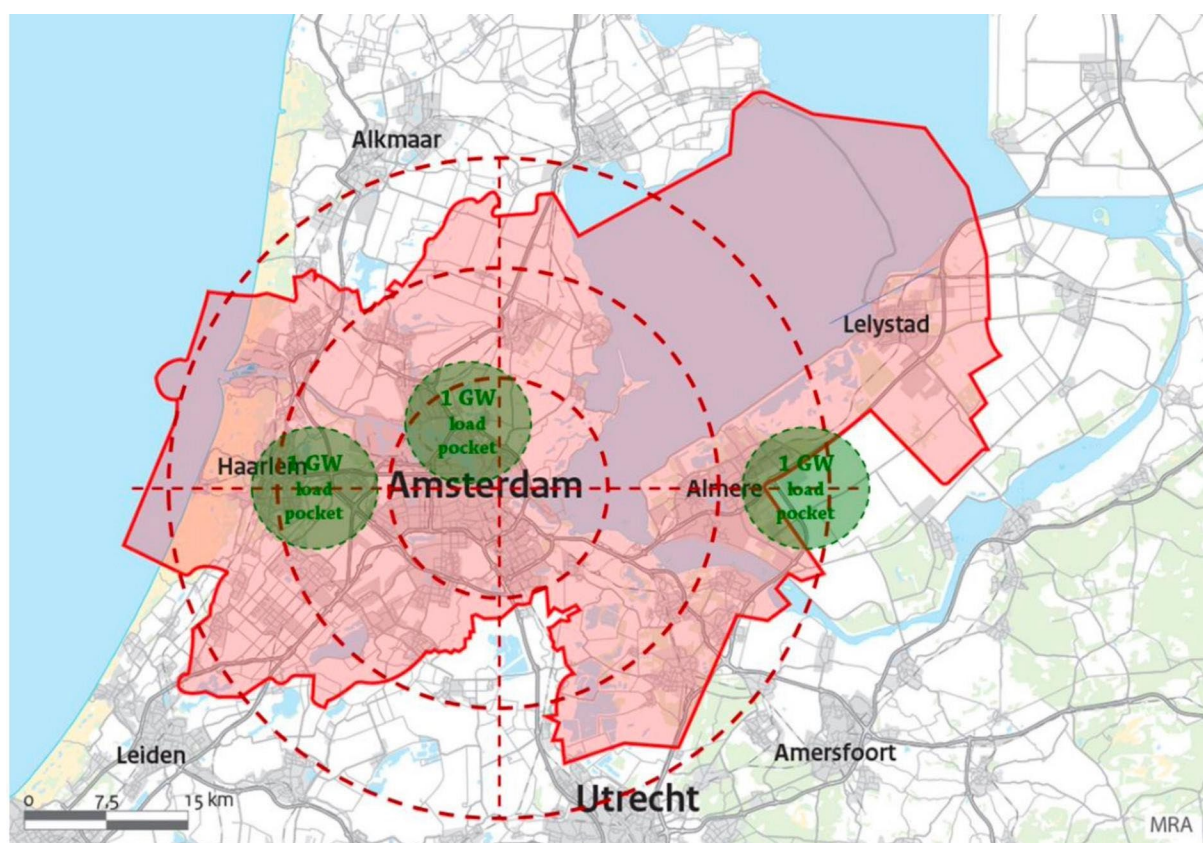
Cette approche mutualisée présente plusieurs intérêts :

- Rendre le réseau électrique plus efficace et robuste au lieu de le faire évoluer au coup par coup, sans vision d'ensemble ;
- Minimiser le bilan écologique du renforcement de réseau : ressources et matières utilisées, réduction du linéaire global d'ouvrages électriques à créer, emprise spatiale restreinte ;
- Limiter la gêne occasionnée par les travaux.

Planification à Amsterdam, tentatives en Ile-de-France

Les municipalités d'Amsterdam, Haarlemmermeer et Hollands Kroon ont décidé d'orienter l'implantation et le développement des centres de données sur leurs territoires pour la période 2022 – 2024, après un moratoire qui leur a permis de mettre les acteurs autour de la table et décider de la meilleure stratégie commune.

Elles ont ainsi encadré l'implantation des centres de données dont la parcelle est supérieure à 2 000 m² et la puissance électrique raccordée supérieure à 5 MW. Le plan de zonage prévoit l'implantation de nouveaux centres de données ou l'extension de centres de données existants **seulement dans les zones d'exploitation identifiées pour la clusterisation des centres de données** et s'ils répondent aux conditions d'implantation conformes à la politique locale et aux conditions associées.



Possibles emplacements d'un nouveau cluster dans la région métropolitaine. Source - D'Cision- "Kansen en bedreigingen voor datacenters in de Metropool Regio Amsterdam (MRA)" Zwolle, avril 2019

La liste des conditions d'implantation est évolutive, en fonction du contexte national et européen. Par ailleurs, des règles d'implantation ont également été établies à l'échelle nationale conduisant, à ce stade, à **l'interdiction de l'implantation de centres de données sur une parcelle supérieure à 10 ha et d'une puissance électrique de plus de 70 MW dans les zones municipales.**

Sur les dimensions Energie, les conditions sont les suivantes (l'eau est présentée dans la partie suivante) :

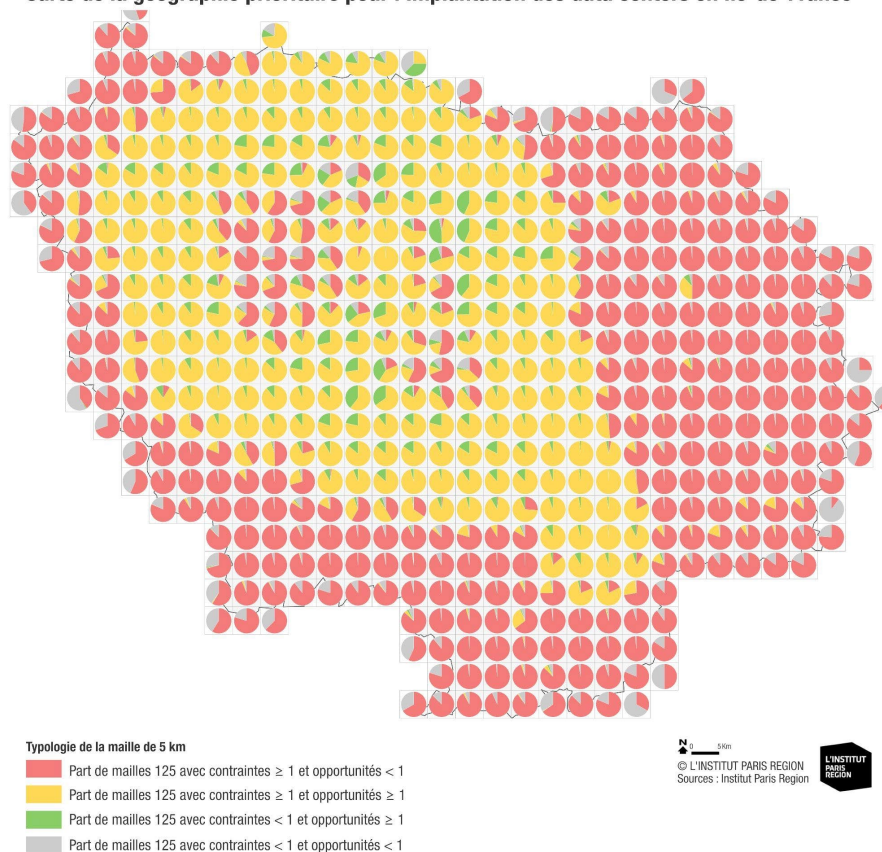
- Disposer d'un **système d'enregistrement et de suivi des consommations énergétiques**. Ce relevé est à garder au moins 5 ans et **doit être mis à disposition de l'autorité compétente**. Il doit donner de la visibilité sur : la consommation totale des centres de données, la consommation totale des équipements techniques, la consommation des équipements informatiques, le rendement des panneaux solaires, la récupération de la chaleur fatale.
- Prévoir toutes les façades aveugles et le toit en couleur claire ou végétalisées afin de limiter l'absorption de la chaleur.
- Présenter le calcul du PUE dont **la valeur cible maximale admise est 1,16**.
- Installer, dans la mesure du possible, des panneaux solaires sur les toits, les parkings et les zones non construites. Un plan d'implantation des panneaux solaires doit être présenté.
- Étudier la possibilité d'utiliser l'énergie éolienne ou solaire si le centre de données se trouve dans la zone de recherche de l'énergie solaire et/ou éolienne ou à proximité immédiate de cette zone.
- Consommer de l'électricité durable. L'opérateur de centre de données doit préciser son fournisseur d'électricité.
- **Mettre à disposition la chaleur fatale** émise pour récupération. Si aucun besoin en récupération de la chaleur fatale n'est émis, l'opérateur de centre de données doit au moins prévoir techniquement la possibilité de récupérer ultérieurement la chaleur fatale émise et de se connecter à un réseau de chaleur urbain.
- **Les nouveaux centres de données d'une puissance électrique supérieure à 80 MVA doivent disposer d'une connexion à un poste source de 150 kV situé à moins de 1 500 m.**
- **Mettre à disposition d'un tiers la capacité électrique excédentaire.**

En Ile-de-France, l'étude de l'Institut Paris Région, publiée en septembre 2023²²⁴, proposait de nombreux outils à la région Ile-de-France, l'État et aux territoires pour planifier l'implantation centres de données et les orienter vers les secteurs géographiques les moins inadaptés. Pendant la production du rapport, le Schéma directeur régional d'Ile-de-France (SDRIF) était discuté, mais la région n'a pas souhaité aller plus loin qu'une orientation réglementaire non-opposable sur la récupération de chaleur fatale et la consommation de foncier, OR qui est cependant la première à évoquer le sujet des centres de données dans un document de planification urbaine régionale.

Une carte (voir ci-dessous) a été notamment produite à la maille, indiquant par une analyse multicritère dont on trouve le détail dans l'étude, les zones les plus adaptées à l'accueil des DC. Ce travail s'est affiné par la suite avec RTE, en fonction des disponibilités électriques de ses postes sources, travail qui reste à ce jour confidentiel.

²²⁴ <https://www.institutparisregion.fr/nos-travaux/publications/le-developpement-des-data-centers-en-ile-de-france/>

Carte de la géographie prioritaire pour l'implantation des data centers en Île-de-France



Récupération de chaleur fatale, de rares projets sur les grands centres de données

État actuel

France

La stratégie française Énergie-Climat²²⁵ vise à multiplier par cinq la récupération de chaleur fatale livrée dans les réseaux, et de valoriser 25 à 29 TWh de chaleur fatale d'ici 2035. En 2021, 898 réseaux de chaleur étaient déployés en France. En dix ans, ils ont contribué à réduire de presque moitié les émissions de gaz à effet de serre, comparativement à l'emploi de chaudières individuelles au gaz.

Les centres de données présentent des gisements importants de chaleur fatale. En 2020, le potentiel de chaleur fatale récupérable via les centres de données était estimé à 1 TWh, soit 100 000 équivalents logements. En 2030, compte tenu de l'augmentation du nombre de centres de données, le potentiel de chaleur fatale récupérable pourrait représenter 3,5 TWh.²²⁶

²²⁵ https://www.ecologie.gouv.fr/sites/default/files/documents/23242_Strategie-energie-climat.pdf

²²⁶ Voir les Cahiers Techniques Valoriser sur le territoire la chaleur fatale des data centers. ADEME

C'est pourquoi la directive européenne sur l'efficacité énergétique (directive (UE) 2023/1791)²²⁷ exige que les centres de données de plus de 1 MW valorisent leur chaleur fatale, sauf si cela n'est pas viable techniquement ou économiquement. En parallèle, tout comme le Code de Conduite de l'Union européenne pour l'efficacité énergétique des centres de données, la norme, non-obligatoire, « Installation et infrastructure de centre de traitement de données » (ISO 50600-2-1 juillet 2021) intègre désormais les synergies énergétiques territoriales.

L'enjeu de la récupération de la chaleur fatale fait l'objet d'une attention plus forte de la part des pouvoirs publics, comme le souligne l'adoption de l'article 28 de la loi du 15/11/2021 visant à réduire l'empreinte environnementale du numérique en France.²²⁸ Cet article impose aux centres de données de valoriser la chaleur fatale ou de respecter un indice de performance énergétique si cela est impossible (texte réglementaire en attente). L'EED inclut aussi de nouvelles obligations. Bien que cette pratique ne permette pas de réduire en valeur absolue les impacts et la consommation d'énergie, elle permet de substituer en théorie une partie de l'énergie nécessaire à un autre bâtiment et une autre entité.

Ile-de-France

Parallèlement, dans le cadre de la procédure d'agrément, une fiche repère produite par la DRIEAT Ile-de-France²²⁹ insiste sur l'obligation de produire une étude exhaustive relative à la possibilité de récupérer la chaleur fatale du projet. La réalisation de cette étude **n'impose pas la réalisation effective de la récupération de chaleur**. Il est à noter que l'ADEME estime à 205 GWh/an le gisement actuel exploitable pour le chauffage de bâtiments récents, l'eau des piscines, le chauffage des serres ou de bâtiments agricoles. Cela représente 5,4 % du potentiel total de développement des énergies renouvelables à l'échelle métropolitaine, selon le PCAEM de la MGP.

Pourtant, il est constaté que les raisons économiques et techniques mises en avant par les promoteurs de centres de données, expliquant la rareté du réemploi de la chaleur fatale, sont rarement démontrées. Elles s'expliquent davantage par le fait que le projet n'a pas été conçu dans un écosystème pensé à cet effet. Les services de la préfecture Ile-de-France, comme la MRAE, constatent qu'aucun projet déposé à ce jour n'assure avec certitude la récupération de chaleur fatale, et signalent que les collectivités locales n'anticipent que très peu ce sujet dans leur planification.

Enfin, il est à noter que la récupération de la chaleur fatale peut faire face à plusieurs obstacles. D'une part, elle nécessite la présence et l'accord d'un gestionnaire local du réseau de chaleur ou de froid, ce qui est aujourd'hui trop peu anticipé. Et d'autre part, elle peut être confrontée à des difficultés techniques importantes (différences majeures de température, longueur et difficulté du raccordement... etc.).

Les impacts sont de trois ordres : d'abord, le gaspillage énergétique dans la mesure où une récupération de cette chaleur pourrait être envisagée ; ensuite, l'aggravation du phénomène de îlot de chaleur pour les riverains habitant à proximité des établissements ; enfin, l'acceptabilité politique et sociale du projet reste difficile, ce gaspillage de ressources générant de l'incompréhension.

²²⁷ <https://eur-lex.europa.eu/eli/dir/2023/1791/oj/fra>

²²⁸ https://www.legifrance.gouv.fr/jorf/article_jo/JORFARTI000044327305

²²⁹ https://www.drieat.ile-de-france.developpement-durable.gouv.fr/IMG/pdf/fiche_repere_dc_fev2022_vdef.pdf

Les données et indicateurs sur la chaleur fatale

Les données sur la récupération de chaleur à l'échelle d'un bâtiment sont liées à :

- La consommation électrique d'un DC, données qu'il est difficile d'obtenir à cette échelle
- Les termes contractuels posés entre l'opérateur et le gestionnaire de réseau, contrat souvent également confidentiel.

Disposer d'une image précise de la récupération de chaleur fatale en France en termes quantitatifs aujourd'hui est donc difficile, cependant, comme souligné souvent par l'Institut Paris Région dans ses travaux, les projets sont très rares. Par ailleurs, il y a beaucoup d'annonces et peu de suivi pour savoir si les projets sont effectivement réalisés et dans quelles configurations.

Les cahiers techniques de l'ADEME sur le sujet, adressés d'une part aux collectivités, d'autre part aux centres de données, sont très complets.

Un indicateur est dédié sur la récupération de chaleur fatale :

L'ERF (*Energy Reuse Factor*) mesure la part d'énergie réutilisée à l'extérieur du centre de données, par exemple sous forme de chaleur. Les centres de données qui sont équipés pour récupérer cette énergie peuvent chauffer des bâtiments, des réseaux de chaleur etc.

$$\text{ERF} = \text{Energie réutilisée} / \text{énergie totale consommée par centre de données}$$

L'ERF varie de 0 à 1, et 1 signifie que toute l'énergie est réutilisée. Par exemple, si un centre de données consomme 10 mégawatts (MW) d'électricité et peut réutiliser 2 MW sous forme de chaleur pour chauffer les bâtiments voisins, son ERF est de 0,2.

1.5.2. Eau

Les consommations d'eau ont et auront un impact de plus en plus fort sur les tensions mondiales et locales sur l'eau. Les niveaux de consommations sont mal connus car peu de données sont transmises par les opérateurs ou centralisées par les opérateurs d'adduction d'eau. Selon les systèmes de refroidissement, les consommations diffèrent comme démontré plus haut.

Sont identifiés deux enjeux principaux :

- **Un enjeu croissant de résilience du système d'approvisionnement en eau**, qui nécessite d'éviter le développement des centres de données dans des secteurs de pression sur la ressource en eau (zones de stress hydrique)
- **Un enjeu de limitation de la consommation d'eau par les data centers**, qui nécessite de trouver un bon équilibre entre efficacité dans l'utilisation de l'énergie (indicateur PUE, power usage effectiveness) et WUE (indicateurs WUE, water usage effectiveness) et qui passent par de nouvelles solutions comme l'utilisation des eaux grises ou des eaux de pluie par exemple.

Plusieurs articles ont récemment été publiés comme par exemple *China ICT running dry? The rise of AI & climate risks amplify existing water risks faced by thirsty data centres*, écrit par Debra Tan et Dr CT Low,

démontrant que le sujet commence à être davantage exploré, ou encore, l'article Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models,²³⁰ cité plus haut.

Les études sur les dérèglements climatiques en France prévoient davantage de vagues de chaleur et de sécheresses. La question de l'utilisation de l'eau est donc stratégique aux échelles locales comme à l'échelle nationale. Mais de quoi parle-t-on exactement lorsqu'on parle de consommation d'eau ?

Définitions

Eau prélevée (water withdrawal)

L'eau prélevée, en anglais "water withdrawal", est celle qui entre dans le centre de données. Cet indicateur de flux mesure la quantité totale d'eau douce en entrée, dans un process, ici le centre de données. Elle peut venir du système d'adduction d'eau locale, en général de l'eau potable (en France, seule la ville de Paris profite de deux réseaux : un réseau d'eau potable et un d'eau non-potable qui sert notamment à nettoyer les espaces publics). La donnée pourrait alors être disponible à travers des factures d'eau de l'opérateur, mais aussi d'outils de mesure au sein du centre de données.

L'eau peut aussi provenir de forages directs depuis la parcelle de l'opérateur, dans les nappes phréatiques accessibles ou de rivières comme à Marseille pour MRS 3, deux centres de données de Digital Realty. En 2018, l'entreprise a en effet obtenu de la Préfecture l'autorisation de détourner pour une durée de sept ans les eaux de la cunette de la Galerie à la Mer²³¹, afin de les récupérer pour refroidir son centre de données.

Eau consommée (water consumption)

$$\text{Eau consommée} = \text{eau prélevée} - \text{eau rejetée dans le milieu}$$

L'eau consommée est la quantité d'eau qui est consommée à chaque étape du cycle de vie du produit (ou service) étudié. Une quantité d'eau consommée est une donnée d'inventaire. **L'eau consommée correspond à la part prélevée qui n'est pas restituée au milieu aquatique au même moment et au même endroit que le prélèvement.** Il n'y a que 3 manières de consommer de l'eau :

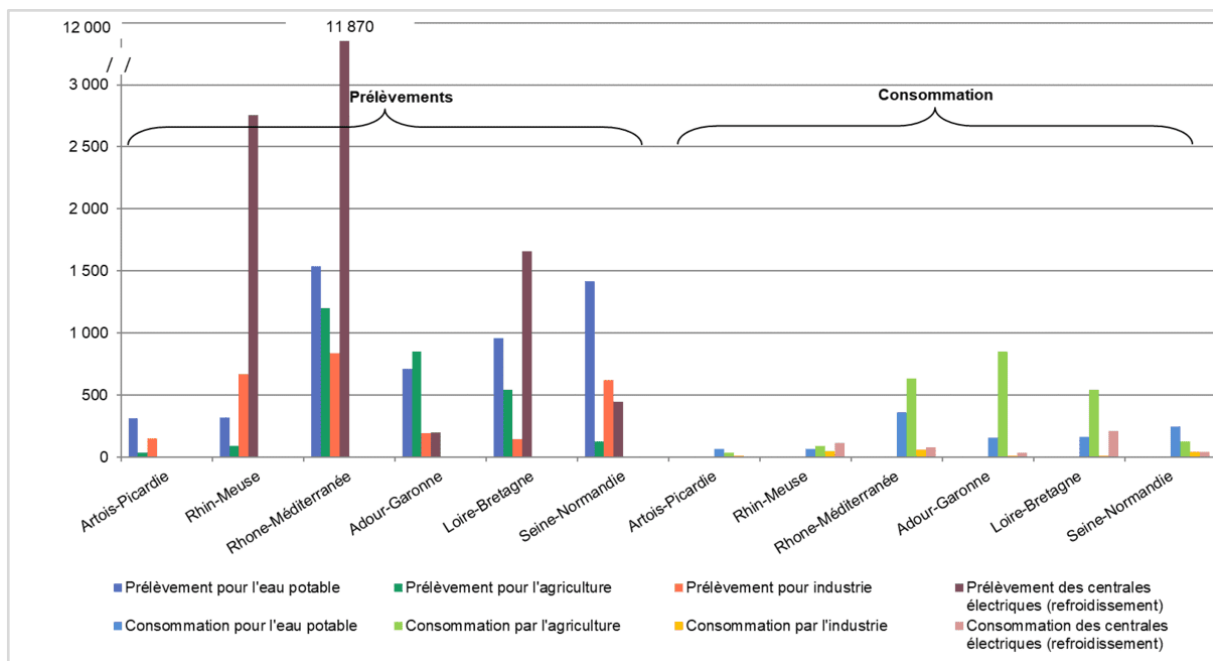
- Évaporation/évapotranspiration : les systèmes de refroidissement adiabatique génèrent par exemple une consommation d'eau plus importante que d'autres systèmes de ce fait.
- Incorporation dans un produit : par exemple, eau minérale, pastèque.
- Transfert dans un autre bassin versant : par exemple : canal, canalisation.

Ces consommations (évaporation, incorporation, transfert) peuvent avoir des effets sur l'environnement, la principale étant une disponibilité réduite de la ressource à une échelle locale plus ou moins large.

Graphique 3 : prélèvements et consommation d'eau douce en France (moyenne 2010-2018) En millions de m³

²³⁰ <https://arxiv.org/pdf/2304.03271>

²³¹ <https://www.laquadrature.net/2024/11/20/accaparement-du-territoire-par-les-infrastructures-du-numerique/> ; <https://www.agam.org/wp-content/uploads/2020/12/Datacenter-Interxion-MRS3.pdf>



source: <https://www.statistiques.developpement-durable.gouv.fr/leau-en-france-ressource-et-utilisation-synthese-des-connaissances-en-2021>

L'État souligne ici que "L'eau consommée correspond à la partie de l'eau prélevée non restituée aux milieux aquatiques. Cette part est très variable selon les utilisations. En moyenne, entre 2010 et 2018, le volume annuel d'eau consommée est estimé à **4,1 milliards de m³ en France métropolitaine** (soit environ 15 % des 27,8 milliards de m³ d'eau prélevée, hors alimentation des canaux), ce qui représente 64 m³/habitant. L'agriculture est la première activité consommatrice d'eau avec 57 % du total, devant l'eau potable (26 %), le refroidissement des centrales électriques (12 %), et les usages industriels (5 %). Cette répartition est variable selon les bassins : l'eau consommée est attribuée majoritairement à l'agriculture dans les bassins Adour-Garonne (80 % du total d'eau consommée), Loire-Bretagne (59 %) et Rhône-Méditerranée (56%), à l'eau potable en Artois-Picardie (59 %) et en Seine-Normandie (53 %), et à la production d'électricité en Rhin-Meuse (36 %)."

Il faut noter une confusion fréquente entre eaux prélevées et eaux consommées dans de nombreux rapports, articles, ou un flou qui ne facilite pas la comparaison des données.

Eaux rejetées/ restituées/ cycle de l'eau

Dans le cas des centres de données, les eaux rejetées sont celles qui sortent du centre de données, et réintègrent soit le circuit des eaux usées, sont traitées en usine d'assainissement puis retournent dans le système hydraulique ; soit sont traitées sur place et rejetées directement dans le système hydraulique. Par ailleurs, **le cycle de l'eau peut aussi être perturbé si les eaux rejetées ne sont pas à la même température que l'eau des rivières ou systèmes naturels concernés** (river cooling par exemple). Enfin, il peut aussi l'être si des polluants sont rejetés dans les milieux naturels ; ce que mesure l'empreinte eau multicritères, mais pas le WUE (voir ci-dessous).

Recyclage des eaux (treated reclaimed water)

L'eau peut être réutilisée de nombreuses fois dans un système de refroidissement en boucles, mais il y a cependant une limite à ce recyclage. En effet, des dépôts calcaires se forment peu à peu, le calcium, magnésium et la silice s'accumulent dans l'eau, de manière de plus en plus concentrée au fil des cycles de refroidissement ; l'eau atteint aussi des niveaux excessifs de conductivité. Il faut alors remplacer l'eau. Chez Equinix par exemple, ce remplacement a lieu tous les six mois.²³² D'autres techniques sont possibles pour limiter la concentration d'impuretés dans l'eau. Ainsi, Digital Realty utilise à Singapour "un système de détartrage électrolytique qui empêche les impuretés de s'accumuler et permet à l'eau de circuler dans le système de refroidissement trois fois plus qu'avant."²³³

Des procédés plus classiques sont sinon mis en place pour garder au maximum l'eau dans des circuits fermés, via deux techniques différentes, dites de "déseimbouage"²³⁴ :

- Chimique (produit)
- Ou mécanique (filtre)

Ce que l'on peut davantage désigner par *eaux recyclées* recouvre l'utilisation d'eaux grises par les centres de données, eaux qui ne sont pas potables mais suffisamment traitées pour être utilisées pour leur refroidissement. Les opérateurs internalisent aussi parfois des unités de traitement de l'eau pour réutiliser l'eau sur site. Amazon Web Services affirme ainsi traiter des eaux grises pour refroidir 16 de ses centres de données en Virginie du Nord et à Santa Clara.²³⁵ Microsoft a co-investi en 2021 dans une unité de traitement de l'eau pour les activités industrielles locales et son centre de données de Quincy, Washington, afin de ne pas utiliser de l'eau potable et de garder les eaux industrielles dans une boucle la plus fermée possible.²³⁶

Données et indicateurs sur l'eau

Périmètre des consommations d'eau : "water usage scopes"

Comme pour le carbone, il y a trois "scopes" pour parler et mesurer les consommations d'eau.

- Scope 1 : eaux consommées sur le site pour le refroidissement
- Scope 2 : eaux consommées hors du site pour la production électrique
- Scope 3 : eaux consommées pour toute la supply chain (chip manufacturing, purification des métaux et des terres rares pour le hardware, extraction des terres et métaux rares...)

²³² Entretien, PA10, octobre 2024

²³³ <https://www.datacenterdynamics.com/en/analysis/keeping-your-cooling-water-clean/>

²³⁴ Entretien avec Wilo, octobre 2024

²³⁵ <https://www.datacenterdynamics.com/en/news/aws-using-reclaimed-wastewater-for-data-center-cooling-at-20-locations/>

²³⁶ <https://www.epa.gov/waterreuse/water-reuse-case-study-quincy-washington>

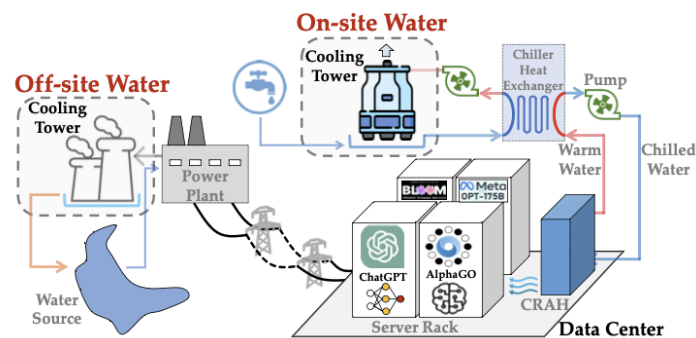


Figure 1: An example of data center's operational water usage: on-site scope-1 water for server cooling (via cooling towers in the example), and off-site scope-2 water usage for electricity generation. The icons for AI models are only for illustration purposes.

Source: Making AI less "thirsty"

WUE, Water Usage Effectiveness : pour les centres de données

L'indicateur le plus usité aujourd'hui pour comparer et représenter la consommation d'eau par bâtiment d'un centre de données ou à l'échelle du patrimoine d'un opérateur est le WUE.

Sur le modèle du PUE, le WUE est calculé en divisant le nombre de litres d'eau utilisés pour l'humidification et le refroidissement par la consommation électrique annuelle mesurée en kWh pour faire fonctionner le matériel IT du centre de données.

WUE on site

$$WUE = W_U / E_{IT}$$

W_U = consommation annuelle de l'eau du centre de données en m3

E_{IT} = consommation annuelle des équipements IT en MWh

W_U (consommation annuelle de l'eau du centre de données en m3) = $W_{in} - W_{out}$

W_{in} = débit entrant annuel

W_{out} = débit sortant annuel

donc

WUE = eau consommée en L par an / électricité consommée en kWh/an pour l'IT
aussi appelé WUE on site

Alors que la valeur optimale du PUE est 1, la valeur optimale du WUE est de 0 L/ kWh, ce qui signifie qu'aucune eau n'est utilisée pour les opérations de DC.

WUE source

Une mesure plus exhaustive du WUE peut inclure l'eau utilisée pour produire l'électricité consommée par le DC.

$$\text{WUE} = (\text{eau consommée en L par an} + \text{eau consommée par an pour la production électrique}) / \text{électricité consommée en kWh pour IT/an}$$

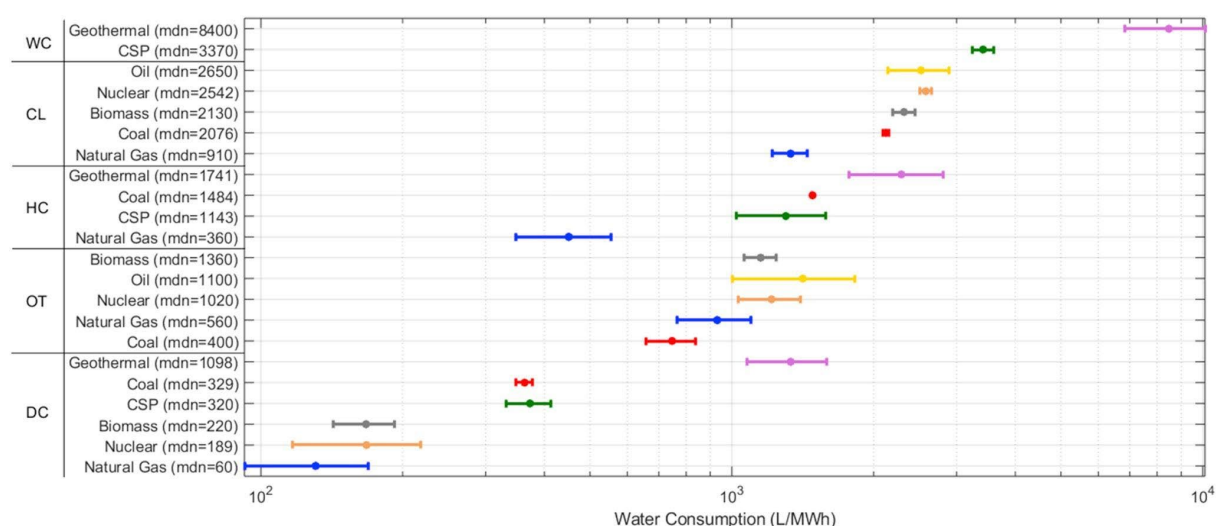
aussi appelé WUE source

L'eau consommée pour produire l'électricité y est calculée grâce à l'intensité hydrique du pays.

$$\text{intensité eau} = \text{eau consommée} / \text{production électrique du parc (l/kWh)}$$

L'intensité eau de la production électrique d'un pays dépend des énergies utilisées : charbon, nucléaire, gaz ou énergie hydraulique.

Le graphique ci-dessous l'illustre²³⁷ :



Explication "Consommation d'eau bleue pour les opérations distinguées par type d'énergie et type de refroidissement. Les points représentent la consommation moyenne d'eau, tandis que les segments de ligne représentent l'erreur standard de la moyenne. L'annotation mdn indique la valeur médiane. L'hydroélectricité, l'énergie éolienne et l'énergie photovoltaïque n'ont pas besoin de refroidissement et ne sont pas incluses. Les codes à deux lettres sont les suivants : WC refroidissement humide, CL refroidissement en boucle fermée, HC refroidissement hybride (combinant refroidissement humide et sec), OT refroidissement à passage unique, et DC refroidissement sec. Les couleurs correspondent au type de combustible pour l'estimation."

Pour identifier cette intensité eau, plusieurs ressources existent, en particulier un rapport très complet publié par le World Resources Institute (WRI)²³⁸. En France, une étude ACV d'EDF sur ces centrales nucléaires montre par exemple qu'elles ont un WUE de 1,3L/kWh²³⁹.

En France, les centres de données n'utilisent pas le WUE source jusque-là, mais le WUE on site.

²³⁷ <https://www.sciencedirect.com/science/article/pii/S1364032119305994?via%3Dihub>

²³⁸ https://files.wri.org/d8/s3fs-public/guidance-calculating-water-use-embedded-purchased-electricity_0.pdf

²³⁹ https://www.edf.fr/sites/groupe/files/2022-06/edfgroup_acv-4_etude_20220616.pdf

Dans la norme cenelec 50600-4-9 qui définit la mesure du WUE, on trouve trois catégories.

Catégorie	Catégorie 1 basique	Catégorie 2 intermédiaire	Catégorie 3 avancée
Débit entrant	Entrée du DC	Entrée du DC	Entrée du DC - eau de pluie + eau nécessaire à la production d'énergie
Débit sortant	Aucune réutilisation (toute l'eau qui rentre sort en tant qu'eau usée)	Réutilisation non industrielle	Réutilisations industrielle et non industrielle
Déclaration supplémentaire	Aucune	Aucune	Niveau de stress hydrique à l'échelle régionale et consommation des sols

Avantages du WUE :

- on site/ source : indicateurs comparables car bien appropriés par les acteurs
- source : plus complet que on site

Inconvénients :

- on site : lacunaire quand il ne s'inscrit pas sur un territoire et ne prend pas en compte le stress hydrique
- Il est à mettre en relation avec le PUE pour donner une image plus juste des arbitrages à faire selon les contextes entre refroidissement eau ou air.

Empreinte eau ou water footprint (WF)

Une mesure plus holistique des consommations d'eau et de leurs impacts s'illustre dans l'empreinte eau ou *waterfootprint* (WF). La WF peut être mono critère ou multi critères (dans le cadre d'ACV). La WF est par essence un indicateur très localisé car en lien avec des situations de stress hydrique, qu'elle soit monocritère ou multicritère.

WF monocritère (aspect quantitatif) :

WF= inventaire des consommations d'eau en m3 X un indicateur de stress hydrique

WF multicritère(aspects quantitatif et qualitatif) :

WF= inventaire des consommations d'eau en m3 X un indicateur de stress hydrique

et inventaire des pollutions d'eau X plusieurs facteurs de caractérisation d'ACV selon les pollutions.

Si la WF multicritères était calculée, il faudrait le faire sur des terrains précis en France : métropole de Marseille ; Nord de l'Essonne... Cela n'a pas de sens de la calculer à l'échelle française. Il y a aussi une notion de saisonnalité.

Avantages :

- Indicateurs relatifs au stress hydrique local, situé
- Indicateur qui prend en compte plusieurs impacts dans sa version multicritères donc plus juste sur les effets globaux d'une activité

Inconvénients :

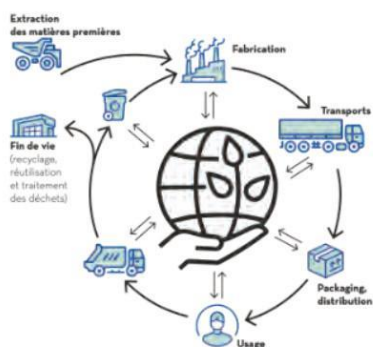
- trouver la bonne échelle pour l'utiliser
- nécessite des données fiables et saisonnières sur le stress hydrique, à la bonne maille (bassin versant ou sous-bassin versant)

Empreinte eau : ce qu'il faut retenir !

1.

Perspective cycle de vie

Une empreinte eau comptabilise tous les impacts relatifs à la consommation et à la pollution de l'eau, dans une perspective cycle de vie (de l'extraction des matières premières, à leur transformation, transport et usage, jusqu'à la fin de vie du produit).



2.

Facile et utile

L'empreinte eau monocritère est, au même titre que le bilan carbone, un indicateur environnemental très efficace pour une communication simple. Elle dispose maintenant d'un indicateur consensuel, AWARE (recommandé par le PNUE, et l'Europe par le PEF) avec des données disponibles à différentes échelles géographiques et prenant en compte la saisonnalité. La mise en œuvre peut être réalisée avec un simple tableur (cf. Bilan Carbone) à condition d'être très vigilant pour réaliser l'inventaire des consommations d'eau (en particulier ne pas mélanger « eau prélevée » et « eau consommée » et bien équilibrer les bilans hydriques). Attention cependant aux dangers de transferts de pollutions liés aux méthodes monocritères (ex. : le meilleur pour préserver la ressource en eau pourrait être le pire en terme de pollutions).

3.

De plus en plus pertinent

Les approches purement volumétriques de l'empreinte eau se sont progressivement améliorées : prise en compte de l'eau consommée (différence entre eau prélevée et restituée) ; prise en compte du stress hydrique local (un m³ d'eau n'a pas la même importance dans une région aride ou tempérée !) ; évaluation des effets de privation d'eau sur les écosystèmes, la santé humaine ou les ressources pour les générations futures.



4.

Empreinte eau multicritère ou ACV ?

La norme ISO 14046 propose d'aller plus loin que l'empreinte eau quantitative (monocritère) en y associant des indicateurs de pollution de l'eau (écotoxicité, eutrophisation, toxicité humaine, etc.). Il s'agit alors d'une empreinte eau multicritère, (voire « exhaustive » si elle inclut aussi les impacts sur la santé humaine en lien avec les pollutions de l'eau). La réalisation de celle-ci ne peut plus se faire avec des outils simples du type tableur et nécessite des logiciels et bases de données d'ACV. Dans ce cas, sachant que la mise en œuvre de l'empreinte eau devient aussi lourde que celle de l'ACV, il est alors bien préférable de basculer vers une ACV multicritère complète afin d'être exhaustif et d'identifier tous les transferts de pollutions.

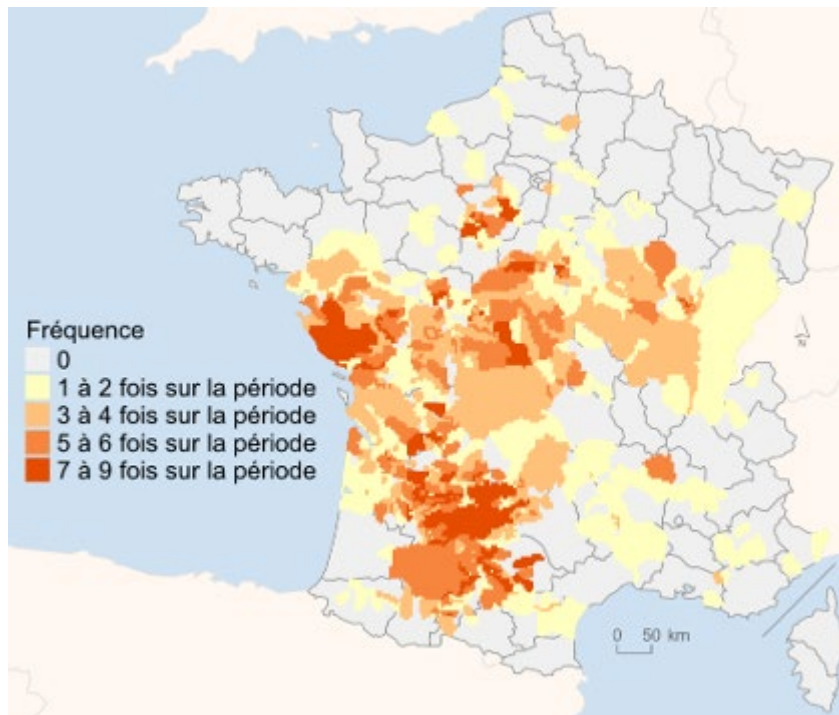


Source : Camille Maesele, Charlotte Pradinaud, Sandra Payen, Philippe Roux. L'empreinte eau - Mémento graphique. , pp. 48, 2021, 10.15454/rx5e-q558 . hal-03207737

Cette approche appliquée aux centres de données aux Etats-Unis a fait l'objet d'un article : *The environmental footprint of centres de données in the United States* écrit par Md Abu Bakar Siddik, Arman Shehabi et Landon Marston, et publié en mai 2021²⁴⁰. Une des conclusions est que : "Notre approche ascendante révèle qu'un cinquième de l'empreinte hydrique directe des serveurs des centres de données provient de bassins versants soumis à un stress hydrique modéré ou élevé, tandis que près de la moitié des serveurs sont entièrement ou partiellement alimentés par des centrales électriques situées dans des régions soumises à un stress hydrique."

²⁴⁰ https://iopscience.iop.org/article/10.1088/1748-9326/abfba1?_sp=b48260d8-0a7b-4784-9d4b-0e1ac60ee727

A titre d'illustration pour la France, la carte ci-dessous montre que certaines régions sont plus fragiles que d'autres.



Légende : fréquence des épisodes annuels de restriction de niveau « crise » des usages de l'eau superficielle d'une durée de plus d'un mois, sur la période 2012-2020 Champ : France métropolitaine et Corse. Sources : ministère de l'Écologie ; ministère de l'Agriculture, 2021. - © Traitements : SDES, 2021

Note de lecture : « 3 à 4 fois sur la période » signifie que des crises d'au moins un mois ont eu lieu 3 ou 4 années sur 9.²⁴¹

Focus sur la question des workloads et du WF²⁴²

L'article de Masanet, Shehabi, Lu, et Lei, intitulé *The WaterFootprint of DC workloads: a review of key determinants*, publié en 2022, aborde l'optimisation du calcul de l'empreinte eau grâce à une meilleure connaissance des workloads. Il se base sur des données serveurs de 2022.

L'article souligne que le WUE qu'il soit *on site* ou *source* ne prend pas en compte les potentiels d'efficacité à l'échelle des workloads, et d'une manière générale l'efficacité du matériel informatique. Il propose donc le calcul d'une empreinte eau définie par le nombre de litres d'eau par workloads (ssj_ops). L'article illustre:

- Les relations entre l'efficacité, les niveaux d'utilisation, les taux de rafraîchissement du matériel et la gestion des serveurs inactifs (comatose),
- Les relations entre WUE site et PUE, en prenant en compte des simulations climatiques et technologiques,

²⁴¹ <https://www.statistiques.developpement-durable.gouv.fr/leau-en-france-ressource-et-utilisation-synthese-des-connaissances-en-2021>

²⁴² <https://www.mdpi.com/2071-1050/7/8/11260>

- Les relations entre les facteurs de consommation d'eau pour le système électrique (WCFs), le WUE source et l'empreinte eau par workload.

Résultats :

- Pas de recette magique mais des stratégies de réduction de la consommation d'eau à penser au cas par cas selon les types de DC, notamment à cause du trade-off bien connu entre PUE et WUE, à penser en fonction du type de territoire, de son climat, de son mix énergétique...
- Mais globalement, augmenter l'efficacité des workloads des serveurs est de nature à réduire l'empreinte eau centres de données ; l'article insiste sur l'importance de prendre en compte cette mesure de l'efficacité des workloads, souvent sous-estimée.

Plusieurs indicateurs sont proposés :

- Direct server energy use/ workload
- Facility-level water use / workload
- Overall water footprint/workload
- Elle peut aussi être réduite grâce à l'utilisation d'ENR.

L'article considère les serveurs comme la composante principale de l'efficacité IT mais il recommande de se pencher aussi sur les GPU, le storage et le network loads.

Limite : l'article ne prend pas en compte tout l'usage d'eau sur le site en relation au territoire et à la ressource locale, ce qui est évidemment structurant en cas de stress hydrique.

Plus largement, on comprend ici l'intérêt de calculer une empreinte eau, et non juste un WUE : c'est une métrique plus écosystémique, contextualisée, qui permet de se poser des questions en termes de localisation idéale des centres de données, non pas seulement d'optimisation (qui aura toujours des limites) sur un bâtiment isolé.

Mesure des consommations d'eau et collecte de données

Pour mesurer les consommations d'eau des DC, trois démarches sont possibles :

- Partir de base de données générales et trier par secteur économique par exemple,
- Partir des consommations constatées centres de données site par site (base bottom-up),
- Partir d'enquêtes dédiées (rares et/ou pas exhaustives).

Aujourd'hui en France, aucune base de données générales sur l'eau comme la BNPE (voir ci-dessous) ne fait le détail sur les centres de données, qui sont donc difficiles à identifier.

Les données sur les consommations d'eau centres de données à l'échelle des sites dépend de la bonne volonté des opérateurs, sachant que les données transmises à l'ARCEP par les 19 plus grands opérateurs dans le cadre de l'enquête annuelle sont couvertes par le SDA, secret des affaires, et ne permet donc aucune exploitation même avec un NDA, Non-Disclosure Agreement.

EED

L'EED présente en revanche un potentiel intéressant de collecte et de calcul de données sur les consommations d'eau, dont l'envoi des données est en attente.

L'EED va notamment collecter les données suivantes, qui seront utiles à la connaissance des consommations d'eau et calcul d'indicateurs :

- Cooling infrastructure redundancy level at room level / at rack level.
- Installed information technology power demand
- Data centre total floor area ("SDC" in m²).
- Data centre computer room floor area ("SCR" in m²).
- Total energy consumption ("EDC" in kWh)
- Total energy consumption of information technology equipment
- Total water input ("WIN" in m3) as defined by CEN/CENELEC EN 50600-4-9 standard WUE
- Cooling degree days ("CDD" in degree-days) for the location of the reporting DC during the last calendar year, by using the methodology used by Eurostat and the JRC or equ. and with a base temperature of 21 C°

Les données publiquement mises à disposition aux États membres seront, pour l'eau les suivantes :

PUBLICLY AVAILABLE INFORMATION IN THE EUROPEAN DATABASE ON DATA CENTRES	
Pursuant to Article 12 of Directive (EU) 2023/1791, the European database shall be publicly available on an aggregated level.	
Data shall be available at two levels of aggregation, namely, at Member State level and Union level.	
Size categories of data centres shall be based on the data centre's information technology installed power as follows:	
(a)	very small data centre: 100-500 kW;
(b)	small data centre: 500-1 000 kW;
(c)	medium size data centre: 1-2 MW;
(d)	large data centre: 2-10 MW;
(e)	very large data centre: > 10 MW.
The following information shall be publicly available:	
(a)	at Member State level:
(i)	number of reporting data centres;
(ii)	distribution of reporting data centres by size categories;
(iii)	total installed information technology power demand (PD _{IT}) of all reporting data centres;
(iv)	total energy consumption (E _{DC}) of all reporting data centres;
(v)	total water consumption (W _{DC}) of all reporting data centres;
(vi)	average PUE for all reporting data centres in the MS territory, average PUE per type of data centre, and average PUE per size category;
(vii)	average WUE for all reporting data centres in the MS territory, average WUE per type of data centre, and average WUE per size category;
(viii)	average ERF for all reporting data centres in the MS territory, average ERF per type of data centre, and average ERF per size category;
(ix)	average REF for all reporting data centres in the MS territory, average REF per type of data centre, and average REF per size category.
For points (vi)-(ix), the aggregation of the sustainability indicators shall be performed with a weighted metric aggregation, using the total energy consumption as the weighting factor.	
For points (vi)-(ix), presentation of aggregated data per type of data centre and per size category will be possible only if the respective category contains data from at least three data centres;	
(b)	at Union level:
(i)	number of reporting data centres;
(ii)	distribution of reporting data centres by size categories;
(iii)	total installed information technology power demand (PD _{IT}) of all reporting data centres;
(iv)	total energy consumption (E _{DC}) of all reporting data centres;
(v)	total water consumption (W _{DC}) of all reporting data centres;
(vi)	average PUE for all reporting data centres in the Union territory, average PUE per type of data centre, average PUE per size category;
(vii)	average WUE for all reporting data centres in the Union territory, average WUE per type of data centre, average WUE per size category;
(viii)	average ERF for all reporting data centres in the Union territory, average ERF per type of data centre, average ERF per size category;
(ix)	average REF for all reporting data centres in the Union territory, average REF per type of data centre, average REF per size category.
For points (vi)-(ix), the aggregation of the sustainability indicators shall be performed with a weighted metric aggregation, using the total energy consumption as the weighting factor.	
ELI: http://data.europa.eu/eli/leg_del/2024/1364/oj	
ISSN 1977-0677 (electronic edition)	

La BNPE²⁴³

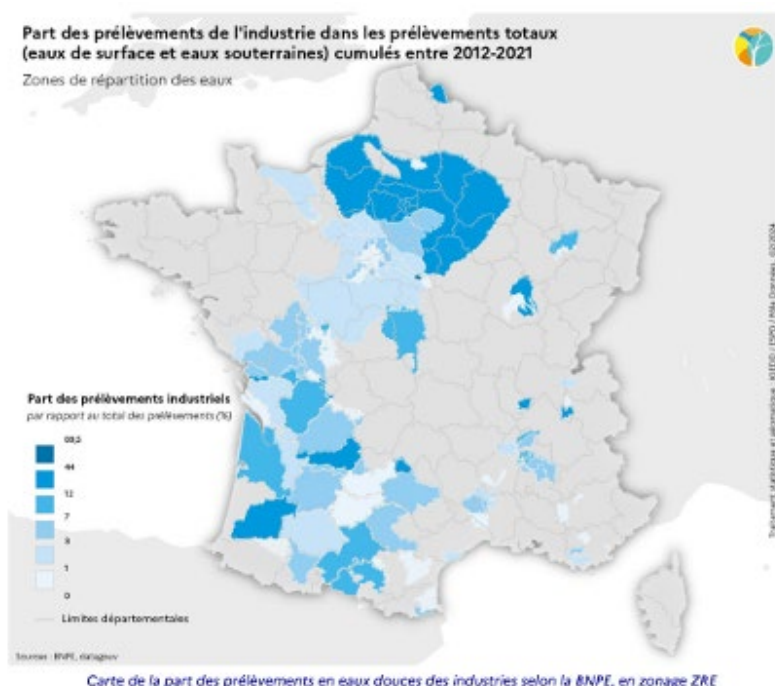
En France, la Banque Nationale des Prélèvements quantitatifs en eau (BNPE) est l'outil national dédié aux prélèvements sur la ressource en eau, pour la France métropolitaine et les départements d'outre-mer.

La BNPE constitue aujourd'hui la source nationale de données sur les prélèvements quantitatifs sur la ressource en eau. Les données présentes actuellement dans la banque sont les volumes prélevés, mesurés ou estimés puis déclarés par les usagers soumis à la redevance pour prélèvement auprès des agences et offices de l'eau. Elle comprend donc les volumes prélevés déclarés supérieurs à 10 000 m3/an (ou 7 000m3/an en ZRE – zone de répartition des eaux). **Elle ne distingue pas les ICPE et ne permet pas d'attribuer les volumes par secteurs d'activité. Par ailleurs, elle ne contient pas les données relatives aux**

²⁴³ <https://bnpe.eaufrance.fr/presentation>

restitutions de l'eau prélevée ou aux exhaures d'eaux de mines qui sont exonérées de redevance. Les données de la BNPE sont contrôlées par les agences de l'eau dans le cadre du contrôle de la redevance.

Rapport sobriété hydrique des industries par le CGE²⁴⁴



Les documents produits pour les plans de sobriété hydrique par types d'industries peuvent aussi apporter des données intéressantes, comme la carte ci-dessus. Elle met en effet en évidence que la part des prélèvements industriels peut être importante dans une ZRE, et qu'il est important de traiter cette zone en priorité du point de vue de l'industrie, notamment en Ile-de-France et dans le Sud-Ouest.

France- collecte de données ARCEP²⁴⁵

L'ARCEP collecte des données auprès de 21 centres de données de colocation/ hébergement en France. Ces 21 opérateurs de centres de données interrogés possèdent près de 150 centres de données en exploitation en France en 2023. Elle a publié en 2025 les données sur l'année 2023.

Il en ressort que, selon ce rapport :

« La quasi-totalité du volume d'eau prélevé par les centres de données en 2023 est de l'eau potable. Le volume d'eau prélevé par les centres de données atteint 681 000 m³ en 2023. La croissance annuelle du volume d'eau prélevé reste très soutenue et similaire à 2022 : elle s'établit à + 19 % en 2023 contre + 17 % en 2022. Le volume d'eau prélevé directement par les centres de données (681 000 m³ en 2023) est faible au regard du volume d'eau consommé indirectement par les centres de données. En 2023, le volume total d'eau prélevé ou consommé par les centres de données (direct + indirect associé à la consommation

²⁴⁴ https://www.economie.gouv.fr/files/files/directions_services/cge/Sobriete_Hydrique_IGEDD_CGE.pdf

²⁴⁵ https://www.arcep.fr/fileadmin/user_upload/observatoire/enquete-pns/edition-2024/enquete-annuelle-pour-un-numerique-soutenable_edition2024.pdf

d'électricité) est estimé à près de 6 millions de m3, soit la consommation annuelle moyenne d'eau en France d'environ 100 000 personnes."

Étude Green IT 2019²⁴⁶

Démarche type ACV simplifiée screening.

Le modèle de quantification des impacts environnementaux repose sur environ 2 000 données primaires. Le modèle a été mis au point entre décembre 2018 et juillet 2019. Le recueil des données a été réalisé entre février et juillet 2019.

Le rapport estime la consommation globale du secteur numérique en 2019 à 7,8 millions de m3 d'eau douce, soit 0,2% des consommations mondiales. Les "centres informatiques" ou centres de données représentent 7% de la consommation d'eau douce du secteur numérique. L'étude prévoit alors qu'entre 2010 et 2025, les consommations d'eau du secteur numérique pourraient être multipliées par 2,4.

En France, les résultats préliminaires de l'étude indiquent que le numérique au global représente, à l'échelle de la nation, l'équivalent de :

- Tension sur l'eau douce : 559 millions de m3 d'eau douce ;
- Eau : 2,2 % de la consommation de la France

La mise à jour de cette étude sortie en février 2025 n'intègre cependant pas d'indicateur eau, cet indicateur souffrant d'une mauvaise affectation des flux.

Résultats des études, enquêtes et collecte de données

France

La collecte de données Arcep aboutit à un WUE site global pour les 19 opérateurs de colocation de 0,4 L/kWh en 2022.

USA

Selon le [United States Data Center Energy Usage Report \(2016\)](#), le WUE moyen aux Etats-Unis était de 1.8 L/kWh en 2016. L'article de David Mytton sur le sujet²⁴⁷ fournit le même chiffre.

Entreprises

Microsoft

Le WUE de Microsoft est de 0.55 L/kWh sur le continent américain, 1.65 L/kWh en Asie Pacifique, et 0.1 dans la zone EMEA. La répartition du parc, les systèmes électriques et les conditions climatiques ont un impact sur l'indicateur.

Au global²⁴⁸, Microsoft voit une augmentation de sa consommation d'eau, qui atteint presque 8 millions de m3 d'eau en 2023.

²⁴⁶ <https://www.greenit.fr/empreinte-environnementale-du-numerique-mondial/>

²⁴⁷ Mytton, D. Data centre water consumption. *npj Clean Water* 4, 11 (2021). <https://doi.org/10.1038/s41545-021-00101-w>

²⁴⁸ <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW1IMjE#page=23>

1.3 Water

Table 9 – Water and effluents (megaliters)¹

	FY20	FY21	FY22	FY23
Total water withdrawals²	7,936	8,068	10,706	12,951
Third-party water	7,831	8,011	10,665	12,926
Surface water	89	41	39	21
Ground water	16	16	2	4
Total water discharges^{2,3}	3,740	3,295	4,307	5,107
Third-party water	3,740	3,295	4,307	5,107
Total water consumption²	4,196	4,773	6,399	7,844

META

Dans le rapport "Soutenabilité" de Meta paru en 2024²⁴⁹, le tableau ci-dessous détaille les données de prélèvement, consommation et "restoration", ce dernier indicateur n'est pas détaillé dans la méthodologie des métriques environnementales. Nous supposons qu'il a trait aux travaux de restauration hydrologique menés par Meta, à la manière d'une forme de compensation environnementale. Google parle de "replenishment".

Les données de consommation, comme pour Google et Microsoft, confirment des consommations à la hausse, notamment depuis 2022.

Water data (megaliters)					
	2019	2020	2021	2022	2023
Water withdrawal	3,430	3,726	5,043	4,893	5,274
Water consumption	1,971	2,202	2,569	2,638	3,078
Water restoration	145	2,250	2,336	2,352	5,889

Water withdrawal, consumption and restoration definitions can be found in our [environmental metrics methodology](#).

Le WUE du parc de Meta est de 0,18 L/ kWh selon le même rapport.

Google

Selon l'article *Making AI less "thirsty"* cité plus haut : "Même en mettant de côté l'utilisation de l'eau dans les installations de colocation louées à des tiers, les centres de données propres à Google ont à eux seuls prélevé directement 25 milliards de litres et consommé près de 20 milliards de litres d'eau du scope 1 pour le refroidissement sur site en 2022, dont la majeure partie était de l'eau potable. Globalement, l'utilisation de l'eau dans les centres de données de Google (prélèvement et consommation) en 2022 a augmenté de 20 % par rapport à 2021, et l'utilisation totale d'eau de Microsoft a même connu une augmentation de 34 % au cours de la même période. Ces augmentations significatives sont probablement attribuées en partie à la demande croissante d'IA."

Récentes évolutions

²⁴⁹ <https://sustainability.atmeta.com/wp-content/uploads/2024/08/Meta-2024-Sustainability-Report.pdf>

Beaucoup d'articles ont attribué à l'IA une croissance importante des consommations d'eau, comme par exemple, *Making AI Less "Thirsty*, cité ici à plusieurs reprises, qui souligne que si les estimations proposées se concrétisent, les **prélèvements** d'eau combinés des scopes 1 et 2 de l'IA dans le monde pourraient atteindre de **4,2 à 6,6 milliards de mètres cubes en 2027**, soit plus que le prélèvement d'eau annuel total de 4 à 6 Danemark ou de la moitié du Royaume-Uni.

L'article envisage cependant les prélèvements, et non les consommations. On voit donc la difficulté à faire des projections sur les consommations.

En parallèle, d'autres acteurs alertent sur le dégonflement possible de la bulle. Le dernier rapport de l'Agence Internationale de l'Energie (voir plus haut) reste prudent sur la consommation électrique (liée à la croissance globale des usages IA). La question reste toujours celle de la localisation et des impacts territorialisés.

Documents et cadre réglementaires en lien avec l'eau

Échelle du site ou du bâtiment

Auto-régulation

On peut mentionner la proposition du Climate Neutral Data Center Pact concernant la préservation de l'eau d'un objectif à 0,4 L/kWh IT (engagement volontaire).

Réglementation ICPE²⁵⁰

Beaucoup de centres de données sont des ICPE et à ce titre doivent respecter les dispositions de la loi sur l'eau qui les concerne. Plusieurs articles, notamment les articles 15 et 31, concernent les centres de données : l'un sur la mesure, l'autre sur le débit maximum autorisé. Ils pourraient être exploités pour collecter des données.

Évaluation environnementale

Les avis des MRAE sur les projets de centres de données évoquent rarement la question de la consommation d'eau et de leurs rejets.

L'évaluation environnementale est réalisée à plusieurs titres, notamment celui d'une procédure de déclaration au titre de la réglementation relative à la loi sur l'eau (rubrique 2.1.5.0 : rejets d'eaux pluviales dans les eaux douces superficielles).

Échelle territoriale

SDAGE

²⁵⁰ <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000204891>

Élaboré au niveau de chaque grand bassin hydrographique (Seine-Normandie, Artois-Picardie, Loire-Bretagne, Rhône-Méditerranée...), le **schéma directeur d'aménagement et de gestion des eaux (SDAGE)** fixe, pour ce bassin :

- Les orientations fondamentales d'une gestion équilibrée de la ressource en eau,
- Ainsi que les objectifs de qualité et de quantité des eaux à atteindre.

Il définit aussi les actions à mettre en œuvre pour améliorer la qualité de l'eau.

Prévu pour 6 ans, le SDAGE est adopté par un comité de bassin et approuvé par le préfet coordonnateur de bassin.

SAGE

Le **schéma d'aménagement et de gestion de l'eau (SAGE)** est un outil de planification, institué par la loi sur l'eau de 1992, visant la gestion équilibrée et durable de la ressource en eau. Déclinaison du SDAGE à une échelle plus locale, il vise à concilier la satisfaction et le développement des différents usages (eau potable, industrie, agriculture, ...) et la protection des milieux aquatiques, en tenant compte des spécificités d'un territoire. Délimité selon des critères naturels, il concerne un bassin versant hydrographique ou une nappe. Il repose sur une démarche volontaire de concertation avec les acteurs locaux.

Il est un instrument essentiel de la mise en œuvre de la directive cadre sur l'eau (DCE). A ce titre pour respecter les orientations fondamentales et les objectifs fixés par la DCE, 68 SAGE ont été identifiés comme nécessaires par les SDAGE approuvés en 2009 (période 2010-2015), 62 SAGE ont été identifiés comme nécessaires par les SDAGE approuvés en 2015 (période 2016-2021), 24 SAGE ont été identifiés comme nécessaires par les SDAGE approuvés en 2022 (période 2022-2027).

ZRE- Zones de répartition des eaux²⁵¹

En France, la ressource en eau est inégalement répartie. **Pour certains territoires, l'eau disponible est inférieure aux besoins de la population, qu'il s'agisse d'une période de sécheresse ou non.** Certains territoires peuvent disposer de peu de ressource en eau disponible naturellement, d'autres peuvent être particulièrement peuplés, et d'autres secteurs peuvent cumuler ces deux réalités. **Les Zones de Répartition des Eaux sont des espaces géographiques sur lesquels ce constat est fait.**

Propositions du CGE (2023)²⁵²

En 2023, le CGE propose des restrictions des prélèvements dans les zones les plus fragiles. Cette piste pourrait être approfondie dans le futur pour mieux planifier l'implantation des centres de données.

Le cas de la planification à Amsterdam et la question de l'eau

Les municipalités d'Amsterdam, Haarlemmermeer et Hollands Kroon ont décidé d'orienter l'implantation et le développement des centres de données sur leurs territoires pour la période 2022 – 2024. Elles ont encadré l'implantation des centres de données ayant besoin d'une assiette foncière supérieure à 2 000 m² et d'une puissance électrique supérieure à 5 MW. Le plan de zonage prévoit l'implantation de nouveaux centres de données ou l'extension de centres de données existants seulement

²⁵¹ <https://www.data.gouv.fr/fr/datasets/zones-de-repartition-des-eaux-zre-metropole/>

²⁵² https://www.economie.gouv.fr/files/files/directions_services/cge/Sobriete_Hydrigue_IGEDD_CGE.pdf

dans les zones d'exploitation identifiées pour la clusterisation des centres de données et s'ils répondent aux conditions d'implantation conformes à la politique locale et aux conditions associées.

Les centres de données doivent réduire leur consommation d'eau. Si la technique de refroidissement retenue pour le centre de données est par recours à l'eau, alors l'ordre de préférence du recours à l'eau doit être le suivant :

- La recherche prioritairement d'une économie maximale sur la consommation d'eau,
- Réaliser un système énergétique selon les principes de puits canadiens mettant en œuvre des gaines enterrées pour utiliser le froid pour le refroidissement
- Utiliser l'eau de pluie, les effluents traités, l'eau saumâtre, ...
- Utiliser de manière optimale les eaux de surface,
- Utiliser de manière optimale des eaux industrielles,
- Utiliser de manière optimale des eaux souterraines,

Si les cas précédents sont insuffisants, utiliser de manière optimale l'eau potable.

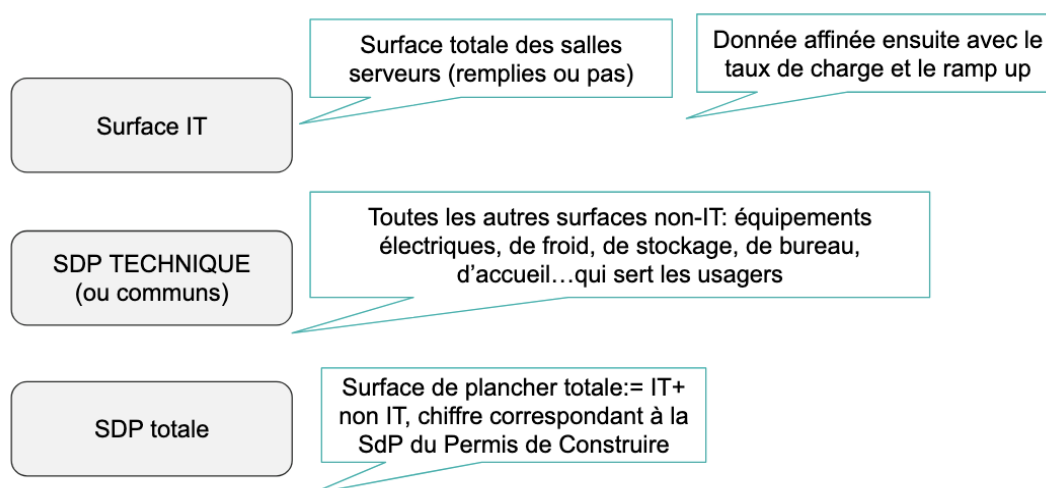
- Promouvoir l'auto-suffisance en conseillant aux opérateurs de centres de données de réaliser une réserve d'eau pour l'eau de refroidissement pour pallier les calamités et les périodes de sécheresse.
- Tenir un registre sur l'efficacité de l'utilisation de l'eau (WUE) et la consommation d'eau, l'utilisation et le type d'eau par mois à conserver pendant au moins 5 ans et à mettre à disposition de l'autorité compétente.

Foncier, spatialités

Définitions

Au sein du centre de données : surfaces en m2

Travailler sur l'empreinte spatiale des centres de données suppose de travailler avec des métriques spécifiques.



Connaître la surface de plancher (SDP) totale ou une des autres surfaces permet, par extrapolation, de calculer les autres surfaces.

Le ratio utilisé par les opérateurs est le suivant :

la SDP totale = 50% IT + 50% pour espaces communs (technique, bureaux, logistiques)²⁵³

D'un point de vue opérateur, les surfaces considérées sont de 3 types²⁵⁴ :

- **"louable" = NRSF = "Net rentable square feet"**

Les « pieds carrés louables nets » d'un bâtiment représentent les pieds carrés

C'est celle-ci que nous utiliserons aussi. Elle correspond à la surface IT louée et sa quote-part d'espaces communs (bureaux, techniques...)

- **"In progress"** : L'espace en développement actif comprend les projets de bâtiments de base et de centres de données en cours.

- **"Next"** : L'espace détenu pour le développement comprend l'espace détenu pour le développement futur de centres de données et exclut l'espace en cours de développement.

²⁵³ Cette règle statistique est établie à partir d'un échantillon de données primaires et d'interviews d'acteurs.

²⁵⁴ [Digital Realty](#)

Parcelle du centre de données

La connaissance de la parcelle sur laquelle est implanté le centre de données est aussi importante.

On mesure ainsi :

- La surface totale de la parcelle ou des parcelles concernées, en hectares ou en m²
- L'emprise au sol du bâtiment qui équivaut en général à la surface de rez-de-chaussée, c'est l'empreinte du bâtiment sur le sol.

On peut en déduire :

- Le coefficient d'emprise au sol (CES) qui donne une idée du taux d'occupation de la parcelle, de sa densité bâtie
- Elle permet ensuite d'identifier des potentiels d'artificialisation des sols si par exemple l'espace est entièrement bitumé pour un parking ou pour accueillir des équipements techniques.

La connaissance de l'emprise au sol permet de calculer une SDP extrapolée en multipliant par le nombre d'étages, que l'on peut connaître par l'observation terrain ou visualisation numérique, en prenant 6 m par niveaux pour les centres de données.

Source de la donnée :

- Base de données bottom up : localisation site, croisement avec cadastre pour identifier la parcelle, puis trouver ou calculer sa surface

D'autres bases sont disponibles :

- Base de données IPR pour l'Ile de France
- Base de données Pappers pour la France

Aujourd'hui, les modalités d'implantation des centres de données ne sont pas les mêmes selon les types :

	Taille	Parcelle	Géographie
Centres de données edge	petit (<500 m ²)	intégré dans un bâtiment/hors parcelle	ultra urbaine
Centres de données colocation retail	moyen (500-2000)	CES élevé	métropolitaine
Centres de données colocation wholesale	grand (2000-10 000)	CES moyen/élevé	métropolitaine
Centres de données hyperscale	très grand (> 10 000 et jusqu'à 100 000)	CES moyen/ bas possibilité de croissance sur site	péri-urbaine

Empreinte foncière centres de données

La notion d'empreinte foncière des centres de données n'est pas explorée à ce stade par les opérateurs ou dans les travaux académiques. Cette notion pourrait cependant être approfondie, en suivant les

réflexions sur les **hectares fantômes**, des historiens de l'environnement et des économistes. Le concept d'« hectares fantômes » (ghost acres ou ghost hectares) a été introduit par l'agronome suédois Georg Borgström (1965), a servi ensuite aux travaux sur la notion d'empreinte écologique ». C'est une grille intéressante pour parler du foncier occupé, détruit, utilisé ou pollué, en dehors de la parcelle du centre de données, que ce soit en France ou en dehors.

L'empreinte foncière globale d'un centre de données serait alors composée des :

- Hectares directs du centre de données = sa/ses parcelles pour faire fonctionner un site
- Hectares indirects du centre de données dans le pays d'implantation :
 - m2 de poste source électrique nécessaires à son fonctionnement
 - m2 ou ha d'usine de traitement de l'eau nécessaires à son fonctionnement
 - hectares de sites de production énergétique nécessaires à son fonctionnement (ferme solaire, éolien, nucléaire, hydraulique...)
- Hectares fantôme :
 - hectares nécessaires pour la construction du bâtiment et des matériels (mines, sites énergétiques pour ces phases, usines pétro/chimiques etc...)
 - hectares pour la phase d'usage centre de données (production énergétique importée)
 - hectares nécessaires pour l'entreposage des déchets (sites décharges)

Tout comme le secteur de la construction, centres de données peuvent chercher à baisser leurs émissions de gaz à effet de serre et impacts environnementaux sur la phase construction :

- Écomatériaux sourcés localement : bois, terre notamment
- À défaut, béton dit bas carbone
- Coordination des interventions sur les espaces publics pour éviter de creuser des tranchées à répétition sur les mêmes axes pour faire passer la fibre.

Impact des usages IA sur les surfaces et les bâtiments, sur les critères géographiques d'implantation

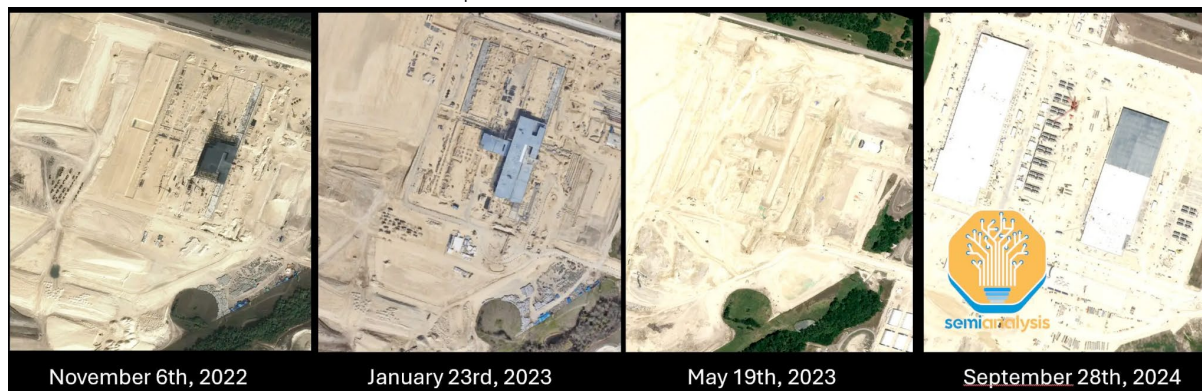
Le développement de l'IA reste grevé de multiples incertitudes. Voici donc plusieurs hypothèses sur la façon dont le déploiement de centres de données dédiés à l'IA, avec des intensités énergétiques très fortes, ainsi que l'introduction d'une part d'usages IA dans centres de données existants pourraient impacter les dimensions bâtimentaires, géographiques et foncières.

Selon nos différents entretiens et l'état de l'art sur la question, il semblerait que :

- Les bâtiments actuels des opérateurs de colocation en particulier ne soient capables que d'accueillir 10 à 15% d'usages IA dans les configurations actuelles
- Centres de données conçus actuellement et qui seront livrés dans 5 à 7 ans seront adaptés aux usages IA, pour le refroidissement en particulier ;
- La limite devient donc la puissance électrique mise à disposition, et non plus vraiment l'espace disponible : on pourrait ainsi imaginer qu'un site devienne partiellement vacant parce qu'il n'est plus possible d'accueillir davantage de racks de GPU, à cause d'une limite de puissance ; on pourrait ainsi imaginer qu'au lieu d'être vacant, l'espace est occupé par un autre usage : bureaux, activités économiques, escape game ou autre... ;

- Dans le cadre du move to cloud et/ou move to colo, un patrimoine de centres de données qui ne sont pas des nœuds télécoms importants (POP) pourraient être démolis pour construire des bâtiments adaptés à l'IA ; ici l'impact carbone et environnemental serait important.

Sur ce point, l'exemple de Meta²⁵⁵ qui démolit un centre de données en cours de construction pour rectifier le tir en construisant un bâtiment adapté à l'IA est assez parlant. Sur des images satellite de 2025, ce centre de données a de nouveau disparu.



RTE utilise aujourd'hui un ratio qui fait équivaloir 1 ha de parcelle de centres de données à 15 MW de puissance raccordée. Si l'on suit les indicateurs qui affirment que l'intensité énergétique des usages IA est 10 fois plus importante, on aurait alors :

$$1 \text{ ha} = 150 \text{ MW}$$

Par ailleurs, on peut se demander si les critères géographiques d'implantation pourraient changer pour ce que l'on appelle parfois les *IA factories*. La proximité avec d'autres centres de données ou des marchés de consommation n'étant pas aussi stratégique, Microsoft a par exemple un temps considéré le retrofit d'une ancienne usine PSA entre Nantes et Rennes, dans un territoire plutôt rural. Une hypothèse serait donc des implantations plus opportunistes possibles dans toute la France tant que les risques environnementaux sont limités et l'électricité disponible.

Il est aussi intéressant de noter l'importance de la planification dans le temps de l'utilisation de l'IA pour l'entraînement des modèles. Il est conseillé de faire tourner les modèles la nuit car il fait plus frais par exemple ("unfollow the sun"). Nous pourrions aussi en déduire que la localisation des *IA factories* sera encore plus guidée par des critères climatiques, avec par exemple peut-être un tropisme vers les pays scandinaves en Europe.

Impact du Zéro Artificialisation Nette (ZAN) sur les consommations futures de foncier par les centres de données en France

La loi Climat et Résilience de 2021, assortie de ses décrets d'application et nomenclatures, limite l'artificialisation des sols et pousse à la reconversion de friches. Cette démarche de Zéro Artificialisation Nette (ZAN) aura des conséquences sur les modalités d'implantation des centres de données dans le futur. Moins d'espaces et de grandes parcelles en extension urbaine sur des terres agricoles, naturelles ou

²⁵⁵ <https://semianalysis.com/2024/10/14/datacenter-anatomy-part-1-electrical/>

forestières devraient être disponibles. Or, les hyperscales cherchent aujourd’hui des sites très vastes. Par exemple, le site de Lisses (Essonne) pour Cloud HQ s’étend sur 13 hectares, celui de NTT (développé par LCP) à Evry et Coudray-Monceaux, sur 14 hectares.

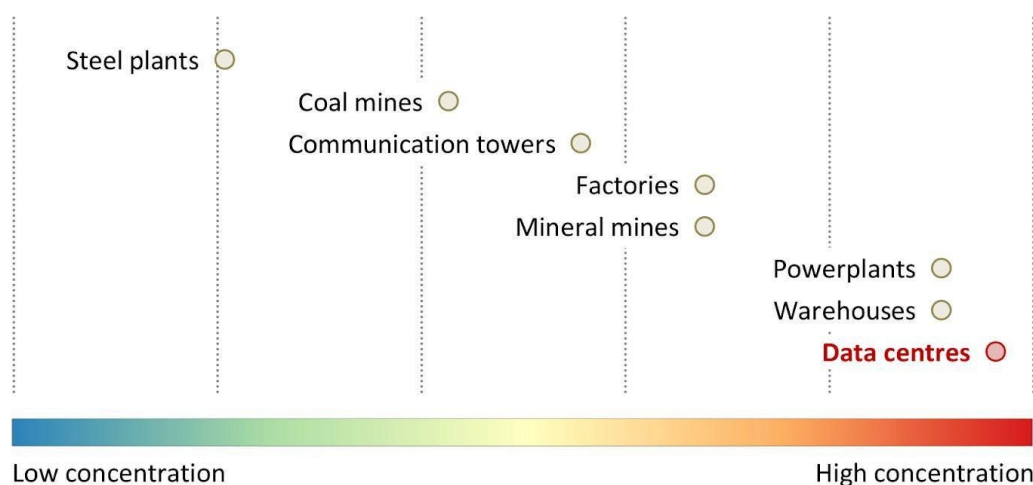
Certains documents de planification régionale comme le SDRIF en Ile-de-France et les SRADDET dans les autres régions, ont en partie émis ces objectifs, qui seront ensuite déclinés dans les plans locaux d’urbanisme.

La loi de “simplification de la vie économique” a été adoptée en juin 2025. Elle ouvre la possibilité, pour certains centres de données, de devenir des PINM, Projets d’Intérêt National Majeur, qui, à ce titre, bénéficierait d’un certain nombre d’exemptions environnementales, énergétiques et urbanistiques.²⁵⁶

Le sujet de la concentration spatiale

Le dernier rapport de l’AIE propose le graphique ci-dessous, qui met en avant la spécificité des centres de données en termes d’implantation spatiale : une très forte logique de concentration, exceptionnelle comparée à d’autres secteurs.

Figure 4.12 ► Spatial concentration of selected types of facilities, United States



IEA. CC BY 4.0.

Data centres have an exceptionally high spatial concentration, which has significant implications for local power grids, given their substantial power requirements

Cette concentration spatiale se retrouve dans les cartographies réalisées en Ile-de-France par l’Institut Paris Région, mais aussi dans les logiques de clusters constatées autour de Dublin, Francfort, Londres ou dans la data center Alley en Virginie du Nord.

Ce phénomène porte donc à des conséquences en termes de congestion électrique, de spécialisation économique, de paysage monofonctionnel, de faible densité d’emplois...

²⁵⁶ Dans son article 15, la loi prévoit que les centres de données pourront bénéficier du label programme d’intérêt national majeur (PINM). À ce titre, le permis de construire des centres de données sera délivré par l’État et non plus par les communes. Ils bénéficieront d’un “raccordement accéléré au réseau électrique” et d’une “mise en compatibilité accélérée, par l’État, des documents locaux d’urbanisme et des documents de planification régionale”. Ils pourront obtenir la reconnaissance anticipée de la ‘raison impérieuse d’intérêt public majeur’ (RIIPM), nécessaire à l’obtention de la dérogation espèces protégées.

Impacts territoriaux et oppositions locales

Moratoires, oppositions locales : les centres de données ne passent plus inaperçus

Le développement des centres de données fait face à une opposition croissante d'associations environnementales et de citoyens et citoyennes concernées.²⁵⁷ Des citoyens et associations se sont organisés contre la construction d'un nouveau complexe de centres de données de 200 MW à County Clare, en Irlande, secteur qui était jusqu'ici l'eldorado des centres de données en Europe. Plusieurs conflits d'usage ont été constatés dans d'autres grandes métropoles européennes ("Sous le feu numérique", 2023, Cécile Diguët, Fanny Lopez) mais aussi aux Etats-Unis, en particulier dans le comté de Loudoun, territoire du plus grand hub de centres de données probablement au monde : 450 hectares et 4 000 MW de puissance raccordée, où 1340 MW de centres de données sont en construction (voir ci-dessous)

Certains ont débouché sur des moratoires ou des équivalents. A Dublin les raccordements électriques sont ainsi suspendus. Au cours des dernières années, la croissance galopante des centres de données dans la métropole d'Amsterdam a créé des situations de tensions inédites, qui s'expriment notamment spatialement (peu de fonciers disponibles, prix élevés) et énergétiquement (infrastructures électriques saturées dans certaines zones). C'est pour ces raisons que les municipalités d'Amsterdam et d'Haarlemmermeer ont décidé en 2019 d'instaurer un moratoire sur les nouvelles constructions de centres de données pour la durée d'un an. Depuis, Amsterdam a adopté un document cadre sur les centres de données, et Haarlemmermeer un document cadre et un plan de zonage (voir plus haut). Singapour a également mis en place un moratoire entre 2019 et 2022. On peut aussi noter que la Commission Nationale du Débat Public a recommandé au gouvernement d'intégrer les projets de centres de données à de plus larges débats démocratiques pour favoriser la participation du public à ces projets structurants et impactants.

Pour autant, des initiatives réussies et plus exemplaires se sont aussi développées. Le centre de données de la Ville de Paris et de l'AP-HP, au-delà de son intégration architecturale au quartier Chapelle Internationale, permet de chauffer les logements voisins grâce à une boucle locale.²⁵⁸ Le tout dernier centre de données d'Infomaniak a été construit sous des immeubles d'habitation à Genève, et chauffe une partie du nouveau quartier de la Bistoquette, au total 6000 ménages.²⁵⁹

Oppositions locales multi dimensionnelles

Depuis quelques années, des collectifs d'habitants, des collectivités locales, des activistes environnementaux se penchent sérieusement sur les impacts négatifs des centres de données. Alors que ces luttes sont longtemps restées ponctuelles, on sent maintenant une consolidation des contestations.

En France, la lutte historique est celle des femmes de la rue du Rateau contre l'extension du centre de données de Digital Realty sur une friche non bâtie, que les habitantes auraient souhaité voir devenir un parc de proximité. La question des concurrences d'usage en zone métropolitaine dense était déjà présente.

²⁵⁷ <https://www.datacenterdynamics.com/en/news/judicial-review-could-block-200mw-data-center-campus-in-county-clare-ireland/>

²⁵⁸ <https://www.paris.fr/pages/la-ville-innove-en-se-dotant-de-son-propre-data-center-6685>

²⁵⁹ <https://bistoquette.ch/article-tdg-16-avril-2024/>

Plus récemment, c'est à Wissous, au Nord de l'Essonne, que certains élus se sont mobilisés contre un projet de Cyrus One pour Amazon Web Services, dont l'implantation par tranches a semblé être une ruse malhonnête pour les habitants, minimisant ainsi la perception des nuisances potentielles.

L'association France Nature Environnement (FNE) a, de son côté, organisé plusieurs événements ces dernières années pour partager la connaissance sur les impacts environnementaux des DC.

À Marseille, certains élus critiquent à la fois les techniques parfois rugueuses d'implantation des DC, et les impacts négatifs sur le territoire, notamment sur la qualité de vie, la densité d'emploi, les consommations électriques mais surtout, les concurrences d'usages. Avec l'association Le Nuage était sous nos pieds²⁶⁰, soutenue par La Quadrature du Net et le Mouton Numérique, un festival techno-critique a été organisé pendant trois jours en novembre 2024²⁶¹.

Comme évoqué plus haut, les oppositions sont aussi de plus en plus fortes en Virginie du Nord, où les impacts des centres de données se font vraiment extrêmement tangibles pour les habitants. L'agence environnementale du Piedmont²⁶² est ainsi très mobilisée pour dénoncer la croissance incontrôlée des centres de données dans plusieurs comtés, et les impacts sur la biodiversité, les paysages, les sols, la consommation électrique, la qualité de l'air... Récemment, la construction d'un centre de données de Microsoft a même abouti à déplacer un petit cimetière de personnes afro-descendantes²⁶³.

Ces mouvements pourraient ralentir l'implantation de centres de données dans certains territoires, si le soutien des élus comme celui des habitants venaient à manquer.

Oppositions locales sur le sujet de l'eau

La ressource en eau attise encore plus d'oppositions tant cette ressource est évidemment vitale. Les mouvements se cristallisent ainsi en Espagne et en Amérique Latine, mais aussi dans certains états du sud des Etats-Unis :

- Le collectif "Tu nube seva mi rio"²⁶⁴ est un collectif espagnol techno-critique qui met en avant les impacts négatifs des centres de données sur la ressource en eau, et leur volonté de s'installer y compris dans des zones de stress hydrique, en Espagne et en Amérique Latine.
- En Uruguay, en pleine sécheresse, alors que l'eau des habitants était imbuvable car en partie salée, un projet de centre de données de Google a suscité une polémique très importante²⁶⁵.
- Au Chili, Google a planifié un projet de centre de données à Cerillos malgré le stress hydrique, projet désormais abandonné²⁶⁶.
- Meta s'est installé à Mesa Arizona, malgré le stress hydrique. En 2021, une des adjointes de la ville s'y opposait, le projet a été quand même validé : *"As we form Mesa's climate action plan and embark upon the first phase of the seven-state drought contingency plan, making cutbacks to agriculture, I cannot in good conscience approve this mega data center using 1.7 million gallons per*

²⁶⁰ <https://lenuageetaitsousnospieds.org/>

²⁶¹ <https://www.laquadrature.net/2024/11/20/accaparement-du-territoire-par-les-infrastructures-du-numerique/>

²⁶² <https://www.pecva.org/>

²⁶³ <https://mimemorial.com/industry/developers-found-graves-in-virginias-woods-authorities-then-helped-erase-the-historic-black-cemetery>

²⁶⁴ <https://tunubeseccamirio.com/>

²⁶⁵ https://www.huffingtonpost.fr/environnement/article/en-uruguay-la-moitie-de-la-population-n-a-plus-d-eau-potable-et-reproche-a-google-de-piller-les-dernieres-gouttes_220702.html

²⁶⁶ <https://www.emol.com/noticias/Economia/2024/09/17/1142980/google-proyecto-data-center-cerrillos.html>

*day at total build-out, up to 3 million square feet on 396 acres," Mesa Vice Mayor Jenn Duff said, in a vote over the Facebook project. She was the only one to vote against the facility."*²⁶⁷

Ces oppositions locales, qu'elles concernent les nuisances liées à l'implantation de nouveaux centres de données supplémentaires, ou plus spécifiquement les tensions auxquels ils contribuent sur la ressource eau, s'amplifient au fur et à mesure que les projets de construction de nouveaux centres de données s'accroissent. Dans ce contexte, on pourrait assister à une forte augmentation des oppositions dans les prochaines années si les projections de croissance de l'IA se confirmaient. Ces oppositions pourraient avoir pour conséquence un ralentissement relatif dans la mise en œuvre des nouveaux projets d'implantation, et/ou une montée en maturité dans la prise en compte de critères environnementaux pour ces projets. Nous envisageons d'intégrer cet aspect dans le scénario prospectif "Coopération territoriale".

²⁶⁷ <https://www.datacenterdynamics.com/en/analysis/data-center-water-usage-remains-hidden/>

2. Analyse qualitative des études prospectives sur l'évolution des consommations des centres de données

2.1. Études institutionnelles

2.1.1. "Transition(s) 2050. Choisir maintenant. Agir pour le climat" ²⁶⁸

Les modèles prospectifs associés concernant les centres de données ont été étudiés et ils serviront de référence par rapport aux nouveaux modèles proposés. Avant d'analyser les résultats il est important de souligner deux spécificités du périmètre de l'étude :

- **Seuls les centres de données de type "co-location" y sont considérés. Les centres de données hébergés dans les entreprises sont exclus.**
- **Les centres de données importés pour un usage français ne sont pas pris en compte.**

La modélisation des centres de données est effectuée dans le modèle VIVALDI qui n'est pas disponible dans le rapport.

Plusieurs scénarios sont étudiés :

Le scénario tendanciel se base sur une prévision de croissance forte du volume de données, sous l'influence de la 5G et de l'IoT, qui n'est pas compensé par les gains en efficacité énergétique. Il aboutit à une forte croissance de la consommation énergétique des centres de données.

Le scénario #1, intitulé "Génération Frugale", correspondait à une frugalité choisie mais aussi contrainte. Il se basait sur un arrêt de la croissance du volume de données à partir de 2025 et sur des gains en efficacité énergétique. Le développement des équipements ne nécessite pas d'évolution des centres de données.

Cette tendance se combinait à une meilleure efficacité énergétique pour aboutir à une baisse de la consommation des centres de données **qui passait de 2 TWh en 2015 à 0,8 TWh en 2050.**

Le scénario #2, intitulé "Coopérations Territoriales" se basait sur *"une évolution soutenable des modes de vie et une économie du partage."*

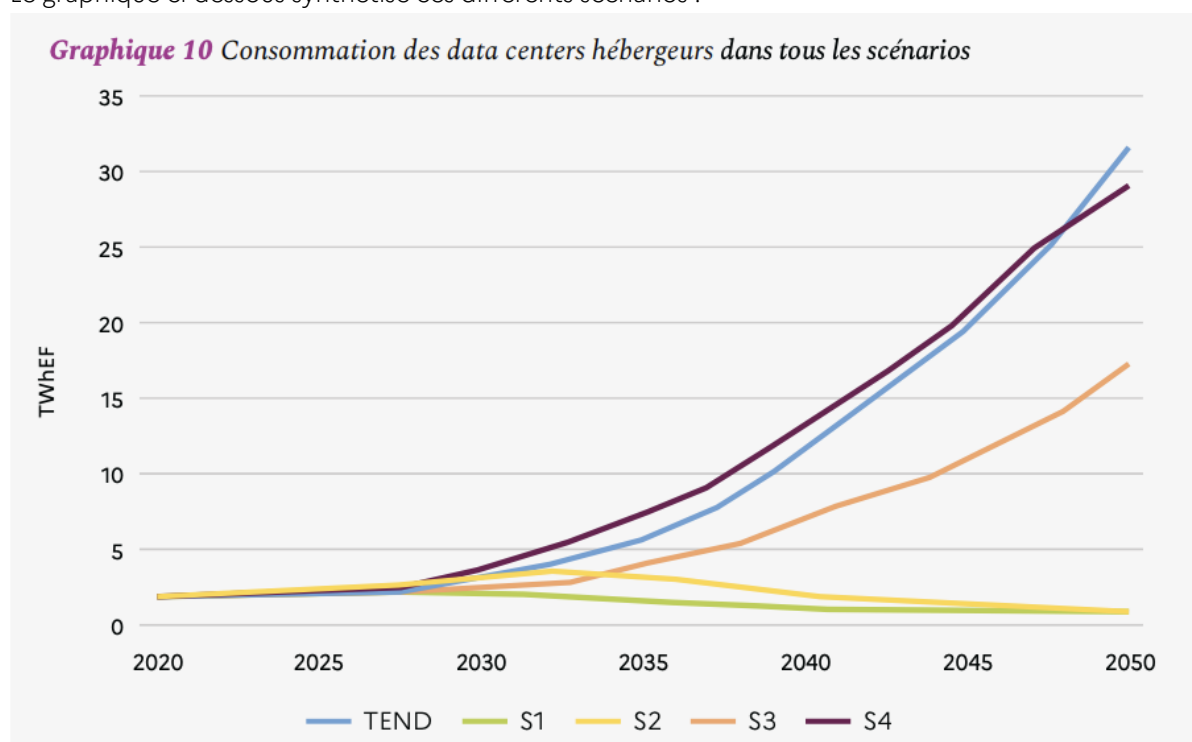
La croissance du volume de données ralentissait à partir de 2030 pour devenir nulle dans les années 2040 tout en se combinant à une meilleure efficacité énergétique. Cela aboutissait à une baisse de la consommation des centres de données **qui passait de 2 TWh en 2015 à 1 TWh en 2050.**

²⁶⁸ <https://librairie.ademe.fr/recherche-et-innovation/5072-prospective-transitions-2050-rapport.html>

Le scénario #3, intitulé "Technologies Vertes", se basait sur une très forte croissance du volume de données en ligne avec le scénario tendanciel, notamment dans les petits centres de données (5G et IoT). L'hypothèse de gain en efficacité énergétique ne compensait pas la croissance du volume de données. Ce scénario aboutissait à une forte augmentation de la consommation des centres de données, **qui passait de 2 TWh en 2015 à 17 TWh en 2050.**

Le scénario #4, intitulé "Pari Réparateur", se basait sur un accroissement sans précédent de la consommation énergétique des centres de données, dû à une forte digitalisation de l'économie. L'efficacité énergétique progressait rapidement mais ne compensait pas la très forte croissance du volume de données. Ce scénario aboutissait à une très forte augmentation des consommations liées aux centres de données pour atteindre **29 TWh en 2050.**

Le graphique ci-dessous synthétise ces différents scénarios :



2.1.2. "Évaluation de l'impact environnemental du numérique en France et analyse prospective" ²⁶⁹

Nous avons eu l'occasion d'étudier en détails ce rapport lors de la mission ADEME sur *L'Évaluation de l'impact environnemental du numérique de la région Grand Est*. Ce travail a abouti à la production de l'article Hubblo sur "Impacts importés des centres de données : l'angle mort des analyses territoriales des impacts du numérique".

Le modèle utilisé se base sur la surface des centres de données. Il utilise la classification suivante :

- Public Local

²⁶⁹ <https://librairie.ademe.fr/consommer-autrement/5226-evaluation-de-l-impact-environnemental-du-numerique-en-france-et-analyse-prospective.html>

- Public national
- Entreprises (hors digital)
- Colocations
- HPC

Le périmètre est donc différent de celui utilisé dans le rapport "Transition(s) 2050".

La prise en compte des centres de données hors "colocation" se base sur un inventaire INSEE des organisations et sur des estimations des besoins en m2 par typologie d'organisation (qui aboutit à une *Superficie interne*) et de l'externalisation de ces besoins (qui aboutit à une *Superficie externe*). Cependant, les sources de ces estimations ne sont pas publiques et viennent essentiellement de l'expertise du groupement qui a réalisé l'étude.

Voici une synthèse de ces hypothèses :

		Nb Entités	Besoins en m2 IT	% externalisé	Superficie interne (m2 IT)	Superficie externe (m2 IT)
Public Local					81885²⁷⁰	30965
	Conseils régionaux	13	50	70%	195	455
	Conseils départementaux	97	50	70%	1455	3395
	SDIS	97	50	70%	1455	3395
	COMUE	69	250	10%	15525	1725
	agglo	258	100	30%	18060	7740
	Mairies plus de 30k hab	277	50	70%	4155	9695
	CHRU	29	150	10%	3915	435
	CHU et centres hospitaliers	550	75	10%	37125	4125
Public national					65000	5000
	Ministères	80	250	0%	20000	0
	EPIC	30	500	10%	13500	1500
	ODAC	700	50	10%	31500	3500
Entreprises (hors digital)					311800	1417200

²⁷⁰ Ce total corrige une "coquille" dans le rapport ADEME/ARCEP Volet 2

	MIC (1-9 salariés)	3700000	0	100%	0	0
	PME (10-49)	172600	5	90%	86300	776700
	PME (50-99)	18100	10	80%	36200	144800
	PME (100-249)	10800	15	80%	32400	129600
	ETI (250-5000)	5700	50	70%	85500	199500
	Grandes Entreprises (> 5000)	238	1000	70%	71400	166600
Colocations		264			414175	0
HPC		18			10800	0
Total					883660	
Colocation - Cloud Total						1453165
dont France						414175

Avec ces hypothèses, les consommations d'électricité des centres de données situés sur le territoire national sont évaluées à 11,59 TWh pour l'année 2020 avec la répartition détaillée ci-dessous :

Type de datacenter	Public Local	Public National	Entreprises	Colocation	HPC	TOTAL
Superficie de salles informatiques m2	81 390 m2	65 000m2	311 800 m2	414 174 m2	10 800 m2	883 165 m2
Densité kW/baie	3	4,5	4	5	15	
Taux de charge	40%	35%	50%	50%	60%	
PUE	1,93	1,93	1,93	1,55	1,17	1,69
Consommations électriques IT TWh	0,34	0,36	2,19	3,63	0,34	6,85
high end servers						435 164 902
mid range servers						940 758 672
volume servers						2 937 479 344
application specific hardware						676 339 197
3.5 Hard disk						400 097 785
2.5 Hard disk						370 022 885
SSD						289 678 796
Storage controller						332 002 326
Network port 1Gbit						5 360 390
Network port 10 Gbit						93 271 508
Network port 40/100 Gbit						177 904 269
Storage Ports						196 726 797
Network port 1Gbit						5 360 390
Consommations électriques Environnement technique TWh	0,32	0,33	2,03	2,00	0,06	4,74
Consommations de fioul m3	66	69,2	422	562	40	1 159
Consommations électriques totales TWh	0,66	0,69	4,22	5,62	0,40	11,59

Tableau 58 - Evaluation des consommations annuelles d'énergie des datacenters par type

À partir de cette base le volet 3 de l'étude présente une analyse prospective à 2030 et 2050.

Les mêmes catégories de centres de données sont considérées mais avec l'ajout de la catégorie "Edge Computing". Elle correspond aux centres de données cloud de proximité. La définition étant encore relativement floue, l'hypothèse retenue pour évaluer la superficie en m2 de centres de données Edge Computing est liée à l'évolution des équipements IoT.

Pour 2030 et 2050, des taux de croissance moyen de la superficie des centres de données ont été appliqués pour faire évoluer les hypothèses de 2020 :

- 4% pour les centres de données publics et ceux des entreprises de moins de 250 salariés
- 6% pour les ETI
- 8% pour les grandes entreprises
- 8% entre 2020 et 2030, puis 3% entre 2030 et 2050 pour les centres de données de colocation/cloud
- 4% pour les centres de données de type HPC

En plus de ces hypothèses de croissance des surfaces, l'étude évalue l'évolution des 3 paramètres identifiés dans l'estimation 2020 :

- la puissance électrique installée par baie
- le PUE
- le ratio entre la puissance électrique réellement appelée sur la puissance électrique disponible

Avec l'ensemble de ces hypothèses, la **consommation d'électricité des centres de données situés sur le territoire national est estimée dans un scénario tendanciel à 16,38 TWh en 2030 et 38,96 TWh en 2050 soit une progression de 38,96% et 236% par rapport à 2020.**

À noter que la part des centres de données traditionnels décroît en passant de 48% en 2020 à 3,9% en 2030 et 0% en 2050.

2.1.3. Études RTE

RTE réalise régulièrement différents exercices de prospective.

Futurs énergétiques 2050 ²⁷¹

L'État français doit réaliser des exercices de planification de long terme depuis 1999²⁷², dans le cadre de ses engagements européens. Ces prospectives ont évolué au fil du temps avec des exercices comme la Stratégie nationale bas-carbone et la Programmation pluriannuelle de l'énergie. RTE a également fait évoluer ses méthodes de prospective en publiant en 2021, un travail approfondi appelé Futurs énergétiques, pour atteindre le Zéro Émissions Nettes (ZEN) en 2050, en se basant sur les objectifs nationaux de la SNBC 2 et de la PPE. La prospective n'est plus un tendanciel mais une scénarisation pour atteindre des objectifs fixés politiquement. Elle inclut des dimensions sociales, environnementales, économiques, foncières...

Dans cette publication²⁷³, l'hypothèse retenue était celle d'une hausse marquée de la demande électrique des centres de données, en parallèle des améliorations en termes d'efficacité énergétique.

Ce scénario conduirait à un triplement de la demande d'électricité des centres de données entre 2019 et 2050, pour atteindre 9 TWh.

Les bilans prévisionnels

Le Bilan Prévisionnel permet d'assurer l'adéquation entre offre et demande, de prévoir une trajectoire de consommation, et de respecter le ratio de défaillance permis au niveau européen (3 heures).

Ce BP est réalisé chaque année. Le BP 2023 est analysé plus bas. Il couvre la période 2023-2030-2035. Le BP 2023 a enrichi et précisé les études réalisées pour Futurs énergétiques 2050. Il inclut aussi 200 scénarios climatiques proposés par une base de données Météo France.

Zoom sur le Bilan Prévisionnel 2023 chapitre 2 Consommation²⁷⁴

Dans, le BP 2023, la consommation des centres de données dépassent largement ce que Futurs énergétiques 2050 avait prévu : d'une situation actuelle autour de **11 TWh (évalué à 10 TWh en 2023, mais remontée à 11 en 2024)**, elle pourrait atteindre

- **15 à 20 TWh en 2030,**
- **et entre 23 et 28 TWh en 2035.**

La trajectoire A-référence est utilisée (scénario « accélération réussie ») pour le SDDR.

²⁷¹ <https://rte-futursenergetiques2050.com/documents>

²⁷² <https://www.ecologie.gouv.fr/politiques-publiques/scenarios-prospectifs-energie-climat-air>

²⁷³ <https://www.rte-france.com/analyses-tendances-et-prospectives/bilan-previsionnel-2050-futurs-energetiques#Lesdocuments>

²⁷⁴ <https://analysesetdonnees.rte-france.com/bilan-previsionnel2023>

Il existe deux autres trajectoires qui ne changent pas drastiquement les projections pour les centres de données :

La trajectoire B « atteinte partielle » où il y a un retard sur la A notamment sur le rythme d'électrification et sur les gains en efficacité énergétique, mais dans le même cadre d'incitation publique.

La trajectoire C est appelée « mondialisation contrariée » : une moindre croissance de l'économie qui se traduit par une moindre augmentation des besoins en énergie (effet volume plus faible), à la fois en lien avec les consommations des industries et du secteur tertiaire, mais également par une moindre consommation des ménages, contraints par la baisse de leur pouvoir d'achat. En outre, ces mêmes difficultés sont de nature à retarder les investissements dans les solutions bas-carbone et ainsi la décarbonation des secteurs du transport, des bâtiments et de l'industrie.

Dans ce scénario, la croissance des nouvelles technologies de l'information et de la communication (NTIC) se poursuit, mais de manière moindre que celle anticipée via les demandes de raccordement : **leur consommation s'établit à 19 TWh environ en 2035, contre 26 TWh dans les trajectoires A et B.**

Les SDDR : schéma décennal de développement du réseau ²⁷⁵

RTE réalise également un exercice de planification appelé SDDR pour planifier ses investissements, et fixer une doctrine. Les dimensions locales et territorialisées de ce document sont très importantes. Son élaboration fait partie des missions légales de RTE, dont le cadre est défini par l'article 51 de la directive européenne « marché intérieur de l'électricité » transposé à l'article L. 321-6 du code de l'énergie et par les articles L. 121- 8 et R. 122-17 du code de l'environnement.

Le SDDR 2025 souligne que 40 projets de centres de données ont déjà un accès au réseau à hauteur de 5,5 GW, et que de nombreux projets sont à l'étude sous l'influence du développement des usages liés à l'IA. ²⁷⁶

“Ils sont principalement localisés en Ile-de-France et dans le Nord de Marseille, en raison de leur proximité avec les centres décisionnels et les infrastructures numériques. Les projets demandent une puissance de 100 à 200 MW et prévoient de réaliser leur montée en puissance sur une durée longue (7 à 10 ans). Néanmoins, depuis quelques mois, le développement de l'intelligence artificielle entraîne une évolution majeure des demandes avec des besoins en puissance très supérieurs (400 à 500 MW, voire 1 GW). Par ailleurs, 25 GW d'études supplémentaires pour des demandes de raccordement du secteur numérique sont en cours d'instruction par RTE. Une partie importante de ces demandes se situe dans des zones urbaines denses, où le réseau est déjà proche de la saturation et où le raccordement peut être complexe (franchissement de voies routières et ferroviaires, habitat dense, sous-sols déjà très fortement utilisés).”

²⁷⁵ <https://www.rte-france.com/analyses-tendances-et-prospectives/le-schema-decennal-de-developpement-du-reseau>

²⁷⁶ <https://assets.rte-france.com/prod/public/2025-06/2025-06-06-sddr-rapport-complet.pdf> p.67

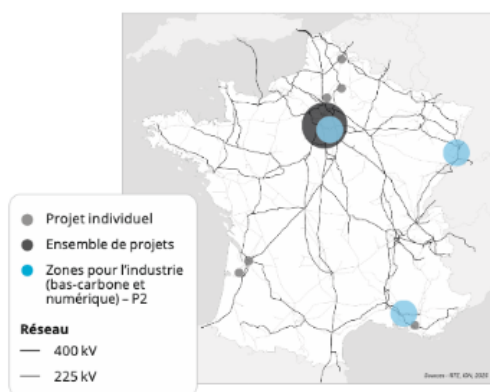


Figure 5.11 – Carte des demandes de raccordement pour les infrastructures numériques (projets ayant *a minima* fait une demande de proposition technique et financière auprès de RTE)

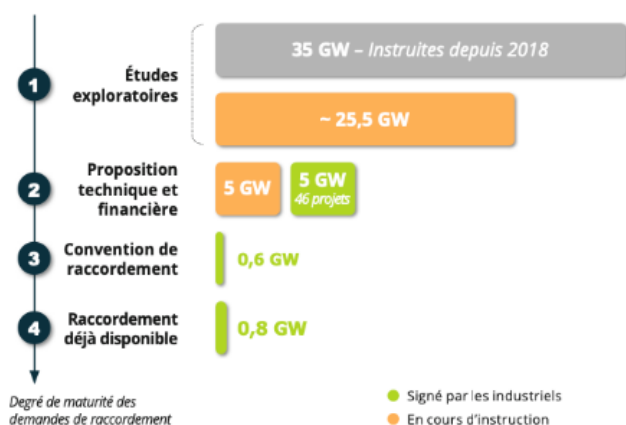


Figure 5.12 – État des lieux des demandes de raccordement pour les infrastructures numériques

2.1.4. Rapport de l'AIE "Electricity 2024, Analysis and forecast to 2026"

Ce rapport de l'AIE est clef puisqu'il infléchit pour la première fois les prévisions de l'agence concernant la consommation énergétique des centres de données. Si entre 2010 et 2020, l'AIE tablait sur une relative stabilité de la consommation énergétique mondiale autour de 200 TWh, en se basant sur des hypothèses de gains en efficacité qui compensaient la forte croissance en terme de trafic et de charge, ce rapport modifie radicalement la courbe d'évolution de la consommation mondiale des centres de données : **il estime une consommation de 460 TWh en 2022 au niveau mondial et prévoit une croissance forte à court terme pour atteindre en 2026 entre 620 et 1050 TWh suivant les scénarios.** Les prévisions 2010-2020 étaient basées sur le modèle Masanet (que nous abordons ci-dessous). Ce changement de trajectoire démontre l'importance de bien comprendre les modèles disponibles, leurs hypothèses et leurs limites.

2.1.5. Les rapports d'orientation stratégique du CIGREF de 2020,²⁷⁷ 2022²⁷⁸ et 2023²⁷⁹

Les rapports du CIGREF sont intéressants pour la dimension prospective. Ils étudient des hypothèses de rupture qui pourraient se réaliser dans le domaine du numérique à échéance moyen terme (10 ou 20 ans). Le rapport de 2023 adresse par exemple des enjeux technologiques (l'informatique quantique, le déploiement de la 6G, ...), environnementaux (tempête solaire, recyclage réglementaire à 100%, ..), géopolitiques (rupture des échanges avec la Chine, repli sur soi des États-Unis, ...), économiques (effondrement des modèles éco des GAFAM, indépendance stratégique de l'Europe, ...) ou sociétaux (surveillance globale des sociétés via le numérique, modification radicale du travail dû au déploiement de l'IA générative, ...).

Cependant, ces rapports ne réalisent pas de projections quantifiées de ces ruptures.

²⁷⁷ https://www.cigref.fr/wp/wp-content/uploads/2020/10/cigref_rapport_orientation_strategique_2020_web_page_a_page.pdf

²⁷⁸ <https://www.cigref.fr/wp/wp-content/uploads/2022/09/Rapport-dOrientation-Strategique-du-Cigref-WEB-CIGREF2022.pdf>

²⁷⁹ <https://www.cigref.fr/wp/wp-content/uploads/2023/10/Rapport-dOrientation-Strategique-du-Cigref-CIGREF2023-WEB-FINAL.pdf>

2.1.6. Energy-efficient cloud computing technologies and policies for an eco-friendly cloud market ²⁸⁰

Ce rapport du Borderstep Institute est à la base de la *"Directive on energy efficiency and amending Regulation (EU)"*.

Borderstep estime que par rapport à 2018, la consommation énergétique des centres de données devrait augmenter de 21% pour atteindre 92,6 TWh/an en 2025 au niveau européen.

Il estime aussi que la part du Cloud passerait de 35% en 2018 à 60% en 2025 tandis que la part des centres de données Edge devrait également augmenter de manière significative pour atteindre, en 2025, 12% de la consommation énergétique de l'ensemble des centres de données de l'UE-28.

Sur le plan géographique, les plus gros centres de données sont situés dans les pays du nord et de l'ouest de l'Europe. Ces régions représentaient 82% de la consommation énergétique des centres de données en 2018.

Le rapport souligne néanmoins que, si tous les potentiels sont exploités, il sera possible de réduire la consommation énergétique des centres de données à des niveaux inférieurs à ceux de 2010. Cela passe par des innovations telles que le refroidissement avancé et l'optimisation pilotée par l'IA mais aussi par des initiatives politiques, telles que des normes d'efficacité plus strictes et des incitations pour les énergies renouvelables.

2.2. Les principaux modèles utilisés

Les différents rapports - notamment à l'échelle mondiale - se basent souvent sur les mêmes modèles issus d'études académiques. Les 5 principales approches sont :

2.2.1. Modèle Masanet

"Recalibrating global data center energy use estimates." de Masanet, E., Shehabi, A., Lei, N., Smith, S., and J.G. Koomey (2020)

Le modèle Masanet se base sur des données de référence de CISCO puis sur une modélisation des centres de données évaluée sur les données des parcs IT américains (Arman Shehabi et al., "United States data center energy usage report", Lawrence Berkeley National Laboratory, 2016).

Le stock de serveurs et la répartition par type de centres de données sont estimés principalement à partir des données CISCO et notamment le "Cisco Global Cloud Index: Forecast and Methodology, 2016-2021: White Paper, 2018" et de la modélisation des serveurs issue de "United States Data Center Energy Usage Report" de Shehabi.

La notion de workload (ou de "compute instance") est définie par CISCO: ".....For the purposes of quantification, we consider each workload and compute instance being equal to a virtual machine or a container."

²⁸⁰ <https://digital-strategy.ec.europa.eu/en/library/energy-efficient-cloud-computing-technologies-and-policies-eco-friendly-cloud-market>

La "baseline" du nombre de workload par type d'infrastructure est également fournie par CISCO. L'étude du rapport CISCO montre que ces données ne sont pas des données primaires mais sont issues de données d'analystes.

Les données sont ouvertes et un modèle Excel est disponible.

Forces : Le modèle est ouvert et un calculateur est disponible. Il se base sur des données "physiques" facilement (re)dimensionnables : des workloads, des volumes de serveurs ou de stockage.

Faiblesses : Le modèle ne se base pas sur des données primaires. Les données initiales sont principalement celles de CISCO qui elles-mêmes se basent sur des données provenant de cabinets d'analystes.

Les données sont anciennes : entre 2010 et 2016, et non maintenues.

2.2.2. Modèle Andrae

S.G.A. Andrae, T. Edler, *"On global electricity usage of communication technology: trends to 2030"*, Challenges, 6 (2015), pp. 117-157

Le modèle Andrae se base sur un ratio de croissance du trafic de données et un ratio de gain en efficacité énergétique.

Les données de référence sont celles de CISCO. Elles datent de 2010.

Les données sont ouvertes et un modèle Excel est disponible.

Forces : Le modèle est ouvert et un calculateur est disponible. Il se base sur des données facilement redimensionnables telles que volume de trafic et gain en efficacité énergétique.

Faiblesses : Le modèle ne se base pas sur des données primaires. Les données initiales sont principalement celles de CISCO.

Les données sont anciennes : entre 2010 et 2016, et non maintenues.

2.2.3. Modèle Belkhir

L. Belkhir, E. Ahmed, *"Assessing ICT global emissions footprint: trends to 2040 & recommendations"*

Le modèle Belkhir se base sur une consommation de référence de 2008 (254 GWh), estimée par l'étude de Vereecken, et par une estimation de croissance annuelle de 10,6% issue des études de marché (2014 et 2015) de Technavio.

À noter que le modèle ignore la phase de fabrication des centres de données - partie IT incluse - dans ses estimations d'empreinte GES globale, alors qu'elle est prise en compte pour les terminaux.

Les principales données primaires de référence datent de 2008.

Les données et hypothèses sont ouvertes mais il n'existe pas de modèle Excel.

Forces : Le modèle et les hypothèses sont clairement décrites. Le modèle se base sur la consommation énergétique qui est facilement manipulable.

Faiblesses : Le modèle ne se base pas sur des données primaires.

Les données sont très anciennes.

2.2.4. Modèle Malmodin

J. Malmodin, D. Lundén, *"The energy and carbon footprint of the global ICT and E&M sectors 2010-2015"*

Le modèle Malmodyn se base sur une croissance de la capacité de "processing" par unité d'énergie combinée à la croissance du nombre de serveurs. Comme le modèle Masanet, il utilise la modélisation IT issue de l'étude Shehabi.

À noter que le modèle inclut aussi l'impact des personnes en charge d'opérer les centres de données. Les données primaires datent de 2015-2016, et proviennent pour partie des clients d'Ericsson. Elles ne sont, de ce fait, pas ouvertes.

Forces : Le modèle est clairement décrit et il se base des éléments "physiques" facilement manipulables tels que le nombre de serveurs et leur capacité de processing.

Faiblesses : Le modèle et les données sont relativement fermés.
Les données sont anciennes même si plus récentes que d'autres modèles.

2.2.5. Modèle Borderstep

Energy consumption of data centers worldwide, Hintemann & Hinterholzer, Borderstep Institute

Le modèle Borderstep se base sur une modélisation bottom-up très structurée des centres de données à partir de données allemandes et européennes. Plusieurs types de centres de données sont identifiés et les données sont mises à jour régulièrement. Cette modélisation est croisée avec des données marché issues principalement de CBRE, Techconsult, IDC and EITO.

Les données primaires de référence s'échelonnent de 2010 à 2017.

Comme pour le modèle de Masanet, les serveurs sont caractérisés suivant trois catégories : les serveurs de base ou d'entrée de gamme (*volume servers*), d'une durée de vie de 5 ans, les serveurs de gamme moyenne, d'une durée de vie de 8 ans, et les serveurs haut de gamme, d'une durée de vie de 12 ans.

À noter que la modélisation Borderstep prend aussi en compte le fait que les serveurs sont souvent utilisés dans une "seconde vie" au sein des centres de données.

Les volumes d'IT répertoriés sont les suivants :

EU28	2010	2015	2018	2020	2025
High-end servers	12,409	16,585	19,241	21,303	26,302
Mid-range servers	845	1,213	1,564	1,873	2,616
Volume servers	153	182	210	228	268
Application specific hardware	n,a,	1,560	2,539	2,880	3,540

Table 40 - Average power consumption of servers in the EU28 in watts (2020, 2025: forecast)

EU28	2010	2015	2018	2020	2025
3.5" hard disk	8.9	8.6	8.5	8.4	8.5
2.5" hard disk	5.9	5.4	5.2	5.2	5.3
SSD	5.6	5.2	4.9	4.8	4.4
Storage Controller	320.1	312.9	302.2	298.7	301.5

Table 41 - Average power consumption of hard disks and storage controllers in the EU28 in watts (2020, 2025: forecast)

Forces : Le modèle est très clairement décrit, avec un fort niveau de détail des éléments manipulés. Un calculateur web²⁸¹ est accessible même s'il n'est pas transparent sur les calculs sous-jacents.

Faiblesses : Seules certaines données sont ouvertes et la plupart ne sont pas ou peu exploitables dans un nouveau modèle. Il n'existe pas de calculateur Excel.

2.2.6. Le modèle Total Energy Model (TEM) de l'EDNA

Le modèle Total Energy Model (TEM) de l'initiative EDNA a permis de documenter de nombreux scénarios qui intègrent comme principaux paramètres le nombre d'objets connectés, la nature des données qui transitent (streaming ou non) et le gain en efficacité énergétique.

La partie centre de données est incluse dans une catégorie appelée "Upstream" qui intègre également la partie WAN. Les parties Streaming et Non streaming sont bien identifiées.

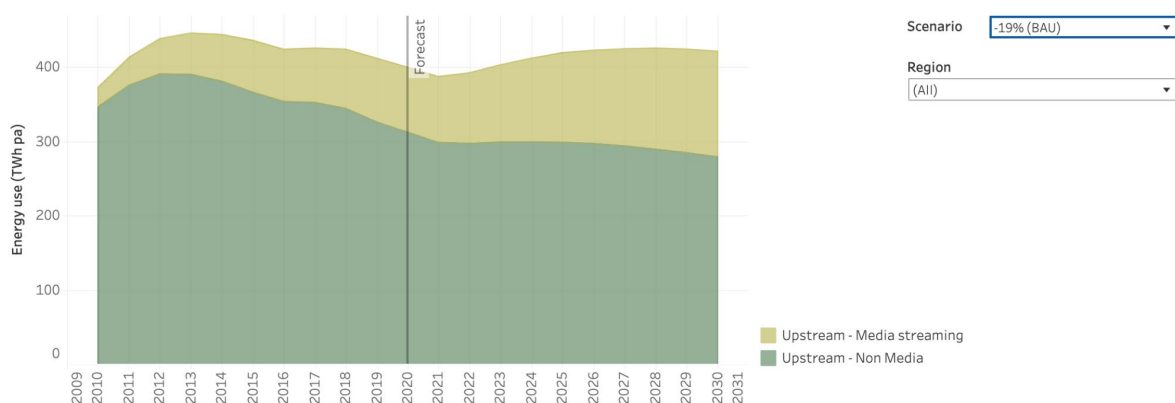
Le principe du modèle est assez similaire à celui d'Andrae qui se base sur le volume de données qui transite (au sens CISCO) et sur l'évolution de l'efficacité énergétique.

Le modèle TEM ajoute cependant des paramètres d'usage qui permettent d'aboutir aux scénarios décrits ci-dessus.

La baseline est l'année 2020. La consommation électrique totale est de 400 TWh dont 88 TWh de streaming.

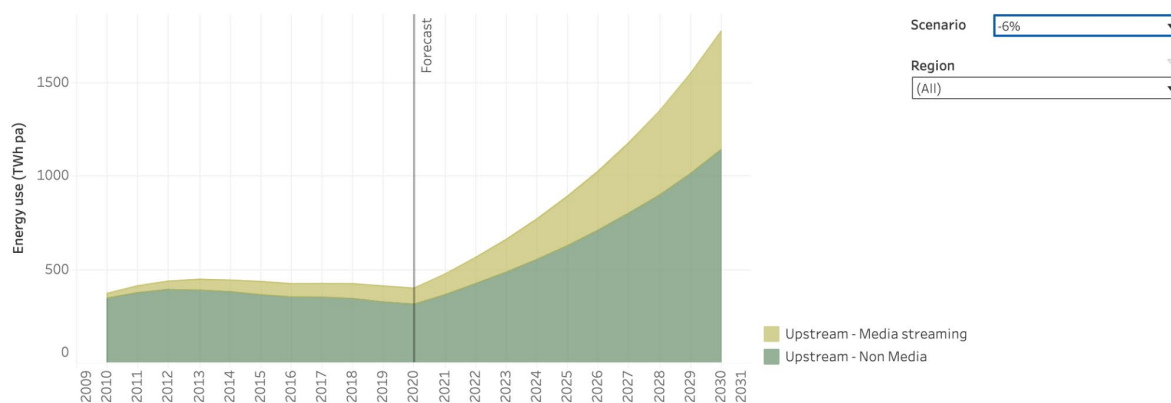
Il est ensuite possible de jouer sur les paramètres d'efficacité énergétique (de 0 à 28% de gain annuel) et de la zone géographique.

Avec la valeur par défaut (19% de gain annuel en efficacité énergétique), on constate une stagnation de la consommation totale pour atteindre 422 TWh en 2030 :



En appliquant un gain annuel en efficacité énergétique plus conservateur de 5%, on note alors une croissance importante pour atteindre 1770 TWh en 2030 :

²⁸¹ <https://www.tempro-energy.de/erbet/results.php>



Le modèle complet sous-jacent n'est pas disponible et ne permet donc pas de modifier les autres paramètres d'usage.

Le constat principal est que l'hypothèse de gain en efficacité énergétique a un impact fort. À noter que le gain annuel de 19%, valeur par défaut (BAU), est supérieur au "best case scenario" du modèle d'Andrae (15%).

Forces : Le modèle est clairement décrit avec une approche "simple", similaire au modèle Andrae : il se base sur le trafic et sur l'évolution de l'efficacité énergétique.

Les données initiales sont plus récentes (2020) que celles des autres modèles.

Il existe plusieurs paramètres d'usage.

Un calculateur web est fourni.

Faiblesses : Peu de données sont ouvertes et la plupart ne sont pas ou peu exploitables dans un nouveau modèle. Il n'existe pas de calculateur Excel.

2.3. Prospectives climatiques

Selon les derniers rapports du GIEC, la fréquence des événements climatiques extrêmes va continuer à augmenter. Les centres de données sont à la fois des victimes potentielles de ces risques (tempêtes, inondations, coupures électriques, dommages matériels...) mais les aggravent également (artificialisation des sols qui stockent moins de carbone, ne laissent pas l'eau s'infiltrer et accélère ainsi le ruissellement, aggravation de la chaleur urbaine par le rejet de chaleur fatale, pollution de l'air notamment dans les clusters quand les générateurs de secours fonctionnent, ce qui devrait augmenter avec des occurrences climatiques extrêmes plus fortes). Sur ces sujets, le 6ème rapport du GIEC paru en mars 2023²⁸² détaille la prospective climatique et environnementale. Un sondage récent de l'Uptime Institute sur les stratégies de résilience des centres de données²⁸³ précise de son côté les réponses et préoccupations de l'industrie.

Plusieurs hypothèses peuvent être tirées de ces travaux du GIEC :

- Les logiques d'implantation des centres de données pourront être de plus en plus liées à l'évitement maximum des risques environnementaux, et se concentrer dans les territoires que l'on suppose moins fragiles,

²⁸² https://www.ipcc.ch/report/ar6/syr/downloads/report/IPCC_AR6_SYR_SPM.pdf

²⁸³ <https://datacenter uptimeinstitute.com/rs/711-RIA-145/images/2023.Sustainability.Survey.Report.pdf?version=0>

- Dans le même temps, si les centres de données continuent à s'installer dans des zones fragiles (sécheresses, tempêtes...), ils risquent d'aggraver les conflits d'usages avec les populations voisines, voire aggraver les risques déjà présents.

2.4. Études prospectives sur la ressource en eau

Les études prospectives sur l'empreinte eau centres de données pourront être mises en relation avec celles concernant plus globalement la ressource eau d'une région, d'un pays.

Dans l'article *The Water Footprint of Data Centers*²⁸⁴, les auteurs écrivent :

"La consommation mondiale d'énergie devrait augmenter de 35 à 40 % et celle d'électricité de 70 % d'ici 2035. En conséquence, les prélèvements d'eau pour la production d'électricité thermique devraient augmenter de 40 % et la consommation d'eau de plus de 100 % d'ici à 2050. À l'échelle mondiale, le secteur de l'énergie devrait voir son empreinte hydrique annuelle augmenter de 37 % à 66 % d'ici 2035, l'empreinte hydrique par unité d'énergie produite augmentant de 5 % à 10 %. Aux États-Unis et en Europe, 50 % et 43 % respectivement des prélèvements d'eau douce sont utilisés à des fins de refroidissement pour la production d'électricité. Cette demande croissante signifie que les consommateurs d'énergie et d'eau sont en concurrence étroite pour les ressources. L'expansion des DC et leur consommation croissante d'énergie concurrencent les acteurs établis."

S'il est plus question ici de prélèvements que de consommations, il existe une relation à projeter entre croissance centres de données et croissance des consommations d'eau, en lien avec l'évolution globale et locale de cette ressource.

Le rapport *"Prospective de la demande en eau en 2050"*, France Stratégie, avril 2024²⁸⁵, fait le point sur les consommations et les prélèvements d'eau actuels, et devrait rendre bientôt une note sur la prospective 2050. La DRIAS de son côté mène le projet *"Futurs de l'eau"* qui a pour vocation de mettre à disposition des projections hydrologiques des eaux de surface et souterraines, réalisées dans le cadre du projet national Explore2*, ainsi que l'ensemble des informations utiles à leur bonne utilisation, sous différentes formes graphiques ou numériques.

²⁸⁴ écrit par Bora Ristic, Kaveh Madani and Zen Makuch, <https://www.mdpi.com/2071-1050/7/8/11260>

²⁸⁵ <https://www.strategie.gouv.fr/publications/prelevements-consommations-deau-enjeux-usages>

3. Méthodologie du modèle

3.1. Définition de l'objectif et du périmètre de l'étude

L'objectif de cette étude Prospective d'évolution des consommations des centres de données en France de 2024 à 2060 est :

1. De permettre aux pouvoirs publics d'appréhender les transformations à l'œuvre actuellement au sein du secteur des centres de données
2. D'anticiper et planifier des politiques publiques au regard de 5 scénarios d'évolution possibles :
 - a. Un scénario tendanciel
 - b. 4 scénarios de transition écologique ayant tous pour objectif d'atteindre l'objectif de Zéro Émissions Nette en 2050, en empruntant des chemins différents pour y parvenir (différents degrés de sobriété, d'efficacité technologique, de recours aux énergies décarbonées et aux éventuelles technologies de compensation...)

Pour cela, cette étude permet de :

1. Dresser un état de l'art des connaissances actuelles sur l'empreinte environnementale des centres de données en France et dans le monde
2. Inventorier et modéliser à l'échelle nationale la consommation des centres de données présents sur le sol français (en Europe), en constituant une base de données des centres de données présents sur le territoire pour l'année de référence 2024.
3. Développer un modèle prospectif et réutilisable permettant d'estimer, par typologie de centre de données et par année jusqu'en 2060, leur consommation énergétique. Ce modèle repose sur des hypothèses d'évolution des usages numériques, de l'implantation des centres en France, ainsi que de l'importation de services depuis des centres de données situés à l'étranger pour répondre à la demande nationale
3. Modéliser 5 scénarios estimatifs afin d'explorer différentes trajectoires possibles :
 - a. Un scénario tendanciel (appelé aussi BAU pour Business As Usual), prolongeant les dynamiques actuelles
 - b. 4 scénarios de transition, basés sur les scénarios "[Transition\(s\) 2050](#)", imaginant des évolutions contrastées :
 - i. Génération Frugale
 - ii. Coopération Territoriale
 - iii. Technologies Vertes
 - iv. Pari Réparateur

3.2. Indicateurs de consommation considérés

3.2.1. Consommation d'électricité

Dans cette étude, nous estimons la consommation énergétique (exprimée en TWh d'énergie finale consommée) des centres de données pour l'année de référence 2024 sur la base méthodologique décrite au chapitre [Modélisation de la situation initiale](#).

Sur la base de cette consommation d'énergie électrique estimée par catégorie de centre de données, nous réalisons dans le modèle CLIK une projection d'évolution des consommations, année après année, pour chacun des 5 scénarios analysés. La méthodologie du modèle CLIK est décrite au chapitre [Court, moyen, long terme : définition de séquences d'évolution](#).

3.2.2. Émissions de Gaz à Effet de Serre

Les émissions de Gaz à Effet de Serre (GES), exprimées en tCO₂eq., mesurent la quantité de gaz contribuant au réchauffement climatique. L'accord de Paris de 2015, signé par la France et l'ensemble des pays européens, engage les États signataires à réduire et limiter leurs émissions de GES afin de maintenir « l'augmentation de la température moyenne mondiale **bien en dessous de 2°C au-dessus des niveaux préindustriels** » et de poursuivre les efforts « pour limiter l'augmentation de la température à 1,5°C au-dessus des niveaux préindustriels »²⁸⁶. Ces objectifs, fondés sur les rapports du GIEC, visent à contenir le réchauffement climatique en deçà de seuils critiques, afin de réduire les risques de franchir des « points de bascule » aux conséquences irréversibles...²⁸⁷

Dans le cadre de cette étude, les émissions de GES des centres de données sont estimées à partir de leur consommation électrique. L'évolution du mix électrique est prise en compte dans le temps dans le modèle, pour estimer l'intensité carbone des émissions de GES liées à la consommation électrique des centres de données. Sont ainsi pris en compte de façon distincte :

- Les émissions de GES induites par la production d'énergie pour le mix électrique français (alimentant les centres de données du territoire français).
- Les émissions de GES induites par la production d'énergie pour alimenter les centres de données situés à l'extérieur du territoire français. Pour ce cas précis²⁸⁸, nous nous sommes basés sur la répartition géographique documentée dans le modèle Masanet²⁸⁹ et pour l'évolution des mix électrique, sur le traitement DGEC à partir des scénarios utilisés dans le sixième rapport d'évaluation du GIEC et hébergés par l'IIASA.²⁹⁰

3.2.3. Consommation foncière

Comme indiqué au chapitre [Foncier, spatialités](#), les centres de données en projet aujourd'hui s'implantent sur des terrains de plus en plus vastes. Toutefois, les indicateurs liés à la consommation foncière des centres de données restent encore peu suivis et communiqués au sein du secteur.

²⁸⁶ <https://unfccc.int/fr/a-propos-des-ndcs/l-accord-de-paris>

²⁸⁷

https://fr.wikipedia.org/wiki/Points_de_basculement_dans_le_syst%C3%A8me_climatique#%C3%89l%C3%A9ments_de_basculement ; <https://climatetippingpoints.info/2022/09/09/climate-tipping-points-reassessment-explainer/> ;

²⁸⁸ Une évolution potentielle du modèle serait de prendre en compte des mix électriques différents selon les scénarios, par exemple un scénario avec un mix plus européen proportionnellement que les autres.

²⁸⁹ <https://github.com/emasanet/GlobalDCAnalysis>

²⁹⁰ <https://data.ece.iiasa.ac.at/ar6/>

Au sein de cette étude, nous avons pu considérer la consommation foncière des centres de données actuellement répertoriés et identifiés dans notre base de données, en France en 2023. Cependant, le manque de données à disposition et de maturité de cet indicateur ne nous a pas permis d'estimer l'évolution de la consommation foncière des centres de données au regard des usages au sein de notre modèle prospectif.

Au regard des connaissances actuelles, cet indicateur nous a donc permis de traduire des informations (de surface de plancher et de surface IT) disponibles en estimation de la consommation d'énergie par tranche de surface IT, pour l'année de référence 2023, en France.

Ces informations sont exprimées en m² de surface de plancher ou en m² de surface IT.

3.3. Principes généraux du modèle

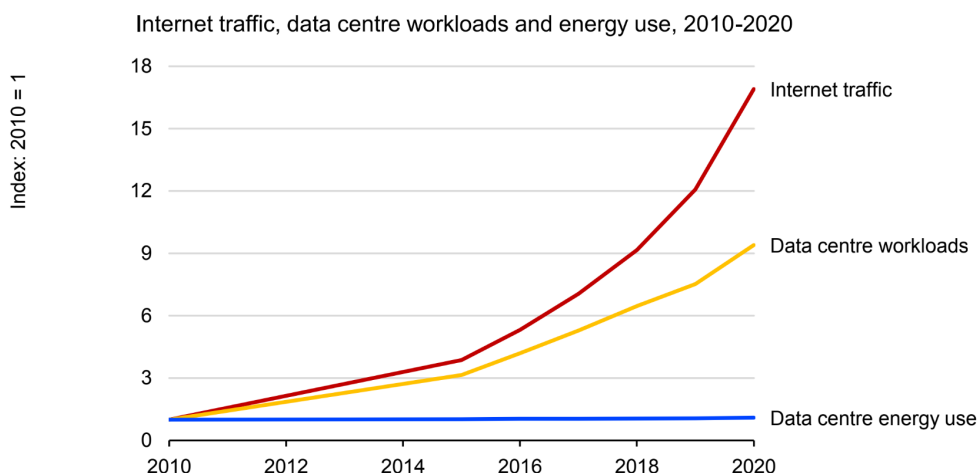
Notre modèle prospectif, nommé modèle CLIK, permet de modéliser des scénarios d'évolution des consommations des centres de données en France sur la base de l'évolution des usages numériques et d'un certain nombre de paramètres, détaillés ci-après.

Nous présentons ici les principes généraux du modèle et les choix méthodologiques retenus pour sa conception et son architecture.

3.3.1. Évolution usage vs efficacité

Quantifier l'usage est complexe. Nous distinguons deux approches dans la littérature. L'approche par workload, que l'on retrouve chez Masanet, Borderstep et dans les rapports de l'IEA, ou l'approche par volume de données transmises (data traffic), utilisée par Andrae, EDNA.

Comme précisé précédemment [dans le rapport](#), la notion de workload est définie par Cisco comme *"un ensemble de ressources, virtuelles ou physiques, affectées à l'exécution d'une application ou à la fourniture d'un service numérique, à un ou plusieurs utilisateurs."* La notion de trafic représente le trafic IP, c'est-à-dire le volume du flux de données transmis via le protocole internet. De la même manière, la notion d'efficacité est complexe à estimer dans une démarche prospective. Ainsi, jusqu'à 2022, l'IEA tablait sur un gain d'efficacité tel qu'il compensait en totalité les croissances d'usage que ce soit en volumes de données ou de workloads :

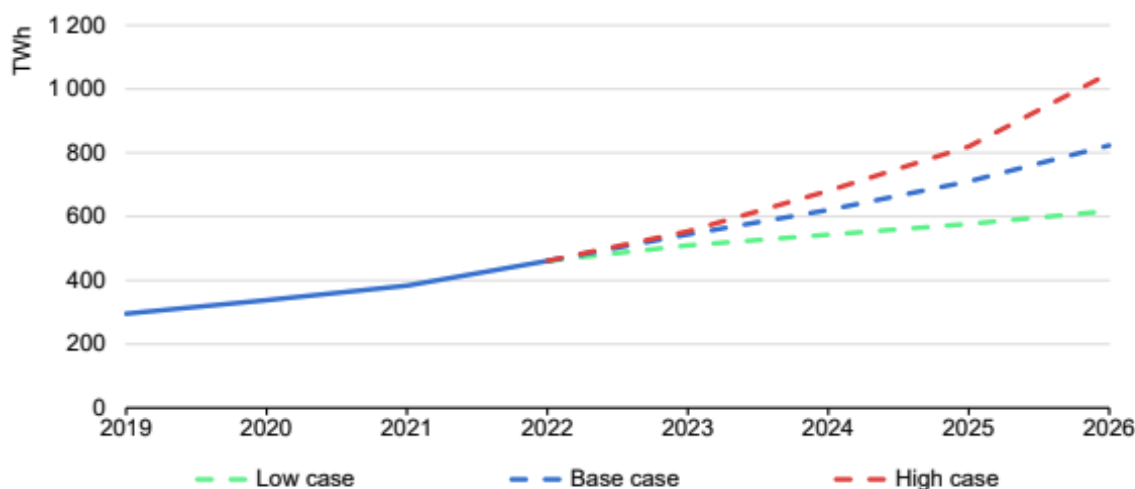


Sources: Masanet et al. (2020). Recalibrating global data center energy-use estimates. IEA (2021). Data centres and data transmission networks; Cisco (2018). Global Cloud Index: Forecast and Methodology, 2016-2021; Cisco (2019). Visual Networking Index: Forecast and Trends, 2017-2022.
 Note: Figures exclude cryptocurrency mining

Globally, data centres used an estimated 200-250 TWh in 2020, or around 1% of global electricity use

Pour autant, en 2024, l'IEA a finalement estimé que ces gains n'étaient à présent plus suffisants et a fourni de nouvelles estimations où les gains en efficacité ne compensaient pas la croissance des usages :

Global electricity demand from data centres, AI, and cryptocurrencies, 2019-2026



IEA. CC BY 4.0.

Il faut noter ici que l'AIE intégrait les centres de données, y compris IA, et les crypto-monnaies. En 2025, elle publie des prévisions concernant uniquement les centres de données (dont IA). Dans son scénario de base, la consommation globale d'électricité des centres de données atteint 945 TWh en 2030, soit presque 3% de la consommation électrique mondiale.²⁹¹

Nous avons tenu compte de ces évolutions dans notre modèle, notamment en faisant les choix suivants :

- Prendre en compte les évolutions des usages ET des gains en efficacité
- Préférer une approche par évolution des workloads pour rendre compte des évolutions potentielles des usages, plutôt que l'approche par trafic IP (la corrélation entre workload et inventaire IT est plus documentée que celle entre volume de données et inventaire IT) .

²⁹¹ <https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai>

3.3.2. Choix de modélisation

Approche surfacique pour la situation initiale en 2024

L'hétérogénéité et les lacunes des données disponibles permettant une estimation des centres de données français, en ce qui concerne les puissances électriques, les consommations d'énergie et d'eau, nous a mené à considérer cette question sous un autre angle.

Nous avons choisi une approche par surface pour l'état des lieux (la *baseline*) des centres de données identifiables, c'est-à-dire ceux pour lesquels l'hébergement IT est la fonction principale du bâtiment.

Ces centres de données comprennent à la fois :

- Les centres de données de co-location : "wholesale" et "retail"
- Les centres de données privés, opérés pour le compte d'une seule entreprise
- Les centres de données publics, opérés pour le compte d'une organisation étatique ou liée à tout autre administration publique.
- Les centres de données HPC conçus spécialement pour le calcul de haute performance (dont l'intensité énergétique par m² de surface est supérieure aux autres catégories)

La description de ces catégories se trouve dans la section [2. Les typologies de centres de données retenues pour cette étude](#), à l'exception des catégories "public" et "HPC" que nous avons ajoutées lors de la collecte, du fait de leurs particularités et de leurs volumes non négligeables.

Les données de surface sont, au sujet des centres de données, les données les plus couramment accessibles et surtout vérifiables que nous ayons identifiés, notamment car elles peuvent être confrontées à travers au moins deux sources : images satellite et données cadastrales, mais aussi avec les données fournies par l'opérateur. Le détail de la méthode de collecte se trouve dans la section dédiée à la [modélisation](#).

Complétion du périmètre considéré, à partir de l'étude ADEME/ARCEP 2022

En complément des données collectées, nous nous sommes appuyés sur les données du rapport ADEME - Arcep 2022 pour traiter les centres de données dits "traditionnels", soit les salles d'hébergement ou baies isolées présentes dans des immeubles de bureau.

Traduction des données de surface et de consommation d'électricité initiales, en inventaire IT

Nous avons adapté le modèle Masanet/Borderstep pour obtenir une modélisation de l'inventaire IT à partir des données de surface et de consommation obtenues grâce à notre inventaire 2024 et des données complémentaires.

Cette conversion des surfaces/consommations initiales en inventaires IT a facilité l'approche prospective, permettant de traiter à la fois des évolutions d'usages et des gains en efficacité.

Estimation des consommations des Blockchains, des cryptomonnaies et de l'IA générative

Nous avons estimé les consommations des Blockchains et cryptomonnaies, ainsi que de l'IA "nouvelle génération" de manière séparée. Par IA de "nouvelle génération" nous considérons les nouveaux

déploiements alimentés par les GPUs de dernière génération. Ces architectures ne sont pas prises en compte dans le modèle global issu du modèle Masanet dont la dernière version accessible date d'avant le gonflement de la vague de l'IA générative. Cela inclut de nouveaux usages ou des mises à jour d'usages existants. Ces déploiements sont parfois effectués sur des centres de données dédiés et conçus à cet effet, mais aussi sur des centres de données existants.

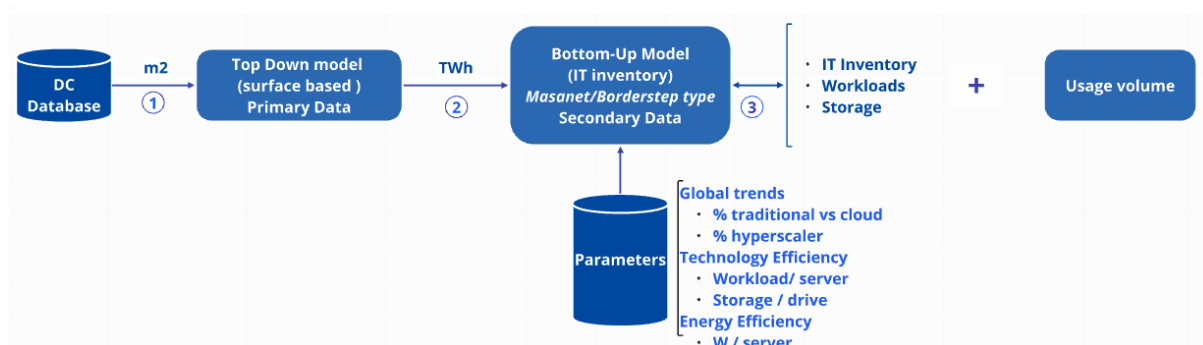
Utilisation d'un modèle orienté consommation

Utiliser un modèle "consommation". Cette approche a été décrite dans l'article d'Hubblo²⁹², mise à jour par l'ADEME en 2023²⁹³ ²⁹⁴. Un document méthodologique qui reprend cette approche est aussi en cours de finalisation au sein du comité mesure Arcep/ADEME.

3.3.3. Architecture globale

Le schéma ci-dessous illustre notre approche en deux étapes complémentaires :

- Une approche **globale** par l'inventaire de données primaires surfaciques des centres de données pour calculer une consommation globale,
- Une approche **locale** et précise du centre de données qui décrit un inventaire IT à partir d'une surface ou d'une consommation d'énergie (données secondaires). Cette approche permet l'évaluation prospective.



Le modèle poursuit deux objectifs :

- **Établir une estimation de la consommation des centres de données à date.** Celle-ci se base sur les surfaces des centres de données collectées et de facteurs de consommation / m², en fonction de leur typologie.
- **Préparer l'évaluation prospective pour les moyens et long terme.** Pour cela nous avons utilisé des modélisations disponibles (Masanet, Borderstep, ...) pour établir un inventaire IT à partir d'une consommation d'énergie électrique. Cela permet de connaître à la fois :
 - La répartition de l'énergie par élément IT

²⁹² [Impacts importés des datacenters : l'angle mort des analyses territoriales des impacts du numérique, Hubblo](#)

²⁹³ [Avis d'experts - Les data centers ou centres de données, ADEME](#)

²⁹⁴ [EVALUATION DE L'IMPACT ENVIRONNEMENTAL DU NUMERIQUE EN FRANCE, Mise à jour de l'étude ADEME-Arcep, Ademe, 2025](#)

- La matérialité derrière le kWh ou le m²
- Le lien avec les usages : workload, storage, hashrate, tflops, ... et d'estimer une évolution de cet inventaire IT en fonction de l'évolution de l'usage et de son efficacité.

De plus, cette modélisation permettrait dans le futur d'obtenir des impacts multi-critères ACV plus précis par kWh de DC, en fonction de leur nature (traditionnel, colocation, hyperscale, IA, ...)

3.3.4. Base de données 2024

Dans le cadre de cette étude, nous avons constitué une base de données des Data Centres en exploitation sur le territoire français. Des informations détaillées ont été collectées concernant **352 Data Centres actifs identifiés**.

Les données collectées proviennent de plusieurs types de sources : bases de données existantes provenant des parties prenantes de l'étude, données cadastrales et actes de ventes des bâtiments, bases de données et cartes publiques, données présentes dans les brochures commerciales des acteurs de la filière, images satellite.

Sources de données et processus de collecte

Voici la liste des sources de données sollicitées :

- La base de données des Data Centres d'Ile-de-France de l'[Institut Paris Région](#) : base de données très riche concernant les Data Centres d'Ile-de-France
- La [carte des Data Centres](#) sur le sol Français et pays frontaliers, de [DCMag.fr](#) : pour l'identification des Data Centres hors d'Ile-de-France
- La carte interactive et l'API [immobilier.pappers.fr](#) : pour la collecte des surfaces de parcelles cadastrales concernées, d'emprise au sol et de la hauteur des bâtiments, l'identification des entreprise détentrices ou utilisatrices des parcelles de terrain
- Google Maps, Open Street Map, [World Imagery Wayback](#), Yandex Maps : pour le repérage des sites, le croisement d'informations avec les données cadastrales et dans certains cas la mesure manuelle de l'emprise au sol des bâtiments.
- [OpenInfraMap](#) : carte open-data des infrastructures, à partir de données collectées collaborativement
- [Baxtel.com](#) : carte mondiale des Data Centres
- [Datacentermap.com](#) : carte mondiale des Data Centres
- [Datacenters.com](#) : base de données mondiale des Data Centres
- [DCByte.com](#) : carte mondiale des Data Centres
- Sites web et brochures des acteurs de la colocation : pour la collecte des puissances installées, des informations relatives à la redondance électrique, aux surfaces ITs, etc.

Le processus générique de collecte et de complétion des données concernant un centre de données est le suivant :

1. Identification d'un site supposé comprendre un centre de données, à partir de l'une des bases de données citées plus haut, ou bien par obtention de l'information par l'un des membres du consortium, ou encore par repérage à partir des outils d'imagerie satellite.



Illustration de l'identification d'un centre de données et de l'extraction de son emprise au sol.

2. Confirmation que le site en question correspond, soit à une parcelle cadastrale qui appartient à l'entreprise opératrice du centre de données ou à l'une de ses filiales, soit que ladite entreprise fait partie des occupants de la parcelle.
3. Le cas échéant, obtention des données à partir de l'API, ou du site web immobilier.pappers.fr : surface de la parcelle (m²), emprise au sol du ou des bâtiments (m²), hauteur du ou des bâtiments (m).
4. Calcul du nombre d'étages du centre de données, à partir d'une règle statistique : au moins 6 m par étage.²⁹⁵
5. Calcul de la surface de plancher (m²) à partir du nombre d'étages et de l'emprise au sol du bâtiment (m²).
6. Si la surface IT est publiquement accessible sur le site ou la brochure de l'opérateur, utilisation de cette valeur. Sinon, calcul de la surface IT (m²), à partir de la surface de plancher et de la règle statistique suivante²⁹⁶ :
$$\text{Surface_IT} = \text{Surface_De_Plancher} / 2$$
7. Obtention des données de puissance électrique totale installée à l'échelle du bâtiment (IT+non-IT) et de redondance électrique sur le site ou la brochure de l'opérateur. Si ce n'est pas disponible, calcul de la puissance installée à partir de la surface de plancher et des ratios de puissance / surface obtenus depuis les autres entrées de la base de données, dans la même catégorie.
8. Obtention de la date de mise en service à partir d'articles de presse ou des informations fournies par l'opérateur.

²⁹⁵ Fondé sur des entretiens avec les opérateurs.

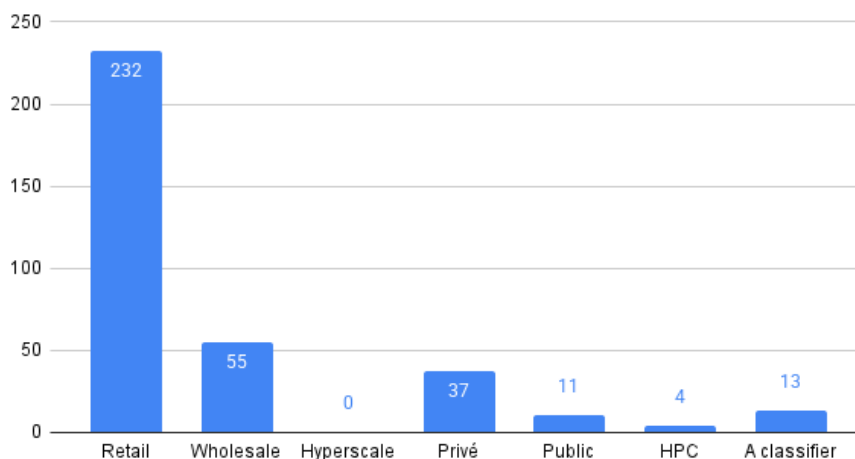
²⁹⁶ Cette règle statistique est établie à partir d'un échantillon de données primaires et d'interviews d'acteurs.

Résultats de la collecte

En nombre de centres de données en exploitation

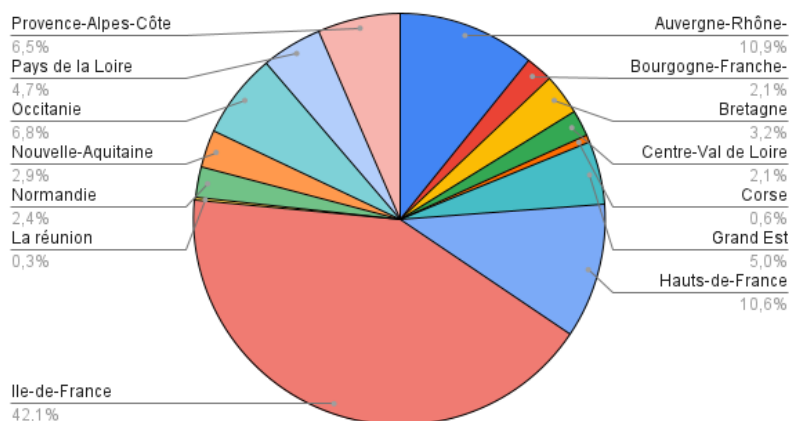
Pour l'année 2024, les résultats de notre travail d'inventaire des centres de données en France a permis de recenser **352 centres de données actifs** (indiqués par la suite "en exploitation").

Nombre de Data Centres en exploitation en 2024



En nombre de centres de données en exploitation, la catégorie largement majoritaire est celle de la colocation de taille modeste, Retail, avec 232 centres de données identifiés, puis Wholesale avec 55 centres de données (colocation jusqu'à 40 MW). Les centres de données identifiés de la catégorie Privé (opérés pour le compte d'une seule entreprise) sont au nombre de 37, ceux des catégories Public et HPC respectivement de 11 et 4. 13 centres de données n'ont pas été catégorisés au moment de l'écriture du rapport par manque de données pour le permettre.

Répartition des en # de datacenters actifs par région

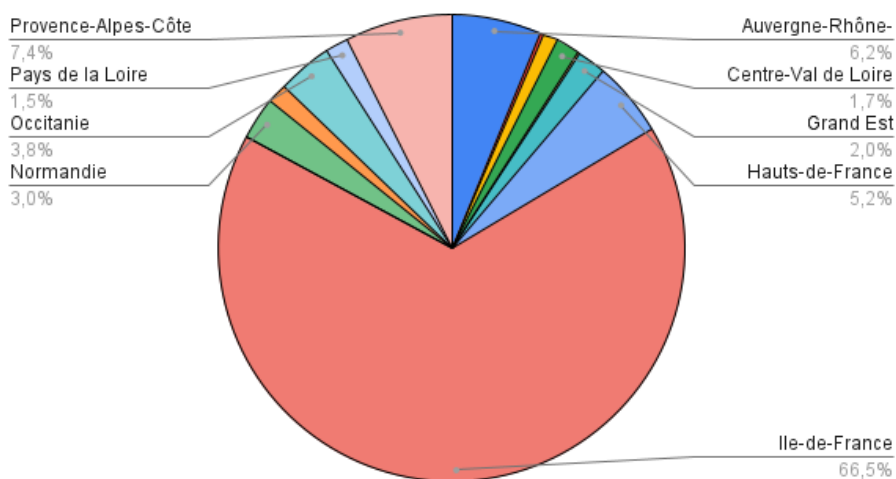


Ventilés par régions, ces centres de données actifs sont en présence dominante en Ile-de-France, mais non majoritaires (42,1 %). Les régions Auvergne-Rhône-Alpes et Hauts-de-France sont en deuxième et troisième position avec respectivement 10,9 et 10,6 %, suivis des régions Occitanie et PACA, avec 6,8 et 6,5 %.

En puissance électrique installée (puissance totale cumulant puissance IT et non-IT)

La puissance installée totale cumulée des centres de données identifiés comme actifs, dans la base de données de cette étude, est de **2,46 GW**.

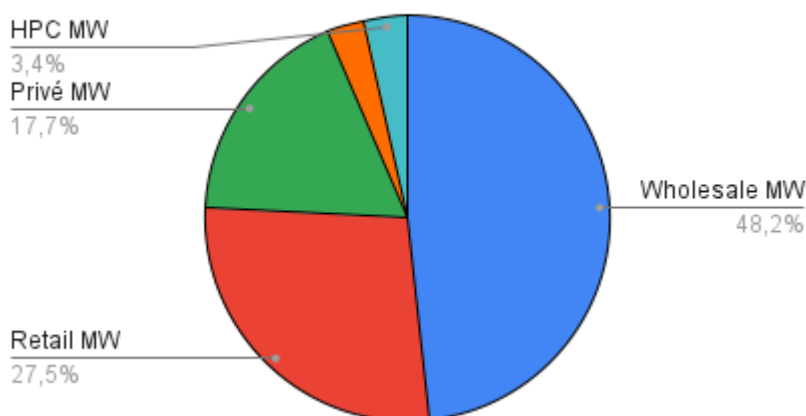
Répartition en puissance installée par région



Sous cet angle la répartition régionale est légèrement différente. Les puissances cumulées font de l'Ile-de-France la région majoritaire à l'échelle du pays avec 66,5% de la puissance. La région PACA est en deuxième position avec 7,4 %, Auvergne-Rhône-Alpes troisième avec 6,2%, suivie des Hauts-de-France avec 5,2 %.

Ce poids de l'Ile-de-France dans ce découpage s'explique notamment par la taille et la capacité des centres de données qui y sont installés.

Répartition des puissances installées

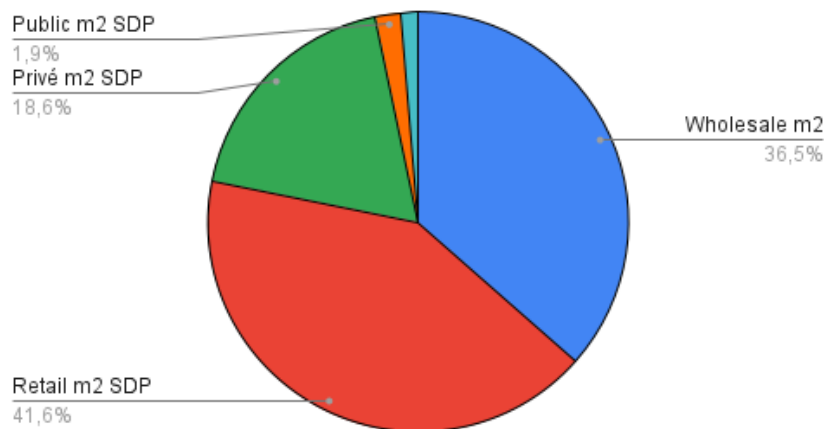


Si l'on ventile cette puissance installée par catégories, ce sont les centres de données Wholesale qui représentent la plus grande part de la puissance installée des centres de données en exploitation, avec 48,2 %, suivie des centres de données Retail, avec 27,5% et Privé avec 17,7%.

En surface

La surface de plancher désigne la surface au sol cumulée des différents étages du centre de données. La **surface de plancher** des centres de données français, en exploitation, enregistrée dans la base de données de la présente étude est de **2 445 642 m²**.

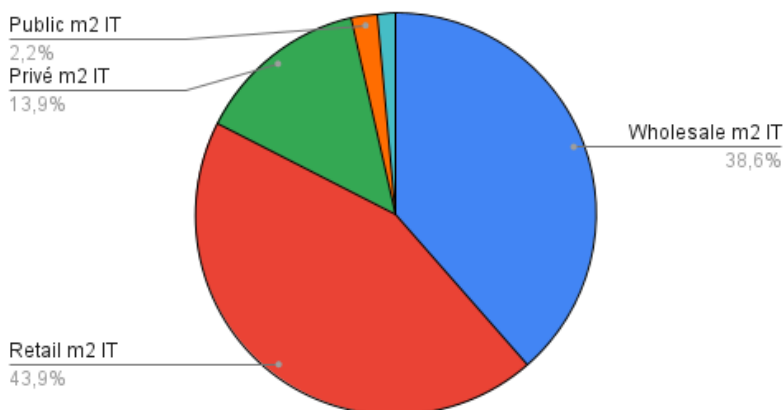
Répartition des surfaces de plancher



Sous cet angle, les centres de données Retail représentent la plus grande part avec 41,6%, ce qui s'explique par leur présence en bien plus grand nombre que toutes les autres catégories (232). Les centres de données Wholesale représentent 36,5 % de la surface de plancher cumulée des centres de données français.

Cette répartition est très proche de celle de la **surface IT** (surface au sol dédiée aux salles d'hébergement des équipements informatiques), dont le total s'élève à **1 125 403 m²**.

Répartition des surfaces IT

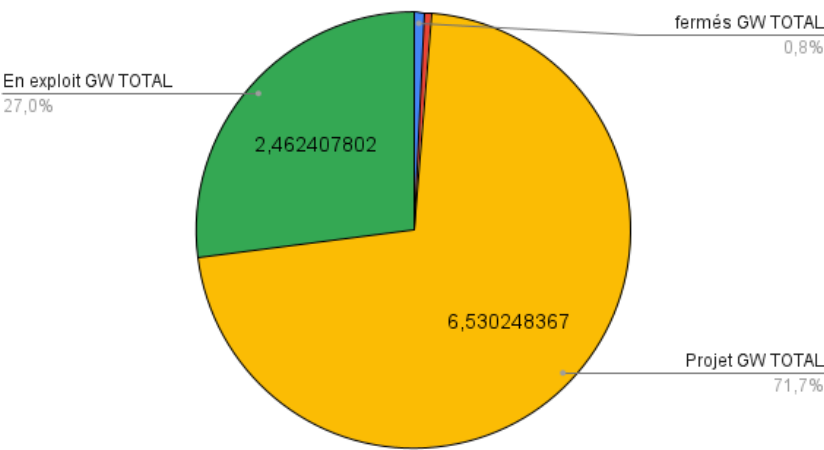


Note sur les projets de futurs centres de données

Un travail **partiel** de comptabilisation des projets de nouveaux centres de données a été effectué lors de cette étude. Au moment de rédaction de ce rapport, les annonces des projets de centres de données en provenance de la presse et des opérateurs de centres de données eux-mêmes, cumulaient 6,5 GW de puissance électrique. Ce total ne fait pas état de la probabilité que ces projets voient effectivement le jour, ni de la progressivité de l'installation de ces puissances. Les puissances annoncées sont souvent des projections à terme, qui sont en fait installées par tranche avec parfois plusieurs années d'écart entre

deux étapes de déploiement. Ce cumul comprend également des projets dont la date de mise en service annoncée va jusqu'à 2035.

Répartition des puissances installées par état d'avancement



Le rapport entre les projets annoncés (6,5 GW de puissance cumulée d'après cet inventaire) et la puissance installée inventoriée en 2024 (2,46 GW) est très important, puisqu'il représente une multiplication par 3,65 de la puissance installée totale du parc de centres de données français avec un total de 8,99 GW, si l'ensemble de ces projets venaient à se réaliser dans les conditions annoncées et à pleine capacité.

Cependant, la volatilité et la non-fiabilité des annonces nous a confortés à **ne pas utiliser ces données de puissances installées** des projets de centres de données **pour nos projections**. La méthode employée dans le modèle prospectif de cette étude est détaillée dans la section [Modélisation bottom-up](#).

Résultats des estimations des consommations pour l'année 2024

Le calcul des estimations des consommations d'énergie effectué pour l'année 2024 est le suivant. Nous avons rassemblé pour chaque catégorie comprenant des centres de données en exploitation en 2024 (Wholesale, Retail, Public, Privé, HPC) :

- Le cumul de la surface IT de la catégorie, en m², présent dans notre base de données
- Le ratio de consommation d'énergie par m² de surface IT, fournis par le CEREN, issus de la base de données européenne en cours de remplissage dans le cadre de l'Energy Efficiency Directive
- Le PUE (Power Usage Effectiveness) Moyen de la catégorie

Wholesale m2 IT	Retail m2 IT	Hyperscale m2 IT
434214,4074	493639,6983	0
Privé m2 IT	Public m2 IT	HPC m2 IT
156812,7225	24911	15825,5

Cumul de la surface en m² IT, en exploitation en 2024, collectée par catégorie de centre de données

ratio consommation IT/surface IT - MWh/m2	Catégorie de surface IT dans la base européenne	Catégorie concernée dans notre calcul
--	--	--

5,21	100 à 500 m2	-
4,43	500 à 1 000 m2	-
5	1 000 à 5 000 m2	Retail
6,09	5 000 à 10 000 m2	Wholesale, Privé, Public
7,3	Plus de 10 000 m2	HPC

Ratios de consommation d'énergie électrique (IT) par m² de surface IT issus de la base de données européenne, au moment de cette étude, fournis par le CEREN

Catégorie concernée	PUE moyen appliqué
Retail	1,5
Wholesale	1,4
Privé	1,4
Public	1,4
HPC	1,3

PUE appliqués par catégorie

Nous obtenons alors une estimation intermédiaire à partir du produit de ces trois jeux de données, à laquelle nous appliquons un nouveau facteur modélisant la part de la surface IT, qui est effectivement déployée en centre de données. Admettons qu'un centre de données puisse théoriquement déployer 1000 m² de salle IT. Il n'a pas nécessairement, à date, mis en place le nécessaire pour que ces 1000 m² accueillent effectivement du matériel informatique. Nous modélisons ce taux de remplissage de la surface IT avec un ratio, calculé en comparant les surfaces IT comptabilisées sur un échantillon de centres de données présents à la fois dans la base de données européenne et dans la base de données de cette étude. Ces ratios n'ont pu être calculés que pour deux catégories et sont respectivement de 0.973 pour la catégorie Wholesale et de 0.773 pour la catégorie Retail.

Nous obtenons ainsi les estimations suivantes :

Consommation d'énergie par catégorie, en 2024, en TWh	Avant modélisation du remplissage de surface IT	Après modélisation du remplissage de surface IT
Wholesale	3,702112037	3,602155012
Retail	3,702297737	2,861876151
Hyperscale	0	0
Privé	1,336985272	1,336985272
Public	0,212391186	0,212391186
HPC	0,150183995	0,150183995
Total	9,103970227	8,163591616

L'estimation totale pour l'ensemble des catégories de cette collecte s'élève à **8,16 TWh**. Dans ce décompte les centres de données de **colocation** (Wholesale et Retail) pèsent pour **6,46 TWh**, les centres de données **privés** pour **1,34 TWh**, les centres de données du secteur **public** pour **0,21 TWh** et les centres de données spécialisés dans le calcul haute performance (**HPC**) pour **0,15 TWh**.

Le périmètre concerné par cette collecte exclut les salles informatiques présentes dans les bâtiments tertiaires. Cette partie est prise en compte en s'appuyant sur la méthode utilisée dans l'étude ADEME/Arcep parue en 2022, prenant en compte les besoins en surface d'hébergement IT des entreprises. Ce faisant, en appliquant les hypothèses du modèle Masanet, c'est-à-dire, une baisse annuelle de 5% pour le volume de workloads opérés par les centres de données "traditionnels", nous prenons en compte une consommation électrique annuelle de 3,5 TWh. Cela inclut des centres de données privés, des centres de données publics et des baies informatiques dans les bâtiments tertiaires. Le cumul de consommation d'électricité sur l'ensemble du périmètre, prenant en compte les centres de données de colocation issus de la collecte et l'extrapolation surfacique issue du rapport ADEME/Arcep, donne une consommation de 10 TWh en 2024, qui sert de point de départ au modèle.

3.3.5. Modélisation bottom-up

En date de réalisation de notre étude, le seul modèle totalement disponible est celui de Masanet²⁹⁷. Nous avons donc basé nos travaux sur celui-ci, notamment pour identifier les inventaires IT équivalents à 1 kWh de consommation IT pour tous les types de centres de données.

Notre année de référence est 2024. Aux inventaires IT nous avons associé des volumes d'usage (workload et stockage) et des caractéristiques d'efficacité.

Nous avons ensuite procédé aux ajustements documentés ci-dessous.

Notre modèle étant ouvert, il sera possible de l'amender à partir d'autres modèles académiques ou suite aux retours de la communauté.

Répartition de la consommation par composant IT

Le modèle de Masanet et al considère les catégories de composant IT suivantes :

- "Volume server": les serveurs d'entrée de gamme
- "Mid-range server": les serveurs de milieu de gamme
- "High-end server": les serveurs hauts de gamme
- "SSD drive": les disques durs de type Solid State Drive
- "HDD drive": les disques durs de type mécanique / Hard Disk Drive
- "Port": le port réseau moyen présent sur un serveur

Le modèle permet d'obtenir une répartition de la consommation électrique par type de composant IT et par type de centres de données :

% energy use by data center type in 2023	Volume Server	Mi-Ranger	High-End	SSD	HDD	Network
Traditional	37,25%	23,11%	27,08%	4,73%	6,07%	1,76%
Cloud (non-hyperscale)	61,96%	9,43%	11,05%	6,58%	8,43%	2,55%
Hyperscale	80,99%	0,00%	0,00%	6,68%	8,56%	3,76%

En incluant les données d'usages (workload, volume de stockage) nous obtenons la répartition par composant IT suivante pour 1 kWh d'IT :

²⁹⁷ <https://github.com/emasanet/GlobalIDCAAnalysis>

Compute :

components for 1kWh of IT	#Workloads of Volume Servers	#Volume Servers	#Workloads of Mid-Range servers	#Mid-Range servers	#Workloads of High-End Servers	#High-End Servers
Traditional	1,4E-03	3,1E-04	5,1E-05	1,1E-05	5,7E-06	1,3E-06
Cloud (non-hyperscale)	5,7E-03	3,7E-04	7,1E-05	4,6E-06	8,0E-06	5,1E-07
Hyperscale	5,9E-03	3,8E-04	0,0E+00	0,0E+00	0,0E+00	0,0E+00

Stockage et réseau :

components for 1kWh of IT	Vol SSD (MB)	# disques SSD	Vol HDD (MB)	# disques HDD	# ports
Traditional	6,0E+03	6,2E-04	1,4E+04	7,0E-04	1,6E-03
Cloud (non-hyperscale)	8,4E+03	8,6E-04	1,9E+04	9,8E-04	1,9E-03
Hyperscale	8,5E+03	8,7E-04	1,9E+04	9,9E-04	1,9E-03

Cette répartition IT par kWh va nous permettre de transformer la consommation électrique de référence en inventaire IT et en volume d'usage pour ensuite faciliter l'exercice prospectif.

Mise à jour des paramètres d'efficacité

Le modèle initial de Masanet et al. inclut des hypothèses d'évolution de l'efficacité énergétique pour chaque catégorie de matériel IT. Concrètement, il s'agit d'hypothèses chiffrées et exprimées en pourcentage, de l'évolution annuelle de la puissance moyenne d'un serveur informatique pour une catégorie marché.

Les hypothèses initiales avançaient 1% de réduction de la puissance moyenne par an pour les "Volume server", 8% d'augmentation annuelle de la puissance moyenne pour les "Mid-range server" et 7% d'augmentation par an de la puissance moyenne des "High-end server". Ce type d'hypothèse est implicitement limité par le fait que la puissance moyenne du matériel en centre de données ne dépend pas uniquement de la puissance moyenne du matériel proposé sur le marché la même année. La capacité à substituer réellement un équipement par un équipement plus récent varie d'une organisation à une autre et dépend fortement de dynamiques économiques, sociales et organisationnelles.

Malgré ces limites, nous avons tenté *a minima* de nous rapprocher le plus possible de l'évolution des puissances moyennes des serveurs, constatée sur le marché. Pour cela, nous nous sommes appuyés sur la base de données SpecPOWER qui fournit des puissances minimum, maximum et moyennes pour plus de 1000 références de serveurs dont les dates de sortie sur le marché vont des années 2000 à l'année 2024.

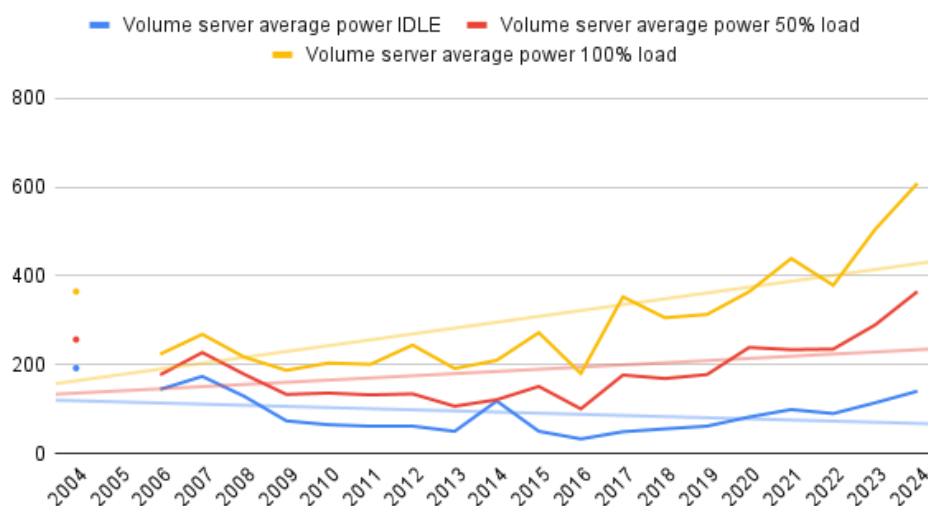
Nous avons extrait²⁹⁸ de cette base les puissances moyennes pour trois catégories de serveur, en fonction de leur date de parution. Les trois catégories sont :

- "Volume server": moins de deux puces de calcul par serveur (CPU et GPU compris)
- "Mid-range server": de deux à 8 puces de calcul par serveur
- "High-end server": plus de 8 puces de calcul par serveur

²⁹⁸ [Fichier de travail des données issues des bases SpecPOWER, The CHIP Dataset, TechPowerUp et BoaviztAPI](#)

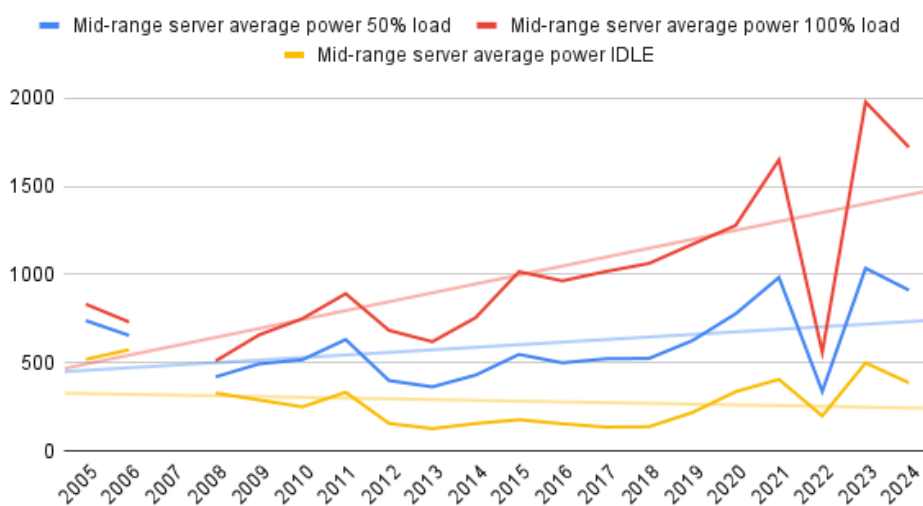
Pour chacune des catégories il est alors possible de tracer l'évolution de la puissance moyenne, pour chaque état d'utilisation du serveur, nous en avons sélectionné trois pour observation : taux de charge à 0% (IDLE), 50%, 100%.

Volume server avg power (W)



Puissance moyenne sur l'ensemble des serveurs de la catégorie "Volume server", de génération en génération, en fonction de leur taux de charge : 0 (IDLE), 50 et 100%

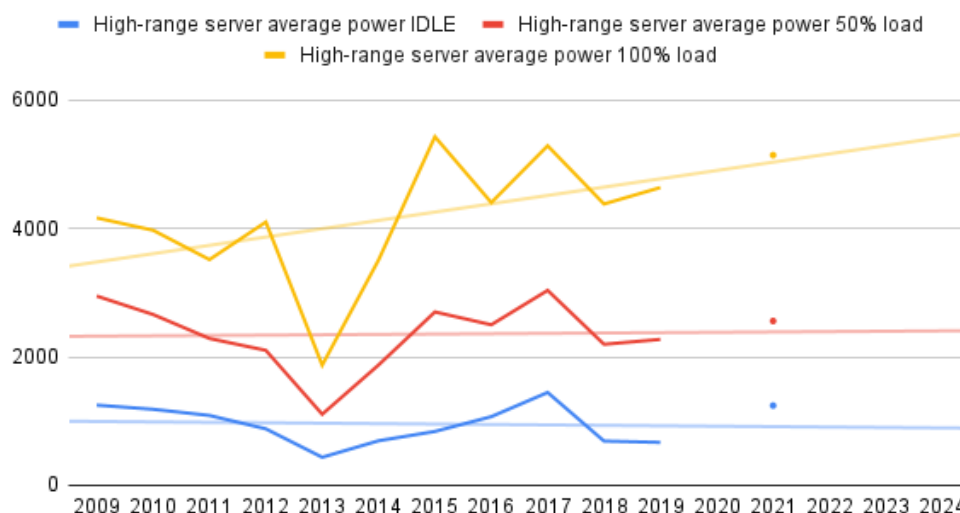
Mid-range server avg power (W)



Puissance moyenne sur l'ensemble des serveurs de la catégorie "Mid-range server", de génération en génération, en fonction de leur taux de charge : 0 (IDLE), 50 et 100% (le creux de 2022 correspond à un défaut de reporting)

Puissance moyenne sur l'ensemble des serveurs de la catégorie "High-end server", de génération en génération, en fonction de leur taux de charge : 0 (IDLE), 50 et 100%

High-range server avg power(W)



Nous avons retenu l'hypothèse de **50% de taux de charge** comme base des nouvelles hypothèses d'efficacité pour le modèle, soit **1,77% d'augmentation par an de la puissance moyenne** des serveurs de type **"Volume server"**, **1,11% d'augmentation par an** de la puissance moyenne des serveurs de type **"Mid-range server"** et **1,17% par an de baisse** de la puissance moyenne des serveurs de type **"High-end server"**. Nous avons appliqué ces hypothèses en lieu et place de celles prévues initialement dans le modèle Masanet ²⁹⁹.

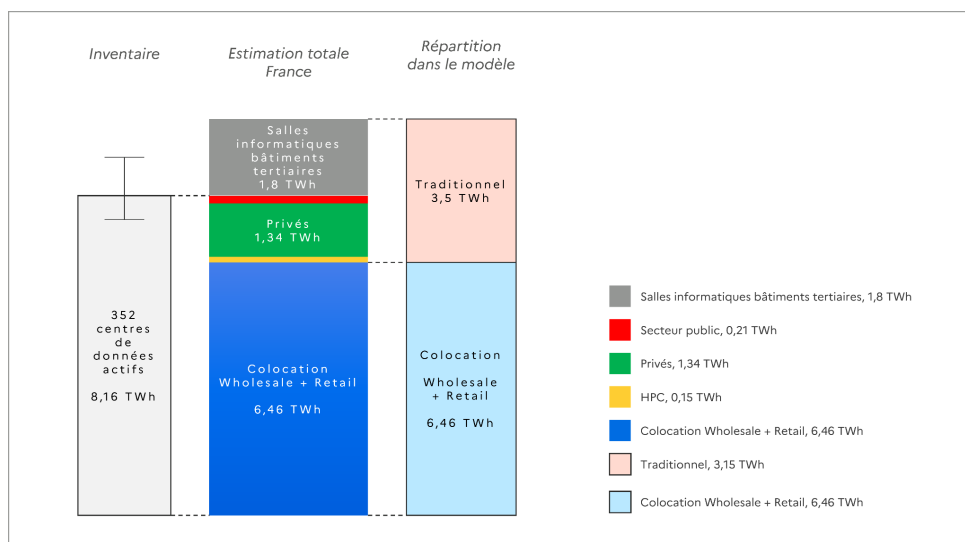
3.4. Approche méthodologique détaillée

3.4.1. Modélisation de la situation initiale

Les données d'entrée du modèle, concernant la consommation d'énergie des centres de données de colocation (compris dans la catégorie "cloud", dans le modèle Masanet/Borderstep), proviennent de la collecte détaillée dans la section [Base de données 2024](#).

Le schéma ci-dessous illustre les données qui ont servi de point d'entrée au modèle, avec d'une part les données issues de la collecte pour la colocation, d'autre part les données obtenues en s'appuyant sur l'étude ADEME/ARCEP 2022 pour les centres de données privés, publics et les salles informatiques et baies de bureau (regroupés dans la catégorie "centres de données traditionnels", dans le modèle).

²⁹⁹ <https://github.com/emasanet/GlobalIDCAalysis>



Représentation schématisée de la répartition des centres de données par typologie dans l'établissement de la baseline du modèle : à gauche, l'inventaire réalisé sur la base de données primaires ; au centre, la répartition sur la base de l'inventaire et en complément le différentiel correspondant aux baies de bureau en s'appuyant sur l'étude ADEME/ARCEP 2022 ; à droite, la répartition conservée dans le modèle CLIK.

Les centres de données cloud "importés"

Comme décrit précédemment, notre estimation est basée sur une approche "par consommation" :

- La prise en compte des "usages en France" inclut les besoins traités sur le "territoire France" et les besoins dont le traitement est effectué hors du "territoire France" (métropolitaine). On parle alors de services importés.
- Ce ratio "interne/importé" sera pris en compte dans les scénarios prospectifs en fonction des évolutions du "move to cloud" et du déploiement des centres de données en France (colocation et hyperscale)
- Faute d'études sur le sujet, les traitements effectués sur le "territoire France" à destination de besoins à l'étranger (services exportés) ne sont pas pris en compte (leur intégration serait compensée par une augmentation équivalente des services importés).

En 2024, la part des workloads traités par des centres de données traditionnels est estimée à 5%. Cela aboutit à la consommation totale des centres de données "cloud" de 12,2 TWh.

Comme détaillé dans l'article d'Hubblo sur la prise en compte des services "importés" ³⁰⁰, la consommation des centres de données "importés" en 2024 est la suivante :

$$\text{Cloud Total} - \text{Cloud France} = 5,8 \text{ TWh}$$

"IA nouvelle génération"

Comme expliqué [au chapitre 2](#), de nombreuses études et publications récentes se sont penchées sur le développement de l'IA générative, ainsi que sur la consommation environnementale des centres de données dédiés et de certains composants clefs, tels que les GPUs ou les serveurs dédiés.

³⁰⁰ <https://hubblo.org/fr/articles/20230918-datacenters-imported-impacts/>

Nous avons fait le choix de considérer séparément l'IA de "nouvelle génération". Cela permet de mettre son évolution en exergue dans l'exercice prospectif, sans qu'il nous soit toutefois possible d'établir cette distinction pour les usages passés dans la baseline.

Le développement de l'IA de nouvelle génération inclut les nouveaux usages permis par l'IA générative (agents conversationnels, assistants etc ...) mais aussi la mise à jour d'usages existants (moteurs de recherche, ...).

Ce secteur en pleine expansion se caractérise par une opacité des données encore plus marquée que celle du numérique "classique". Certaines recherches sont basées sur l'estimation de la consommation électrique moyenne de l'inférence (Schneider, De Vries, ...), tandis que d'autres ont adopté une approche plus traditionnelle fondée sur des estimations d'inventaire. C'est le cas de l'étude de Berkeley Lab³⁰¹, qui nous semble à ce jour la plus fiable.

Cette étude permet d'estimer la consommation des centres de données spécifiques à l'IA générative³⁰² sur le territoire des Etats Unis. Elle représente environ 40% de la consommation totale des centres de données sur le territoire.

En supposant que les États-Unis sont globalement exportateurs de services d'IA générative, on considère un ratio de 20% de consommation totale des centres de données du territoire français alloué à l'IA.³⁰³

Cela conduit, **pour l'année 2024, à une estimation de 3,3 TWh consommés par les centres de données dédiés à l'IA générative pour un usage français.**

La blockchain

La nature distribuée des technologies Blockchain rend l'estimation de sa consommation énergétique très difficile.

Après une étude de l'état de l'art, détaillée au Chapitre 2, nous avons retenu la méthodologie documentée par le Cambridge Bitcoin Electricity Consumption Index (CBECI) de l'Université de Cambridge.³⁰⁴

Leur approche distingue deux types de Blockchain (appliquée à la cryptomonnaie)

- Le Proof of Work sur la base de l'exemple de Bitcoin
- Le Proof Of Stake sur la base de l'exemple d'Ethereum

Le suivi de l'estimation des consommations est fourni par le site dédié :

³⁰¹ <https://escholarship.org/uc/item/32d6m0d1>

³⁰² comprenant entraînement et inférences

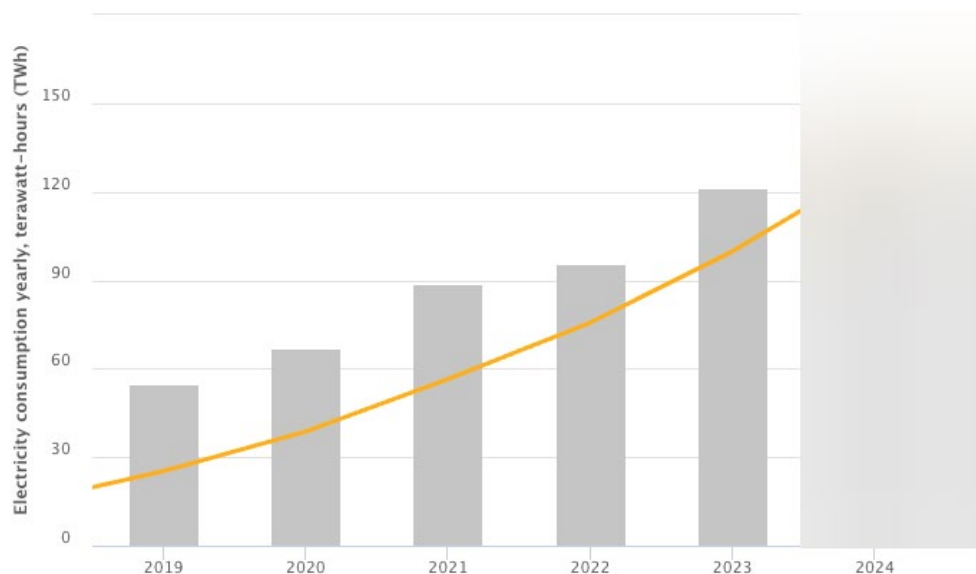
³⁰³ hypothèse CLIK

³⁰⁴ <https://ccaf.io/cbnsi/cbeci/ghg/methodology>

Total Bitcoin electricity consumption

Select a time window by clicking and dragging to define the start and end dates on the timeline below

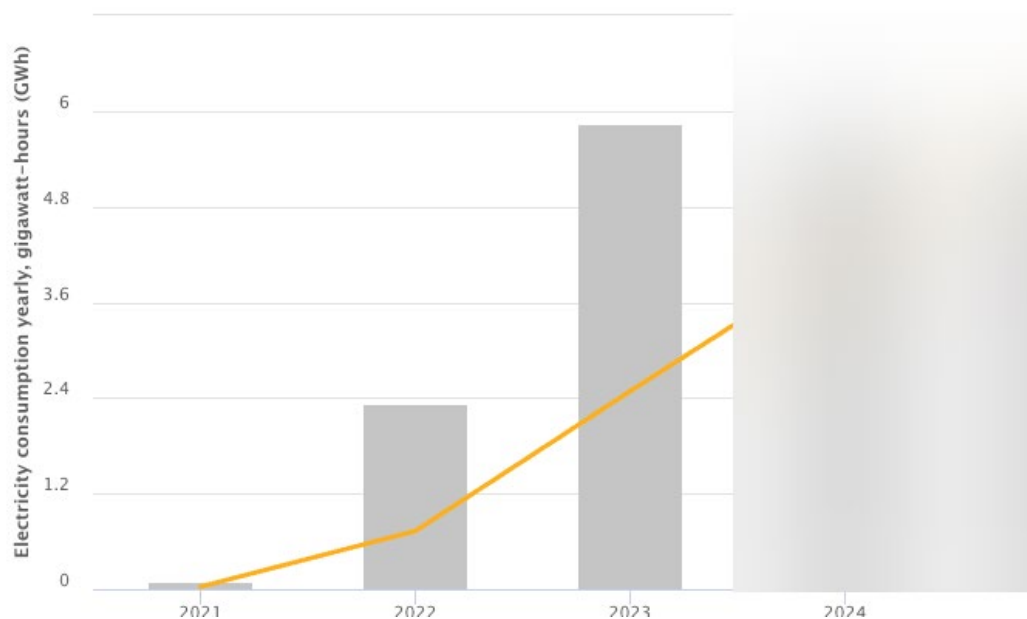
Yearly Monthly



Total Ethereum electricity consumption

Select a time window by clicking and dragging to define the start and end dates on the timeline below

Yearly Monthly



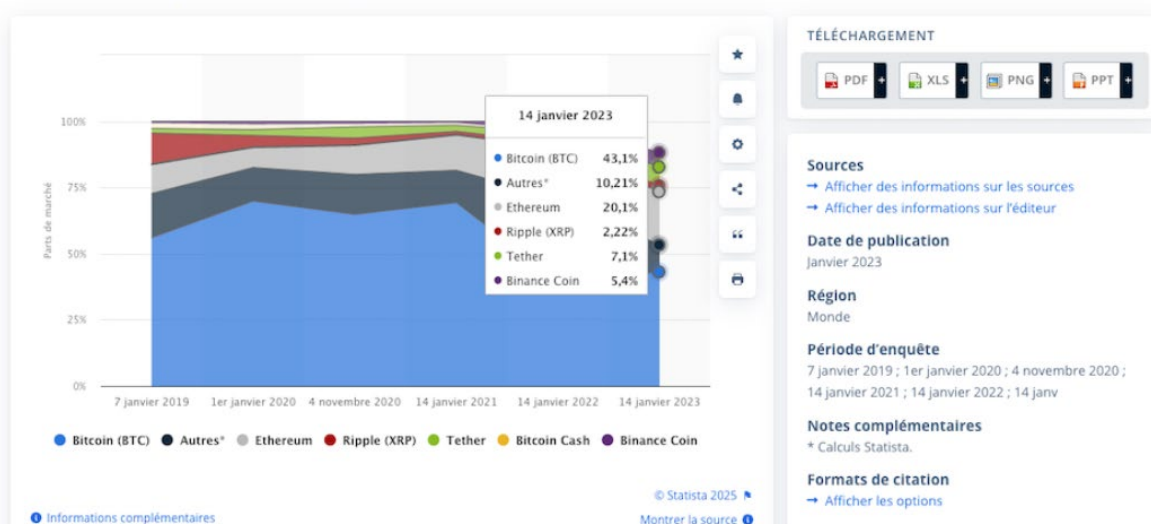
L'outil fournit également le suivi des efficacités (en kWh / kHashs) pour le PoW ainsi qu'une comparaison des efficacités entre PoW et PoS.

Cela nous permet d'identifier les consommations ainsi que les volumes d'usage en équivalent Hashs (pour le PoS) pour le Bitcoin et Ethereum :

	Bitcoin (PoW)	Ethereum (PoS)
Conso 2023 (TWh)	121	5,85E-03
Efficacité 2023 (kWh/kH)	8611,111111	
Progression annuelle efficacité	15,48%	
#THeq (2023)	1,41E-02	3,40E-03
Efficacité PoS/PoW (%)		0,02%

Nous avons ensuite utilisé des statistiques des parts de marché de Bitcoin et d'Ethereum³⁰⁵ :

Répartition des parts de la capitalisation boursière des principales crypto-monnaies dans le monde de janvier 2019 à janvier 2023



Nous utilisons le PIB³⁰⁶ comme indicateur de référence pour ramener les estimations mondiales à celles de la France :

Total Crypto 2024	World	France
Conso PoW	192,1	5,5
Conso PoS	9,29E-03	2,67E-04
Volume Heq (PoW)	2,23E+07	6,41E+05
Volume Heq (PoS)	5,39E+06	1,55E+05
Efficacité PoW (kWh/kH)	8611,1	8611,1
Gain annuelle efficacité PoW	0,2	0,2
Efficacité PoS (kWh/kH)	1,7	1,7

³⁰⁵ Source : Statista <https://fr.statista.com/statistiques/803748/parts-capitalisation-boursiere-principales-crypto-monnaies/>

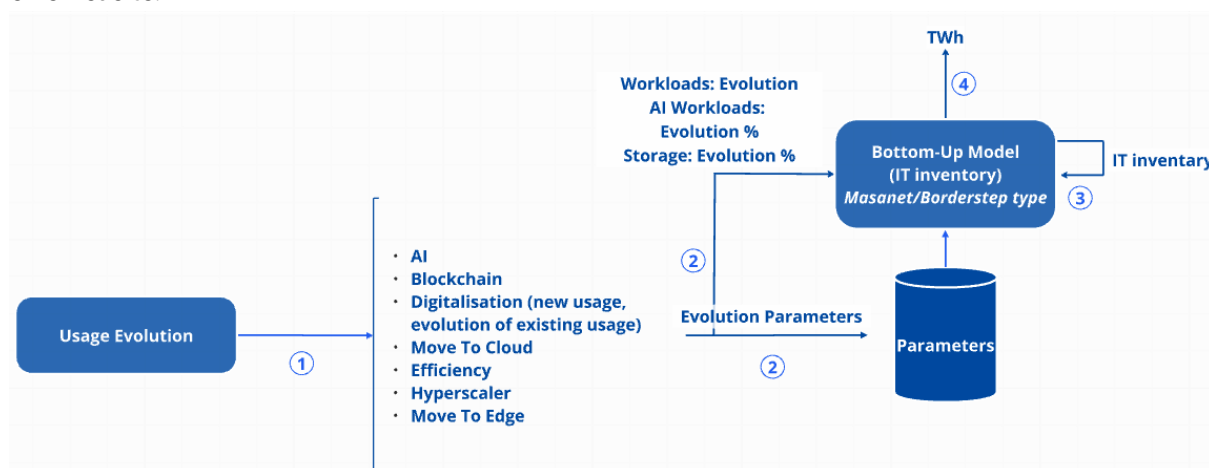
³⁰⁶ La carte fournie par CBECI fournissant une répartition géographique de la production (du minage) et non de la consommation, nous utilisons le PIB comme proxy. Par ailleurs, l'approche production montre bien que la part française est anecdotique.

Ainsi, la consommation estimée de la Blockchain pour un usage français en 2024 est de 5,5 TWh. Cette consommation estimée liée à l'usage français de la blockchain, est par hypothèse considérée comme 100% importée, la consommation de la blockchain sur le territoire français étant considérée comme anecdotique par les acteurs du secteur.

3.4.2. Court, moyen, long terme : définition de séquences d'évolution

Approche globale

L'approche présentée dans le chapitre précédent établit une base pour l'année 2024, à partir de laquelle nous pouvons modéliser différentes trajectoires en faisant principalement varier, d'une part, l'évolution des usages du numérique et des tendances technologiques ("move to cloud", ...), et d'autre part, les gains en efficacité.



A noter que nous avons fait le choix de permettre des évolutions non linéaires, contrairement aux principaux modèles académiques disponibles (Masanet, Andrae, etc.).³⁰⁷

Dans un premier temps, nous allons décrire les tendances initiales qui s'appliquent à tous les scénarios pour les premières années. Celles-ci sont basées sur l'historique proche, que ce soit pour les usages (workloads etc.), les tendances globales (Move To cloud) ou les gains en efficacité (IT classique, GPU, Hash rate etc.).

Les évolutions moyen terme (jusqu'à 2035) seront ensuite définies en fonction des scénarios prospectifs considérés.

³⁰⁷ Ce point est important pour limiter le biais de linéarité. Ainsi, les évolutions des usages peuvent connaître des pics, des creux, ou tout autre effet, les gains d'efficacité peuvent subir une accélération ou bien un plafond ; l'évolution des workloads en France ou au global peut varier en fonction des effets produits par diverses mesures politiques et/ou économiques, et l'atteinte de seuils d'usages. Ainsi, l'ensemble de nos scénarios modélisant une croissance des usages tiennent compte d'une forme d'inflexion de cette croissance des usages au fur et à mesure des années, pour limiter les effets d'une croissance en pourcentage de l'année précédente. Par exemple, pour le scénario tendanciel, alors que dans l'hypothèse faite dans le modèle Masanet, les workloads augmenteraient de 20% chaque année, nous avons ici fait l'hypothèse que les usages des workloads des centres de données "historiques", croîtraient en 2025 et 2026 de +22% par an, puis subiraient une légère inflexion à +20% de croissance en 2027 et 2028, puis à +18% entre 2029 et 2034, puis à partir de 2035 (avec taux de croissance des workloads à +15% par an), cette inflexion se poursuivrait régulièrement pour se stabiliser à 7% par an en 2048 et les années suivantes jusqu'à 2060.

Pour les évolutions plus long terme (2035-2060), compte tenu du haut niveau d'incertitudes sur l'évolution des technologies IT, nous avons choisi un niveau de modélisation plus macro, en considérant un seul niveau d'efficacité par type de centre de données en lieu et place des efficacités par composant IT (types de serveurs, de disques etc ...).

Modèle bottom-up 2024-2035

Nous détaillons dans cette section les principaux paramètres constitutifs du modèle CLIK, qui s'est inspiré de la structure du modèle Masanet et a réajusté ses paramètres à la période actuelle et les particularités du périmètre France.

Centres de données "historiques"³⁰⁸

Évolution des usages du numérique et des tendances technologiques

La croissance annuelle constatée du volume de workloads dans le modèle Masanet³⁰⁹ sur les 5 dernières années est de 20%. De même, la croissance annuelle du volume de stockage est de 33%.

Concernant, l'évolution des centres de données dits traditionnels³¹⁰, on observe une baisse annuelle de 5% du volume de workloads.

On en déduit la croissance annuelle des workloads opérés par les hyperscales et les centres de données de colocation.

Ces hypothèses du modèle Masanet ont servi de base à notre réflexion et la constitution des paramètres du modèle CLIK.

Évolution sur le territoire français

Notre modèle intègre en input, les données 2024 mentionnées précédemment [ici](#).

Nous distinguons 3 lignes concernant l'évolution des consommations des centres de données sur le territoire français, au sein du modèle :

1. Une première ligne concerne l'évolution globale des Workloads par année. Cette évolution étant globale, elle concerne l'ensemble des centres de données qui répondent aux **usages** des Français, que ces usages soient alimentés par des centres de données présents sur le sol français ou à l'étranger. Cette évolution permet de donner la tendance globale des évolutions des usages.
2. Une seconde ligne permet de déterminer l'évolution des workloads des centres de données de type colocation en France, sur la base des données 2024 en input et au regard de la tendance globale mentionnée au point précédent. En d'autres termes, cette ligne permet de modéliser (en pourcentage d'évolution année après année sur la base des données 2024) l'évolution de la consommation des centres de données de type colocation en France, année après année, en tenant compte de la demande (usages).
3. Une troisième ligne permet de déterminer l'évolution des workloads des centres de données de type hyperscale, en France, sur la base des données 2024 en input. Cette ligne fonctionne de la même manière que la ligne précédente, en étant alimentée par les données 2024 de l'implantation des centres de données de type hyperscale en France. Spécificité liée au cas français : ce type de centre de données n'étant pas encore implanté et actif en 2024, nous avons choisi une donnée de base

³⁰⁸ Préexistants au modèle CLIK

³⁰⁹ <https://github.com/emasanet/GlobalDCAnalysis>

³¹⁰ Pour mémoire, cette catégorie dans le modèle CLIK inclut les centres de données privés, les centres de données publics, les baies de bureau et les HPC.

proche du 0 mais non nulle, sans quoi aucune hypothèse d'évolution de l'implantation de ces centres de données ne pourrait être utilisée.

Note : Le modèle CLIK ne prend pas en compte l'exportation des services. Il est limité par la demande d'usage : si le nombre de centres de données en fonctionnement en France dépasse les besoins réels des utilisateurs français, la consommation excédentaire liée à cette capacité supplémentaire n'est pas intégrée dans le modèle.

3.4.3. Modélisation du potentiel de récupération de chaleur fatale

La modélisation s'appuie sur les résultats produits par le modèle, concernant les consommations d'électricité des centres de données dans les différents scénarios, jusqu'en 2035.

Nous nous appuyons sur la répartition projetée des technologies de refroidissement sur le marché des centres de données français, en partant des données 2024 fournies par l'ADEME provenant de l'étude sur le refroidissement en centre de données, menée par Critical Building en parallèle de cette étude.

Cette projection nous donne la répartition suivante, la cible de répartition à 2035 est une hypothèse fournie par l'ADEME.

Répartition projetée	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035
Groupes frigorifiques à condensation à air + terminaison à air	40.00 %	39.0 9%	38.18 %	37.2 7%	36.3 6%	35.4 5%	34.5 5%	33.6 4%	32.7 3%	31.82 %	30.91 %	30.0 0%
Groupes frigorifiques à condensation à eau + terminaison à air	54.00 %	53.1 8%	52.36 %	51.5 5%	50.7 3%	49.9 1%	49.0 9%	48.2 7%	47.4 5%	46.6 4%	45.8 2%	45.0 0%
Détente directe dans l'air	3.00%	2.77 %	2.55 %	2.32 %	2.09 %	1.86 %	1.64 %	1.41 %	1.18 %	0.95 %	0.73 %	0.50 %
Détente indirecte dans l'eau (chaude)	3.00%	2.77 %	2.55 %	2.32 %	2.09 %	1.86 %	1.64 %	1.41 %	1.18 %	0.95 %	0.73 %	0.50 %
Direct Liquid Cooling	0.00%	1.82 %	3.64 %	5.45 %	7.27 %	9.09 %	10.91 %	12.7 3%	14.5 5%	16.36 %	18.18 %	20.0 0%
Immersion Cooling	0.00%	0.45 %	0.91 %	1.36 %	1.82 %	2.27 %	2.73 %	3.18 %	3.64 %	4.09 %	4.55 %	5.00 %

Répartition des technologies de refroidissement dans les centres de données français avec projection de 2024 à 2035

Nous avons combiné ces répartitions avec le potentiel de récupération de chaleur fatale permis par chaque technologie, en nous appuyant sur les données de l'étude sur la récupération de chaleur fatale en centre de données, également commandée par l'ADEME et en cours lors de notre étude.

Ce potentiel, exprimé en pourcentage pour 2020 et 2030 dans l'étude initiale, a été projeté jusqu'à 2035 en s'appuyant sur des hypothèses de l'ADEME :

	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035
Boucle d'eau glacée basse température < 20 C et boucle d'eau glacée moyenne température < 40 C	20.00 %	20.91%	21.82 %	22.73 %	23.64 %	24.55 %	25.45 %	26.36 %	27.27 %	28.18 %	29.09 %	30.00 %
Refroidissement par free cooling direct à air	5.00 %	5.45%	5.91%	6.36%	6.82 %	7.27 %	7.73 %	8.18 %	8.64 %	9.09 %	9.55 %	10.00 %
Refroidissement terminal par eau	40.00 %	41.82%	43.64 %	45.45 %	47.27 %	49.09 %	50.91 %	52.73 %	54.55 %	56.36 %	58.18 %	60.00 %
Refroidissement terminal par bain d'huile	60.00 %	60.91%	61.82 %	62.73 %	63.64 %	64.55 %	65.45 %	66.36 %	67.27 %	68.18 %	69.09 %	70.00 %

Le potentiel net de récupération de chaleur fatale exposé dans la section [Modélisation du potentiel de récupération de chaleur fatale](#) est obtenu en appliquant la somme des potentiels par technologie, chacun étant calculé avec la formule :

$$PotentielNet_{technologie A / scénario X} = Énergie consommée en France_{scénario X} \times Répartition_{technologie A} \times Potentiel_{technologie A}$$

3.5. Modèle 2035-2060

Le modèle bottom-up de CLIK permet de modifier les gains en efficacité par composant IT et par type de centre de données : 3 types de serveurs, disques SSD et HDD, ports réseau, etc.

C'est le choix que nous avons effectué pour modéliser les scénarios prospectifs de 2023 à 2035.

Pour les estimations de 2036 à 2060, il ne nous paraît pas pertinent de gérer ce niveau de détail. Notre approche est de prendre la moyenne sur les 5 dernières années (2031-2035) des consommations et du volume d'usage (workloads, AI workloads, Hashs) pour obtenir des gains en efficacité plus macros en kWh/workload pour chaque type de centre de données.

4. Scénarios : approche de révision des scénarios Transition 2050 et scénario tendanciel

4.1. Approche globale

Le travail de scénarisation s'inscrit dans la continuité de la mise à jour menée parallèlement par d'autres départements de l'ADEME pour actualiser l'exercice Transition(s) 2050. Les quatre scénarios alternatifs au scénario tendanciel s'appuient sur les mêmes visions prospectives d'un futur possible. Par souci de clarté et de cohérence, ils reprennent les mêmes intitulés que ceux utilisés précédemment.

Plusieurs dimensions sont communes à tous les scénarios :

- **La trajectoire nationale d'adaptation climatique** (scénario de référence de la TRACC, parue en 2025) : +4°C de réchauffement climatique en France métropolitaine en 2100, +3°C en moyenne mondiale
- **Les évolutions démographiques** modélisées par l'INSEE (scénario central appliqué à tous les scénarios) : la population commence à décroître à partir de 2050

Chaque scénario est construit en plusieurs parties qui font système, certains paramètres font l'objet d'une description qualitative sans intervenir dans le modèle, d'autres sont traduits dans le modèle :

- **Contexte global du secteur numérique**
 - Cadres de régulation, contexte géopolitique, configuration de la gouvernance du secteur numérique (dimension politique)
 - Évolutions des usages numériques des particuliers comme du monde économique (dimensions sociale et économique)
- **Application du contexte globale au secteur des centres de données**
 - Rythme et trajectoires de créations de nouveaux DC
 - Types de centres de données créés (hyperscale, traditionnels, colocation, IA, edge³¹¹)
 - Géographie des implantations
 - Stratégies d'adaptation au changement climatique
 - Intégration des attentes en termes de souveraineté
 - Porteurs du changement
- **Impacts centres de données sur les ressources**
 - Sols
 - Eau
 - Énergie
- **Caractéristique des évolutions technologiques : informatique**
 - IA et IA Gen
 - Évolution de la numérisation
 - Move to cloud
 - Blockchain et crypto monnaies
 - Quantique
- **Caractéristique des évolutions technologiques : chaud et froid**
 - Techniques de refroidissement
 - Récupération de chaleur fatale

4.2. Scénario tendanciel

- **Contexte global du secteur numérique**

Gouvernance

Les politiques publiques interviennent peu dans la régulation écologique des marchés technologiques. Elles interviennent davantage sur l'information et la régulation juridique (libertés numériques, protection des données personnelles). Sur le plan technologique et économique, les politiques publiques sont incitatives (dérogations environnementales, fiscalité avantageuse (TICFE, accélérations des procédures). Les problématiques écologiques sont marginalement ou pas adressées. La nouvelle donne géopolitique

³¹¹ Il faut noter que les centres de données edge, étant donné le flou de leur définition, ne sont pas inclus en tant que tel dans les projections. Nous avons cependant identifié leurs tendances de croissance de façon qualitative dans la scénarisation.

de 2025 (politiques douanières et guerre économique des Etats-Unis, réarmement des armées européennes) pousse l'Union Européenne et la France à vouloir développer fortement leurs propres entreprises numériques, tablant ainsi sur moins de régulation, plus d'incitations et l'accroissement de l'usage des nouvelles technologies et de la donnée liées à la défense.

Usages

L'extension du périmètre des usages concernés par la numérisation se fait par la croissance des usages existants : en lien avec la croissance démographique, la multiplication des supports matériels, la banalisation des pratiques numériques au sein de toutes les tranches d'âge et de tous les secteurs (sport, éducation, culture, services publics, défense...), et le move to cloud généralisé des entreprises. Elle se réalise également avec le développement d'usages nouveaux et récents : captation accrue de l'attention par les Réseaux Sociaux (RS), IA générative grand public, streaming... On note en particulier la poursuite de l'adoption et de l'intégration de l'IA dans des outils du quotidien (mainstreaming) ; l'intensification de l'usage de l'IA générative grand public (toutefois légèrement ralentie par une hypothèse d'un passage au payant à partir de 2027) ; la diffusion progressive de l'IA dans les grandes entreprises (moins dans les TPE-PME) et son intégration forte dans les activités de défense, d'assurance et de services financiers.

La stabilisation puis la baisse des usages numériques sont fortement liées à la baisse démographique à partir de 2050, mais aussi à la saturation du marché de la numérisation et les tensions croissantes sur les ressources.

- Application du contexte globale au secteur des centres de données

Mode de développement : très forte croissance des centres de données

Il s'agit ici de la poursuite des tendances observées dans la Base de Données CLIK sur la période 2020-2025 (projets compris).

Géographie

Les centres de données sont localisés dans les grandes métropoles régionales de l'hexagone. La concentration francilienne initiale se résorbe très progressivement, les territoires ruraux et périurbains des grandes métropoles accueillant les hyperscale notamment concernant des usages spécialisés comme l'IA ou son entraînement. Ce sont plus précisément les territoires suivants qui sont concernés : métropoles du Grand Paris et région Ile-de-France (Essonne, Seine et Marne, Seine St-Denis) ; métropoles Aix-Marseille, Nantes, Toulouse, Bordeaux, Lille-Roubaix-Tourcoing, Strasbourg, Rennes, Montpellier.

L'implantation des centres de données ne prend pas en compte les enjeux sur les évolutions climatiques des territoires (risques spécifiques locaux liés aux dérèglements climatiques).

Adaptation climatique

L'adaptation est avant tout technique, à l'échelle du bâtiment, pour minimiser les impacts des canicules notamment (isolation, refroidissement, toits blancs...). Les centres de données utilisent les méthodes de refroidissement et d'adaptation qui sont les plus rentables indépendamment de leur impact environnemental. Les systèmes de refroidissement peuvent ainsi être surdimensionnés pour pouvoir gérer le réchauffement global. Cela peut être considéré comme une mal adaptation. L'expansion du marché des centres de données est tellement rapide et importante que les opérateurs de centres de données estiment qu'intégrer le critère du risque climatique pourrait les ralentir, ou les conduire à ne pas exploiter certains fonciers par exemple.

Souveraineté

La politique publique vise à développer des champions nationaux, mais la dépendance vis-à-vis des Etats-Unis en termes d'importation pour le traitement et le stockage des données reste importante au regard des usages.

- Impacts centres de données sur les ressources

Sols

On prévoit une forte croissance de la consommation et du rythme d'artificialisation des sols par centres de données et les postes sources électriques.

Eau

Avec un très fort développement des clouds et hyperscales, combiné à l'augmentation du risque de canicule, certaines analyses anticipent une pression accrue sur l'eau dans les années à venir. Ce phénomène serait accentué si le recours à l'adiabatique se développait, l'adiabatique contribuant à accroître les déséquilibres hydrologiques, tout particulièrement dans les zones déjà soumises à un stress hydrique. Cependant, la variabilité du prix de l'électricité rend difficile l'anticipation des choix stratégiques du secteur sur le recours à l'adiabatique.³¹²

Système énergétique

Le fort développement des centres de données nécessite de nouveaux investissements pour le renforcement des réseaux de transport et de distribution d'électricité. Les générateurs de secours sont davantage utilisés (lors des événements climatiques extrêmes qui génèrent des coupures d'électricité plus fréquentes), ce qui entraîne des effets néfastes sur la qualité de l'air. A partir de 2030, la diffusion des générateurs HVO (biocarburant) limite l'impact sur l'air. Certains centres de données mettent volontairement leur batteries et équipements de secours à disposition pour aider le réseau à la marge.

- Caractéristique des évolutions technologiques : informatique

IA Générative et IA Générative optimisée³¹³

Traitement et inférences : l'évolution des usages de l'IA se traduit par un doublement des workloads d'IA générative entre 2024 et 2025, puis le taux de croissance annuel baisse régulièrement jusqu'à atteindre 20% par an en 2034 et 2035.³¹⁴

L'hypothèse de répartition des workloads est à 90% IA générative et 10% IA générative optimisée, pour chaque année à partir de 2026 (0% IA générative optimisée en 2024 et 2025). Comme décrit précédemment, ce paramètre permet de rendre compte de la répartition des workloads entre IA générative (très gourmande) et IA générative dite "optimisée", c'est-à-dire dont la configuration technique lui permet d'atteindre un niveau de consommation bien plus faible, comparable à l'IA précédent l'arrivée de l'IA générative. Cela veut dire que le scénario tendanciel considère que 10% de l'IA

³¹² Une hausse importante des prix de l'électricité se solderait par une utilisation croissante de l'eau pour refroidir (systèmes adiabatiques) pour réduire les dépenses énergétiques.

³¹³ Dans le modèle CLIK, afin de rendre compte de manière distincte des évolutions des usages avant et après la généralisation de l'usage de l'IA générative de la part du grand public, nous avons créé des paramètres quantitatifs "Gen AI workloads". Un de ces paramètres permet de rendre compte de l'évolution dans le temps de la charge de travail (workloads) des serveurs dédiés à l'IA générative en particulier (plus intensives électriquement que les autres types de calculs) que ces serveurs dédiés se situent dans des centres de données dédiés à l'IA ou non. Un autre de ces paramètres permet de rendre compte de la répartition des workloads entre IA générative (très gourmande) et IA générative dite "optimisée", c'est-à-dire dont la configuration technique lui permet d'atteindre un niveau de consommation bien plus faible, comparable à l'IA précédent l'arrivée de l'IA générative.

³¹⁴ L'hypothèse d'évolution annuelle des usages est modulée de façon non linéaire : en effet, avec un modèle linéaire d'évolution en % par rapport à l'année précédente, avec une hypothèse de 100% d'augmentation d'une année à l'autre, on assisterait chaque année à un doublement des usages. Dans les faits, une telle croissance n'intervient que très rarement et ponctuellement, dans le cas de l'essor massif d'un nouvel usage (ici l'IA générative). Nous modulons donc à la baisse le taux de croissance annuelle pour rendre compte d'une évolution des usages qui subit un essor important à un instant T, puis au fur et à mesure qu'il entre dans les pratiques, stabilise sa croissance.

généralisée à partir de 2026 est "optimisée" pour réduire sa consommation. L'évolution globale des workloads d'IA générative tient donc ainsi compte que 10% de ces workloads ont une consommation optimisée.³¹⁵

Stockage : l'essor de l'IA générative participe, via la forte croissance du stockage de données humaines et synthétiques, à l'essor des besoins en volume de stockage.

Numérisation

La digitalisation de l'économie et des usages français se traduit par une augmentation globale de 20% par an des workloads de calcul et d'une augmentation de 33% par an jusqu'à 2035 du volume de stockage global liés aux usages français.

Move to cloud

Accélération du move to cloud des entreprises et des individus : elle se traduit par une baisse de -15% / an des workloads des centres de données dits traditionnels et une augmentation de +0,4% des workloads au sein des centres de données hyperscale.³¹⁶

Blockchain et crypto monnaies

La croissance se poursuit : le marché devrait passer de 12 milliards de dollars en 2022 à 42 en 2025 (voir [plus haut](#) dans le rapport), mais en la matière, comme nous l'avons montré, on note que la majorité des capacités de calcul des centres de données sont importées par la France.

Cela se traduit par l'hypothèse d'une augmentation de +30% par an des hashes blockchain jusqu'en 2030, puis une baisse du taux de croissance par paliers de 2% jusqu'à +10% de croissance par an en 2040, puis +9% de croissance annuelle entre 2041 et 2050 et +8% de croissance par la suite.

Proof of Work : majoritaire aujourd'hui, plus le temps passe, plus il est difficile de basculer vers du Proof of Stake.³¹⁷ Hypothèse de 80% PoW pour 20% de PoS chaque année.

Quantique

Bien que l'on observe aujourd'hui des innovations telles que la puce Majorana, leur visibilité à court terme demeure limitée ; leurs impacts significatifs ne devraient se manifester qu'à l'horizon 2035-2040, dans le cadre des technologies émergentes. Les seuls effets sensibles du quantique avant 2035 sont liés à l'évolution dite "post-quantique" des méthodes de cryptographie : l'objectif de ces méthodes est de protéger dès aujourd'hui, des méthodes de décryptage quantique de demain. Cette évolution des méthodes entraîne une augmentation des besoins en calcul, ni mesurable ni modélisable à date dans le cadre de cette étude.

- Caractéristique des évolutions technologiques : chaud et froid

Refroidissement

Sont utilisés dans ce scénario les techniques suivantes :

- Concernant la production de froid
 - Avec groupes frigorifiques à condensation à air ou à eau et boucle d'eau glacée, principalement
 - À la marge et de manière combinée avec un système classique de production de froid : utilisation de l'air extérieur avec free-cooling

³¹⁵ Pour en savoir plus, consulter les chapitres [Méthodologie du modèle](#) et [Hypothèses](#)

³¹⁶ Hypothèse CLIK

³¹⁷ En effet, les choix techniques réalisés au passé encadrent et pré-déterminent les choix techniques futurs. Plus un choix technique est prédominant et en place depuis longtemps, plus l'effort à fournir pour s'en extraire est important et coûteux.

- À la marge, mais de manière croissante, en complément avec les deux autres systèmes cités, refroidissement adiabatique
- Concernant la terminaison (refroidissement en salle IT)
 - Le refroidissement à air reste majoritaire dans les datacenters non spécialisés pour l'IA générative
 - Le refroidissement par liquide (DLC) est la norme pour les centres de données spécialisés pour l'IA générative, de même que pour les centres de données dits HPC
 - À la marge, le DLC est implémenté dans certaines salles IT de centres de données de colocation, privés ou public, dans sa version "door cooling", soit jusqu'à la porte puis prolongé par un refroidissement par air amélioré au niveau des machines

4.3. Scénario Génération Frugale (S1)

Dans le scénario Génération Frugale, les pouvoirs publics mettent un terme à des usages numériques jugés inutiles et nocifs. Les technologies légères sont privilégiées, peu gourmandes, et quand c'est possible, des alternatives plus sobres remplacent le numérique. Le tout-numérique laisse place à une approche par plus de socialisation humaine. L'intelligence artificielle est réservée à des domaines critiques comme la santé, le climat, la défense ou la recherche. La course aux centres de données est temporairement freinée par un moratoire, afin d'épargner les territoires dont l'équilibre environnemental est déjà fragile. Enfin, les villes repensent leurs espaces publics : on y bouge un peu, ou beaucoup, on y sort et on y favorise les initiatives collectives... loin des écrans.

- Contexte global du secteur numérique

Gouvernance

Les politiques publiques ont un fort impact et mobilisent à la fois interdictions, incitations et régulations. Les interdictions concernent notamment les usages les plus toxiques du secteur numérique (captation de l'attention sur les réseaux sociaux, incitation à utiliser l'IA systématiquement et sans consentement des utilisateurs, publicité invasive...), et l'installation de centres de données dans les territoires aux fortes vulnérabilités environnementales (eau, sols, qualité de l'air, Ilots de Chaleur Urbains...).

Les régulations ont pour objectif une numérisation raisonnée et limitée de certains secteurs comme l'éducation ou les services publics. Les conditions d'implantation des centres de données sont liées à la stratégie climatique, et obligent à la concertation au plus tôt dans le projet ou encore le fléchage de l'IA uniquement pour les usages à haute valeur ajoutée sociale et environnementale.

Les incitations visent à proposer des alternatives basse technologie ou humaines à certains services par exemple la création de davantage de guichets humains pour les services de la CAF, ou le partage de matériel informatique, à renforcer la recherche publique pour plus de dispositifs légers (par exemple favoriser l'affichage papier dans les transports plutôt que les écrans), mais aussi à fortement orienter les populations, en particulier les enfants, vers des activités de plein air et de lien social (piétonnisation de rues, développement de squares et parcs, espaces capables pour une diversité de sports).

Usages

Le contexte réglementaire, une fois mis en place à partir de 2030, limite la numérisation de nouveaux usages (éducation, services publics, culture, objets non-connectés), réduit certains usages existants (réseaux sociaux, publicité) et en réoriente un certain nombre vers des dispositifs non-numériques (guichets humains, supports analogiques), mutualisés (tiers lieux numériques par exemple) ou de basse technologie avec des langages de programmation légers. Rapidement, dès 2026, des limitations à l'usage de l'IA sont mises en place. Elle est réservée à certains usages critiques : santé, climat, défense, recherche. Son

déploiement doit se faire systématiquement avec le consentement des utilisateurs. Le public des enfants est particulièrement ciblé : les espaces publics extérieurs des villes leur permettent de jouer dehors en sécurité, des activités sportives et créatives sans numérique sont massivement proposées. À partir de 2030, les usages se sont fortement transformés, **le discernement dans les choix technologiques est le maître mot. Les technologies numériques ne sont pas utilisées systématiquement mais là où elle apporte une véritable plus-value, selon des critères décidés collectivement visant l'intérêt général.**

- Application du contexte globale au secteur des centres de données

Modes de développement

Un moratoire sur l'implantation de nouveaux centres de données est mis en place : les centres de données se développent très peu, à l'exception de ceux dont le chantier est déjà entamé, et dont les chantiers s'accroissent, poussés par la crainte du moratoire. Des conditions drastiques sont nécessaires pour l'implantation de nouveaux centres de données. Par ailleurs, la baisse des usages du numérique liée à la forte évolution des modes de vie et de la place du numérique dans la société, a un impact à la baisse significatif sur la demande, tant au niveau des centres de données en France que sur l'usage importé des centres de données présents hors du territoire. Les effets de cette baisse de la demande sont plus sensibles à moyen terme que les effets du moratoire, qui eux le sont à plus long terme.

Géographie

La baisse des usages entraîne la diminution du nombre de centres de données. Aucun centre de données ne s'installe dans les territoires peu denses. Le développement de nouveaux centres de données est limité et leurs surfaces de plancher et emprises au sol sont plafonnées.

Il y a moins de centres de données en Ile-de-France ; une légère croissance dans les métropoles régionales et les villes moyennes, là où les risques environnementaux sont limités.

La localisation des quelques centres de données supplémentaires et des centres de données relocalisés découle d'une planification écologique, climatique et spatiale, fondée sur le croisement de données environnementales et de connaissances terrain. Alors que l'Ile de France accueillait en 2025 40% des centres de données français, on assiste à un déplacement vers des territoires moins risqués, en Ile-de-France ou ailleurs (Comme suite à l'ouragan Sandy en 2012 à NYC, quand de nombreux centres de données s'installent dans le New Jersey pour échapper aux inondations). Des centres de données ferment dans la région capitale, notamment les plus anciens, installés en zones inondables.

Adaptation au changement climatique

La stratégie d'implantation des centres de données intègre comme critère majeur les aléas climatiques. Par exemple, pas d'installation dans les secteurs où les aléas de type Îlot de Chaleur Urbaine sont élevés ; où les risques de submersion, inondation par débordement ou ruissellement sont avérés ; ou encore où la ressource en eau est fragile. **Le maître mot est "éviter".**

Souveraineté

La baisse des usages numériques est une stratégie de recherche de l'autonomie technique par d'autres moyens et de réappropriation technique et des usages.

- Impacts des centres de données sur les ressources

Sols

La consommation et la dynamique d'artificialisation des sols sont très faibles puis nulle à partir de 2030.

Eau

La baisse des usages et les localisations de centres de données en fonction de critères climatiques et environnementaux limitent très fortement l'impact sur la ressource en eau.

Energie

La baisse des usages et corrélativement celle des centres de données limitent leur impact sur le système électrique.

- Caractéristique des évolutions technologiques : informatique

IA Générative et IA Générative optimisée

Traitement et inférences : l'évolution des usages de l'IA se traduit par une croissance très forte des workloads IA en 2025, alors que l'influence du changement de mode de vie poussée dans ce scénario n'a pas encore de prise ; puis nous faisons l'hypothèse d'un tassement important de cette croissance au fur et à mesure que ces changements s'opèrent, sans toutefois aller jusqu'à une réduction du nombre de workloads par an. Ainsi, la croissance du nombre de workloads IA par an se tasse progressivement pour atteindre +20% par an en 2030, puis +8% par an en 2035, puis se stabiliser à +5% par an à partir de 2045. L'hypothèse de répartition des workloads est à 80% IA générative et 20% IA générative optimisée, pour chaque année.

Stockage : Les limitations apportées à l'IA générative participent à la limitation des besoins en volume de stockage.

Evolution de la numérisation

La mise en place d'une politique forte en termes de sobriété numérique et d'alternatives non-numériques aux usages précédemment digitalisés se traduit par une limitation de l'augmentation globale de 20% par an des workloads de calcul liés aux usages en France, en 2024 et 2025, puis nous prenons une hypothèse de tassement de cette croissance jusqu'à 0% en 2034, puis une baisse jusqu'à -10% par an à partir de 2038.

Concernant le volume de stockage, nous prenons l'hypothèse d'une augmentation de 33% par an jusqu'en 2025 puis un tassement régulier de cette augmentation jusqu'à 0% en 2031 et une baisse de 3% par an du volume de stockage global liés aux usages en France.

Move to cloud

Dans le scénario Génération Frugale, nous explorons une stabilisation des proportions entre centres de données de type hyperscale, colocation et traditionnel, avec un ralentissement du move to cloud. L'hypothèse est faite d'un arrêt de l'érosion des centres de données traditionnels versus ceux du cloud en 2031. Cela se traduit par une hausse des workloads colocation, une baisse de 5% / an des workloads des centres de données dits traditionnels jusqu'en 2031 puis une stabilisation et une stagnation (+0% / an) des workloads des centres de données hyperscales - en proportion du total des workloads. Les proportions passeraient donc de 55% / 40% / 5% des workloads respectivement en hyperscale / colocation / traditionnel, en 2024 à 55,7% / 40,8% / 3,5% en 2031 et les années suivantes dans ce scénario. À noter que cette répartition vaut pour l'ensemble des workloads alimentant les usages des Français, pas uniquement les workloads des centres de données présents sur le territoire.

Par ailleurs, nous faisons l'hypothèse dans ce scénario, de la définition d'enveloppes limitées de stockage pour les individus, les acteurs publics et les entreprises : cela se traduit par une limitation de la hausse des workloads au sein des centres de données hyperscale, et la baisse globale des workloads, sous l'effet conjoint de cette mesure et d'autres.

Blockchain et crypto monnaies

Dans ce scénario, un recul de l'intérêt pour les usages de la Blockchain et des crypto-monnaies se dessine. Cela se traduit ici par une hypothèse de baisse de tassement de la croissance des hashes (workloads Blockchain), qui passe d'une hypothèse de 10% de croissance par an en 2024 à 5% de croissance par an en 2030, puis 4% par an entre 2036 et 2050 et 3% par an par la suite.

Proof of Work : majoritaire aujourd'hui, plus le temps passe, plus il est difficile de basculer vers du Proof of Stake. Nous prenons une hypothèse de 80% PoW pour 20% PoS entre 2024 et 2026, et modélisons un retournement progressif des proportions pour atteindre 20% PoW pour 80% PoS en 2032.

Quantique

Le quantique se développe marginalement, principalement pour la recherche. Son traitement s'inscrit dans la perspective des technologies émergentes à partir de 2035.

- Caractéristique des évolutions technologiques : chaud et froid

Refroidissement

Changements sur le parc actuel :

- Dans ce scénario, on privilégie les techniques de air cooling et free-cooling
- Un volet réglementaire impose le respect strict de la norme EN 50600-4-2 : audit et publication annuelle du PUE, rapport d'audit avec niveau de calcul à la clef
- Des objectifs chiffrés de baisse du PUE plus drastiques que le décret tertiaire sont édictés.³¹⁸

4.4. Scénario Coopération Territoriale (S2)

Comme Génération Frugale, le scénario Coopération Territoriale cherche à réduire notre dépendance au numérique. Ici la coopération et l'incitation sont les leviers du changement. Les projets de centres de données sont développés en concertation avec les territoires et encadrés par des normes - notamment en matière de consommation ou de surface - limitant ainsi leur prolifération et protégeant les écosystèmes fragiles. Des coopérations entre industriels favorisent des démarches d'Écologie Industrielle et territoriale. L'intelligence artificielle est principalement dédiée aux secteurs prioritaires, comme la santé, le climat, la recherche, la défense et l'optimisation logistique. En optant pour une politique incitative, Coopération Territoriale soutient celles et ceux qui souhaitent se tourner vers des pratiques à faible impact environnemental, tout en laissant à chacun la liberté de rester autant connecté s'il le souhaite.

- Contexte global du secteur numérique

Gouvernance

Les outils de régulation et d'incitation sont davantage mobilisés que ceux de l'interdiction.

Les politiques publiques interviennent dans la régulation des marchés technologiques pour les orienter vers des usages vertueux environnementalement. Une stratégie nationale concertée d'implantation des centres de données est mise en place sur le territoire français, elle compose avec les singularités de chaque territoire et leurs capacités à accueillir certains types de centres de données. Elle limite les effets clusters grâce à une stratégie de déconcentration numérique. Les outils de régulation et d'incitation sont avant tout mobilisés pour à la fois éviter l'envolée du trafic de données et encadrer les impacts

³¹⁸ Cf. modèle CLIK

environnementaux du numérique : réglementation de la captation de l'attention et de la collecte de données par les services numériques, incitation à l'écoconception des services numériques, ...

Usages

Une dichotomie apparaît à partir de 2027 entre des usagers numérique-intensifs d'une part, et des numérique-décroissants d'autre part, qui sont de plus en plus nombreux à partir de 2030. Les premiers ont très fortement digitalisé toutes leurs activités, utilisent de multiples supports et objets connectés, tandis que les seconds s'inscrivent dans une recherche de sobriété de leurs usages numériques (limitation du nombre d'équipements individuels et du temps passé par jour sur les écrans, volonté de réaliser certaines activités quotidiennes hors numérique : vie sociale et culturelle, achats, ...). Cette dichotomie se retrouve dans le monde économique et chez les acteurs publics. La situation suscite ainsi un débat de société fort et aboutit à des choix communs vers plus de coopération territoriale dès 2030 pour concilier la diversité des usages et leurs évolutions tout en les réduisant : des acteurs locaux qui mutualisent leurs besoins et infrastructures, branchées sur des productions énergétiques locales ; des lieux mutualisés d'accès au numérique ; des services pour réparer et mutualiser le matériel informatique...

Les usages de l'IA sont prioritairement axés sur les secteurs de la santé, du climat, de la défense, de la recherche et pour améliorer le métabolisme territorial et les flux logistiques, en lien avec une stratégie de planification écologique.

- Application du contexte globale au secteur des centres de données

Modes de développement

Le développement des centres de données est un peu moins contenu que dans le scénario Génération Frugale.

Géographie

Une politique de déconcentration numérique de l'Ile-de-France se met en place, à l'image de la politique de déconcentration industrielle de la région parisienne dans les années 1970-80. Elle vise à ce que chaque région/département dispose de centres de données en lien avec ses besoins. La localisation des centres de données découle d'une volonté de répartition équitable des centres de données dans les territoires, en fonction de leurs besoins en emploi, calcul et en lien avec leurs énergies locales. Ils se développent notamment dans les territoires moins dotés aujourd'hui : Nantes, Toulouse, Besançon, Clermont-Ferrand, Reims...

Adaptation au changement climatique

Les centres de données co-investissent dans des actions d'adaptation climatique à l'échelle de leurs territoires d'implantation pour minimiser les effets d'îlots de chaleurs urbains, mais aussi en se raccordant aux réseaux urbains de chaud et de froid (comme à Stockholm où les centres de données donnent la chaleur fatale du RCU en échange d'un refroidissement gratuit³¹⁹). Leurs localisations sont concertées pour éviter les concurrences d'usage sur le foncier et la ressource électrique. **Le maître mot est "équilibrer".**

Souveraineté

Les politiques publiques visent à développer des champions européens.³²⁰

- Impacts centres de données sur les ressources

Sols

³¹⁹ <https://www.institutparisregion.fr/nos-travaux/publications/limpact-spatial-et-energetique-des-data-centers-sur-les-territoires/>

³²⁰ Voir par exemple le projet Eurostack : https://www.euro-stack.info/docs/EuroStack_2025.pdf

La consommation et l'artificialisation des sols sont très faibles.

Eau

La baisse des usages, mais aussi le raccordement de centres de données sur les réseaux urbains de froid limitent fortement l'impact sur la ressource en eau.

Energie

De nouveaux acteurs du chauffage sont créés, les centres de données deviennent des équipements hybrides. Des systèmes énergétiques plus locaux se dessinent.

- Caractéristique des évolutions technologiques : informatique

IA Générative et IA Générative optimisée

Traitement et inférences : après l'engouement pour l'IA du début des années 2020, l'hypothèse est faite dans ce scénario d'un tassement de la croissance des usages de l'IA : 35% de croissance des workloads IA en 2028, 25% en 2029, 20% en 2030, 15% en 2032, puis 10% par an entre 2033 et 2037, puis 8% par an par la suite.

Nous prenons l'hypothèse de répartition suivante des workloads : 80% IA générative et 20% IA générative optimisée, pour chaque année à partir de 2027 (pas d'IA générative optimisée auparavant).

Stockage : Les limitations apportées à l'IA générative participent à la limitation des besoins en volume de stockage.

Evolution de la numérisation

La limitation des usages intervient mais de façon moins marquée que pour le scénario Génération Frugale : cela se traduit par une augmentation globale de 20% par an des workloads de calcul liés aux usages en France, de 2024 à 2026, puis nous prenons une hypothèse de tassement de cette croissance jusqu'à 0% par an en 2035 et pour les années suivantes.

Concernant le volume de stockage, nous prenons l'hypothèse d'une augmentation de 33% par an jusqu'en 2026 puis un tassement régulier de cette augmentation jusqu'à 0% en 2031 puis une baisse de 3% par an du volume de stockage global liés aux usages en France par la suite.

Move to cloud

Externalisation mesurée vers des opérateurs régionaux de cloud :

Cela se traduit par une hausse des workloads colocation plus que hyperscale au niveau des centres de données présents sur le territoire français. Dans le modèle CLIK, qui tient compte des workloads importés, cela se traduit par une baisse de 5% par an des workloads des centres de données dits traditionnels et une augmentation de +0,2% par an des workloads des centres de données hyperscales - en proportion du total des workloads.³²¹

Définition d'enveloppes limitées de stockage pour les individus, les acteurs publics et les entreprises, en lien avec des priorités environnementales régionalisées : cela se traduit par une limitation de la hausse des workloads au sein des centres de données hyperscale, et la baisse globale des workloads, sous l'effet conjoint de cette mesure et d'autres.

³²¹ Si sur les 10 premières années du modèle, cette tendance maintient dans une forte stabilité la proportion des workloads colocation (40% à 40,1% du total des workloads entre 2024 et 2034), l'essor global des hyperscales entraîne par la suite un léger tassement de la colocation dans les proportions.

Blockchain et crypto monnaies

Il y a ici une croissance légère des usages Blockchain et crypto. Cela se traduit par une hypothèse de croissance de 5% par an des hashes jusqu'en 2046, puis 4% par an par la suite.

Concernant la répartition entre Proof of Work et Proof of Stake, nous faisons l'hypothèse d'une absence d'évolution dans la répartition actuelle (80% Proof of Work et 20% Proof of Stake).

Quantique

Le quantique se développe marginalement, principalement pour la recherche. Son traitement s'inscrit dans la perspective des technologies émergentes à partir de 2035.

- Caractéristique des évolutions technologiques : chaud et froid

Refroidissement

De même que pour le scénario Génération Frugale, dans Coopérations Territoriales :

- Des interconnexions plus fortes sont opérées entre centres de données et réseaux de chaleur avec des opportunités de refroidissement en utilisant l'eau froide qui vient du réseau urbain qui en sortie va dans le réseau urbain d'eau chaude
- Des objectifs chiffrés de PUE moyen sont édictés, légèrement moins ambitieux que dans Génération Frugale.
- Sur les nouveaux centres de données IA : il y a des incitations à développer le DLC sur tout le matériel des centres de données (pas que IA), mais des systèmes de DLC partiels sont aussi mis en place (voir exemple AWS³²²)

4.5. Scénario Technologies Vertes (S3)

Dans le scénario Technologies Vertes, les politiques publiques misent avant tout sur l'innovation pour développer des énergies qui rejettent peu de gaz à effet de serre avec l'espoir de maintenir l'expansion continue des usages numériques via l'innovation et l'efficacité plutôt que la sobriété. L'intelligence artificielle s'installe durablement dans notre quotidien : au travail, dans nos loisirs ou encore dans nos services publics. Pour répondre à cette demande croissante, les centres de données se multiplient un peu partout sur le territoire : les enjeux territoriaux et environnementaux de leur implantation passent au second plan.

- Contexte global du secteur numérique

Gouvernance

Les politiques publiques interviennent avant tout pour inciter à l'innovation dans le domaine énergétique dans l'optique de minimiser l'impact du numérique, et des autres secteurs par le numérique, sans réduire les usages, et à favoriser l'implantation de centres de données en France à travers l'argument du mix énergétique français peu carboné.

Il s'agit du développement du nucléaire, d'énergies renouvelables et de récupération, et décarbonées, dont la récupération de chaleur fatale ; du développement d'hydrogène et de méthane vert, mais aussi

³²² <https://datacentremagazine.com/technology-and-ai/aws-builds-custom-liquid-cooling-in-11-months-for-ai-chips>

de l'organisation de l'effacement³²³. Des avancées technologiques importantes sont suscitées pour que les équipements de secours soient moins polluants, et plus efficaces.

Usages

La production importante d'énergies décarbonées et les usages d'optimisation permettent aux usages numériques de croître tout au long du modèle. Avec une dynamique démographique à la baisse à partir de 2050, on pourrait s'attendre à une baisse ou une stabilisation des usages du numérique. Toutefois, celle-ci est limitée par le fort développement des interactions de machines à machines, notamment via l'IoT et l'IA. Les usages se développent ici de façon relativement proche du scénario tendanciel. On note une forte utilisation de l'IA notamment pour optimiser les réseaux électriques et la connexion de multiples sources décentralisées d'énergie.

- Application du contexte globale au secteur des centres de données

Modes de développement

Ce scénario présente une très forte croissance des centres de données, en particulier en France via une politique favorisant leur implantation sur le territoire avec l'argument du mix électrique peu carboné.

Géographie

Ce scénario est proche du scénario tendanciel quand on observe la dynamique de consommation des centres de données en France, mais avec un tropisme fort vers la façade Ouest de la France, où de nouvelles productions électriques se déploient (éolien maritime, fermes solaires³²⁴).

Adaptation au changement climatique

L'adaptation technique des bâtiments, comme dans le scénario tendanciel, se couple à des investissements dans des unités de production d'énergie décarbonée, notamment locales, afin de réduire les impacts du secteur sur le climat et l'environnement.³²⁵

Souveraineté

Le succès de la transition énergétique constitue un levier de négociation avec les acteurs numériques non-européens pour qui l'exemplarité énergétique française est un facteur fort d'attractivité. Il devient donc possible de demander des conditions d'implantation plus ambitieuses : intégration urbaine, énergétique, architecturale...

- Impacts des centres de données sur les ressources

Sols

Le scénario Technologies vertes est similaire au tendanciel sur la consommation et l'artificialisation des sols : elles sont très importantes.

Eau

La haute disponibilité de l'électricité incite à ne pas trop utiliser l'adiabatique pour refroidir. L'impact sur la ressource en eau est donc relativement faible.

³²³ En cas de pic sur le réseau électrique, certains utilisateurs peuvent s'effacer : soit réduire ou arrêter temporairement leur activité ; soit avoir recours à des dispositifs de secours comme les génératrices pour ne plus utiliser l'électricité du réseau et limiter le pic. La réduction peut se faire par incitation tarifaire par exemple.

³²⁴ <https://www.rte-france.com/projets/nos-projets>

³²⁵ L'adaptation vise à la fois la réduction des GES mais aussi l'habitabilité des territoires, moins risqués, mieux adaptés, et notamment un lien plus fort entre ressources locales et consommations locales.

Energie

Les centres de données deviennent des acteurs énergétiques majeurs : co-investisseurs, co-producteurs et consommateurs massifs. Le réseau doit s'adapter fortement aux nouvelles productions locales décentralisées. Dans ce scénario, on observerait ici une baisse des générateurs de secours au fuel grâce à la massification des biocarburants (HVO), le développement de la pile à combustible, et à l'acceptation du seul double raccordement électrique comme sécurité (sans générateurs)³²⁶. Mais comme dans le tendanciel, de nouveaux investissements sur le renforcement des réseaux de transport et de distribution d'électricité sont nécessaires.

- Caractéristique des évolutions technologiques : informatique

IA Générative et IA Générative optimisée

Traitement et inférences : l'évolution des usages de l'IA se traduit par une croissance annuelle de 100% en 2024, puis le taux de croissance annuel diminue régulièrement jusqu'à atteindre 15% par an en 2035 et 10% par an par la suite. Une part de la croissance annuelle de l'IA est poussée par les usages liés à l'optimisation des énergies.

L'hypothèse de répartition des workloads se situe à 90% IA générative et 10% IA générative optimisée, pour chaque année à partir de 2027.

Stockage : L'essor de l'IA générative participe, via la forte croissance du stockage de données humaines et synthétiques, à l'essor des besoins en volume de stockage.

Évolution de la numérisation

La numérisation des usages en France se poursuit, avec une part de ces usages alloués à l'optimisation énergétique et la production d'énergies vertes : cela se traduit par une augmentation de 20% par an des workloads de calcul entre 2024 et 2026, puis cette augmentation se poursuit entre 2027 et 2038 à des taux soutenus de jusqu'à une stabilisation à 5% de croissance annuelle en 2039 et les années suivantes.³²⁷

Concernant le volume de stockage, hypothèse d'une augmentation de 33% par an jusqu'à 2035 du volume de stockage global lié aux usages en France.

Move to cloud

Dans ce scénario, il y a une accélération du *move to cloud* des entreprises et des individus. Les pouvoirs publics proposent des incitations au *move to cloud* vers les opérateurs produisant et utilisant des énergies vertes grâce à un système de labellisation comme il est fait par exemple pour le cloud souverain avec le label "SecNumCloud". De nouveaux centres de données de colocation, spécialistes du cloud, s'implantent en France, de même que des centres de données hyperscale. Dans le modèle CLIK, cela se traduit dans le mix de centres de données qui alimentent les usages des Français, par une baisse de 13% / an des workloads des centres de données dits traditionnels et une augmentation de +0,5% des workloads au sein des centres de données hyperscale.

³²⁶ Par exemple pour certains usages comme l'entraînement de modèles d'IA. L'usage des HVO est une ambition affichée par certains opérateurs pour le développement des futurs centres de données.

³²⁷ Il faut tenir compte ici comme pour tous les paramètres du modèle CLIK qui renvoient à une évolution annuelle, que chaque année est en pourcentage d'évolution par rapport à la précédente. Ainsi, il faut 4 ans d'une croissance de 20% des workloads par an pour doubler une quantité de workloads, ou 5 ans pour une croissance annuelle de 15% par an.

Blockchain et crypto monnaies

La croissance mondiale se poursuit : le marché devrait passer de 12 milliards de dollars en 2022 à 42 en 2025, cependant beaucoup de *mining* proviennent d'une capacité importée pour la France en la matière. Cela se traduit dans le modèle CLIK par une hypothèse de croissance de 25% par an des hashes jusqu'en 2030, puis atténuation de la croissance à 20% par an jusqu'en 2033, 15% par an jusqu'en 2037, puis 10% par an en 2038 et 2039, puis stabilisation à 8% par an par la suite.

Les proportions Proof of Work / Proof of Stake actuellement de respectivement 80% et 20% passent progressivement à respectivement 60% et 40% entre 2026 et 2030 et maintiennent cette proportion par la suite.

Quantique

Le quantique commence à se diffuser dans le domaine énergétique à partir de 2035.

- **Caractéristique des évolutions technologiques : chaud et froid**

Refroidissement

Dans ce scénario, les incitations concernent :

- Le free cooling
- Le DLC pour centres de données IA
- L'immersion cooling
- Le couplage avec récupération de chaleur
- L'aide à l'investissement et le développement des piles à hydrogène

4.6. Scénario Pari Réparateur (S4)

Le scénario du Pari Réparateur repose sur l'idée que les solutions aux problèmes climatiques viendront des technologies et qu'elles atteindront leur maturité à temps. Dans ce scénario, l'innovation n'a pas pour objectif premier de réduire les émissions de gaz à effet de serre à la source, mais plutôt de permettre la poursuite de nos modes de vie en misant sur des technologies capables d'éliminer les émissions une fois dans l'atmosphère. Concernant les centres de données plus spécifiquement, ce scénario envisage de compenser la forte hausse des émissions dues à leur développement par les baisses des émissions d'autres secteurs d'activité du fait d'une numérisation accrue. Ici, l'intelligence artificielle connaît un essor rapide, portée par l'émergence de nouveaux usages et la généralisation de la robotisation.

Les centres de données se multiplient et s'agrandissent, profitant d'une réglementation assouplie. Les territoires s'organisent pour répondre à la demande énergétique croissante. Cette dynamique entraîne une forte augmentation de la consommation d'énergie et des émissions de gaz à effet de serre. Pour y faire face, les politiques publiques misent sur l'innovation, en soutenant le développement de technologies de captation ou de compensation carbone, dans l'espoir d'atteindre — comme dans les trois autres scénarios de Transitions 2050 — la neutralité carbone à long terme.

- **Contexte global du secteur numérique**

Gouvernance

Les politiques publiques interviennent très fortement pour inciter à l'innovation au sein des marchés technologiques, notamment en lien avec la recherche de solutions d'adaptation au changement climatique.

Les politiques publiques régulent très peu les impacts environnementaux du numérique, car les innovations technologiques ont pour objectif de venir les compenser.

L'environnement est de plus en plus monitoré (capteurs), ce qui permet d'anticiper les événements extrêmes et de faire de la gestion de crise.

Les centres de données deviennent l'infrastructure la plus stratégique du pays avec le réseau électrique. À ce titre, les politiques publiques de facilitation et d'accélération de leur développement sont mises en place.

Usages

Ce scénario est relativement similaire au tendanciel lorsqu'on observe les consommations des centres de données présents sur le sol français, mais une consommation importée proportionnellement légèrement supérieure dans ce scénario que sur le scénario tendanciel. Bien que difficilement modélisables, des effets rebonds liés à un usage accru du numérique et de l'IA pour le monitoring des technologies climatiques et la robotisation de la société sont à anticiper. L'adoption de l'IA générative s'intensifie fortement, marquée par une généralisation de l'usage de modèles de grande taille et de haute complexité, sans réelle contrainte liée à leur consommation énergétique ou à leur coût d'utilisation. Ce sont des modèles d'IA polyvalente (non-spécialisée), capables de réaliser une multitude de tâches dans des contextes très variés. On observe ici une généralisation et une intensification de l'IA dans toute la société : dans la sphère privée mais aussi dans la sphère professionnelle où l'utilisation de modèles d'IA génératives généralistes et spécifiques est largement diffusée progressivement dans l'ensemble des métiers et des entreprises, pour automatiser le maximum de procédés répétitifs. La robotisation assistée d'IA est également adoptée dans l'industrie, la mobilité ou dans la vie usuelle. Les gros modèles généralistes sont la norme avec le développement de modèles plus spécifiques, ainsi qu'avec le développement d'agents permettant d'appliquer l'IA à tous les usages du numérique.

De plus, de nouveaux usages numériques émergent avec l'arrivée de nouveaux types d'IA :

- "L'IA générale ultime ou Superintelligence artificielle (ou même IA puissante) se distinguerait par ses capacités cognitives surpassant celles des humains, ayant notamment une conscience." ³²⁸
- "Le transhumanisme est un mouvement qui, en s'appuyant sur les progrès de la biologie et de l'intelligence artificielle, défend l'idée de transformer ou dépasser l'homme pour créer un post-humain, ou un transhumain, aux capacités supérieures à celles des êtres actuels. Cette transformation s'envisage au niveau individuel, mais aussi collectif, conduisant alors à une humanité nouvelle. Différentes facultés physiques ou mentales et cognitives de l'être humain seraient concernées : il verra dans l'obscurité, ne connaîtra plus la fatigue et ne se cassera pas le col du fémur en glissant... Ses capacités intellectuelles seront décuplées et sa mémoire prodigieuse." ³²⁹

- **Application du contexte global au secteur des centres de données**

Modes de développement

On note ici une très forte croissance des centres de données.

³²⁸ [Comprendre l'IA pour la démystifier](#)

³²⁹ [Transhumanisme : de l'illusion à l'imposture | CNRS Le journal](#)

Géographie

Ce scénario est similaire au tendanciel.

Adaptation au changement climatique

Ce scénario est similaire au tendanciel.

Souveraineté

Dans ce scénario, des investissements lourds des pays membres de l'UE sont consentis pour consolider des champions européens du numérique et développer une infrastructure commune de centre de données, robuste et capacitaire, répartis sur tout le territoire européen, exploités et utilisés par des acteurs européens. Ces choix permettent de réduire la dépendance aux Etats-Unis sur le stockage et le traitement de données, et d'améliorer le mix énergétique utilisé.

- Impacts centres de données sur les ressources

Sols

Ce scénario est similaire au tendanciel.

Eau

Sans régulation sur le WUE et du fait d'une incitation forte favorisant l'implantation et la gestion des centres de données en France, ce scénario envisage une explosion des systèmes adiabatiques, y compris là où les nappes sont fragiles. L'impact sur la ressource en eau, les petits cycles de l'eau et l'eau potable est donc très fort, même si les eaux de sortie sont parfois réutilisées pour d'autres usages industriels.

Energie

Ce scénario suit une courbe proche du scénario tendanciel, quoique légèrement inférieure concernant la consommation électrique des centres de données présents sur le territoire français ; puis son accélération, légèrement supérieure à celle du scénario tendanciel, fait passer la consommation électrique des centres de données présents en France au-dessus du scénario tendanciel dans ce scénario Pari Réparateur. Si l'on observe la consommation électrique globale de ce scénario, en tenant compte des émissions importées, ce scénario Pari Réparateur suit une croissance de ses consommations légèrement supérieure à celle du scénario tendanciel.

- Caractéristique des évolutions technologiques : informatique

IA Générative et IA Générative optimisée

Traitement et inférences : l'évolution des usages de l'IA se traduit par une croissance annuelle de 100% en 2024, puis le taux de croissance annuel baisse régulièrement jusqu'à atteindre un premier palier de croissance de 20% par an en 2032 et se stabiliser à 9% de croissance annuelle à partir de 2038.

L'hypothèse de répartition des workloads est à 90% IA générative et 10% IA générative optimisée, pour chaque année à partir de 2027.

Stockage : l'essor de l'IA générative participe, via la forte croissance du stockage de données humaines et synthétiques, à l'augmentation des besoins en volume de stockage.

Move to cloud

La digitalisation de l'économie et des usages en France entraîne une hausse annuelle des workloads estimée à 20% par an en 2025 et 2026, puis à 16% par an jusqu'en 2032, puis cette croissance annuelle se

poursuit à des taux plus faibles mais soutenus jusqu'à atteindre 10% en 2044 puis une stabilisation à 6% par an à partir de 2050.

Parallèlement, le volume de données stockées liées aux usages en France augmente en moyenne de 33 % par an jusqu'à 2035.

Blockchain et crypto monnaies

La croissance mondiale se poursuit plus fortement que dans le scénario tendanciel. Cela se traduit dans le modèle CLIK par une hypothèse de croissance de 45% par an des hashes jusqu'en 2031, puis on assisterait à une baisse de la croissance progressive jusqu'à une stabilisation à 10% par an à partir de 2039.

Les proportions Proof of Work / Proof of Stake actuelles restent identiques à celles actuelles (respectivement 80% / 20%).

Quantique

Le quantique se diffuse à une diversité de secteurs économiques après 2035.

- Caractéristique des évolutions technologiques : chaud et froid

Refroidissement

Dans ce scénario, les systèmes de refroidissement suivants sont utilisés :

- Air cooling
- Free cooling, et abadiatique pour le cloud
- Fort développement du DLC pour les centres de données IA (et comme centres de données spécifiques IA sont en croissance, le DLC prend une part plus importante de manière relative)
- Fort développement de l'immersion cooling

5. Hypothèses

5.1. Évolution de l'efficacité des serveurs

Sur la base des tendances initiales décrites au chapitre [Mise à jour des paramètres d'efficacité](#), nous avons modélisé l'évolution des efficacités de la façon suivante pour chacun des 5 scénarios :

Efficacité énergétique - puissances effectives	Volume servers ³³⁰	Mid -range server ³³¹	High -end ³³²	SSD ³³³	HDD ³³⁴
Hypothèses initiales Masanet	-1%	+8%	+7%	-3,7%	-5,3%
Hypothèses appliquées aux scénarios 2023-2035	+1,77%	+1,11%	-1,17%	0,00%	5%

Comprendre : Par exemple, pour les volumes servers, l'hypothèse proposée par le modèle Masanet est une réduction de la consommation de 1% par an par gain d'efficacité énergétique. L'hypothèse prise dans le modèle CLIK est une augmentation de la consommation d'énergie de 1,77% par an, sur la base des paramètres d'efficacité constatés³³⁵.

Le tableau ci-dessous rend compte des hypothèses prises en termes d'efficacité énergétique :

Efficacité d'usages		Workloads/serveurs		Stockage/Drive (TB/Drive)	
		Traditionnel	Cloud	SSD	HDD
Hypothèses initiales Masanet		+8%	+8%	+10%	+10%
Hypothèses appliquées aux scénarios 2023-2035	Scénarios Tendanciel, Technologies Vertes, Pari Réparateur	+8%	+8%	+10%	+10%
	Scénarios Génération	+10%	+10%	+12%	+12%

³³⁰ Serveurs d'entrée de gamme, cf. traductions dans le [glossaire](#) au début de l'étude

³³¹ Serveurs de milieu de gamme

³³² Serveurs hauts de gamme

³³³ Disques durs de type Solid State Drive

³³⁴ Disques durs de type mécanique / Hard Disk Drive

³³⁵ Au chapitre [Mise à jour des paramètres d'efficacité](#)

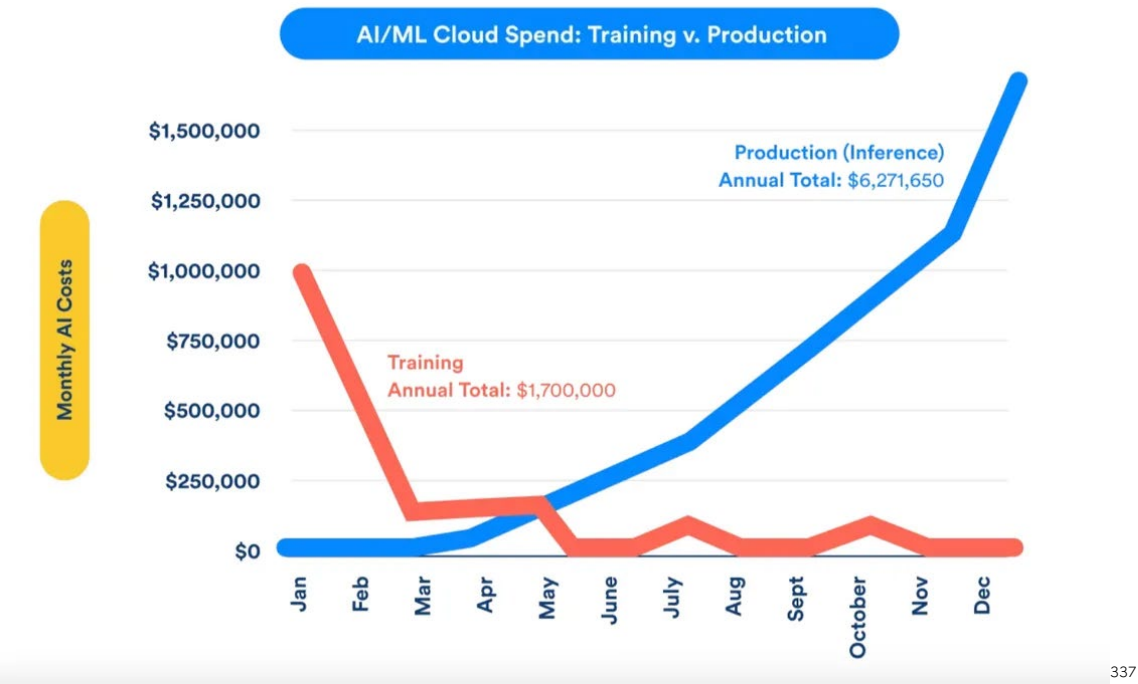
	Frugale, Coopérations Territoriales				
--	---	--	--	--	--

Concernant les efficacités d’usages, aucune information plus récente n’ayant permis de comparer les hypothèses proposées par Masanet, nous avons repris ces hypothèses pour la plupart des scénarios. Pour les scénarios Génération frugale et Coopérations Territoriales en revanche, nous avons retenu l’hypothèse d’un gain d’efficacité énergétique de 2 points par rapport aux autres scénarios. Ce gain repose sur une amélioration relative de l’efficacité des usages, qui serait liée à l’adoption systématique de bonnes pratiques par les architectes des clients des centres de données (conteneurisation, mise en place de VM, ...) ayant pour effet une augmentation forte du taux d'utilisation des serveurs existants. Ces bonnes pratiques permettent d’augmenter le nombre moyen de workloads par serveur, réduisant ainsi le besoin global en ressources matérielles.

5.2. Usages de l’IA de “nouvelle génération”

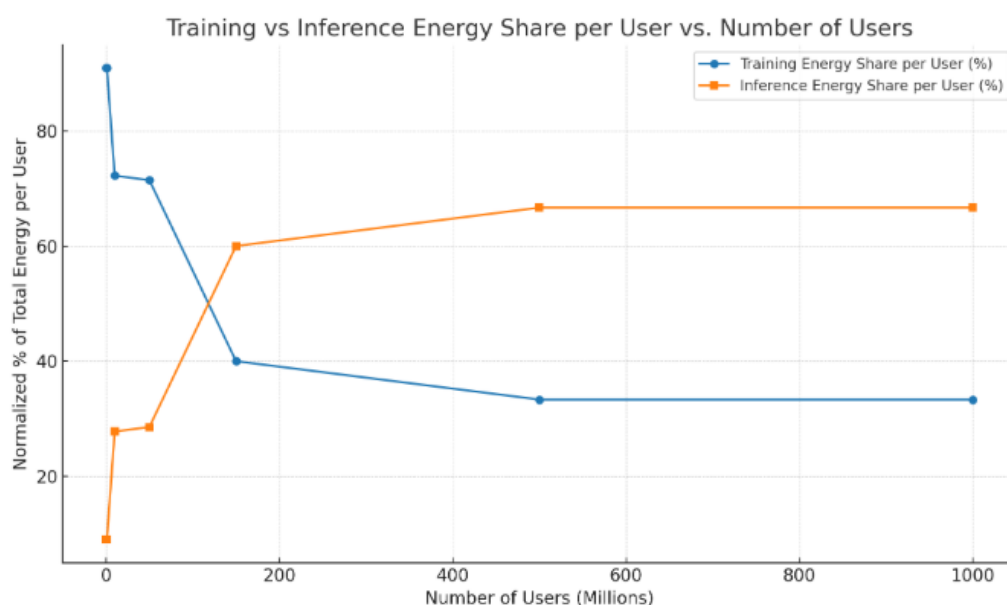
5.2.1. Inférences vs Apprentissage

Si de nombreuses études ont intégré les phases d’apprentissage dans les modèles d’estimation, les plus récentes ont tendance à ne prendre en compte que le poids de l’inférence. Elles partent du principe que le coût énergétique ou financier de l’apprentissage est rapidement amorti dès lors que le modèle est largement utilisé.³³⁶



³³⁶ Ce n’est pas le cas pour tous les modèles, en revanche, cette assertion reste globalement bonne si on cherche à modéliser le cas général et non le cas spécifique par modèle.

³³⁷ <https://www.ankursnewsletter.com/p/the-real-price-of-ai-pre-training>



L'analyse s'est donc focalisée sur l'évaluation de l'impact de la phase d'inférence.

Évolution du nombre d'inférences (AI Workload)

La consommation d'énergie par inférence varie fortement en fonction de la taille et du nombre de paramètres d'un modèle, mais aussi de la nature du traitement (simple question versus génération d'image).

Pour le travail prospectif, nous avons distingué deux types d'inférences (que nous appellerons *AI Workloads* pour éviter les confusions) et deux moyennes de consommation, sur la base des articles de Sasha Luccioni et notamment la synthèse de 2024 *"Power Hungry Processing: Watts Driving the Cost of AI Deployment?"*³³⁸

- "Gen AI" Cela correspond aux inférences de réponses à des "prompts" que l'on retrouve dans les diverses applications de type IA générative. **Consommation IT / AI workload de type Gen AI : 4 Wh**
- "Optimized Gen AI" Cela correspond à une consommation optimisée de l'IA Gen, **qui permettrait de répondre aux besoins d'IA de manière plus sobre, avec une algorithmie plus proche des IA générative optimisée de type Machine Learning, en moyenne bien moins gourmandes. Consommation IT / AI workload de type Machine Learning : 0,4 Wh**

Nous considérons que, pour notre année de référence, les centres de données dédiés à l'IA de "nouvelle génération" traitent principalement d'IA Générative, l'IA générative optimisée pouvant être déjà présente dans les autres catégories de centres de données.

Le nombre de "AI workloads" que l'on considère pour un usage français est donc de

$$\frac{(\text{Consommation_Totale_IA_Nouvelle_Génération}/\text{PUE})}{(\text{Consommation_IT_Workload_NouvelleGeneration soit } 3,3\text{E}+12 \text{ Wh} / 1,2) / 4\text{Wh}} = 6,8\text{E}+11$$

³³⁸ <https://arxiv.org/pdf/2311.16863>

Il sera possible dans certains scénarios de considérer que ces mêmes centres de données adaptent le type d'IA aux besoins.

5.2.2. Les tendances initiales

L'étude de Berkeley³³⁹ nous permet d'identifier des tendances sur l'évolution des consommations et des efficacités :

	2023	2024	2028	Gain annuel
Consommation des serveurs, 2 GPU et + (TWh)	40	60	250	44,27%
Average Power (kW)	3,5	4	5,1	7,82%
Efficacité (Wh/inference)	4	3	0,95	-25,00%
#inferences	1,00E+10	2,00E+10	2,63E+11	92,36%

À court terme, une forte croissance des usages de l'IA générative est anticipée pour tous les scénarios, avec un doublement des usages entre 2024 et 2025, ainsi qu'un gain en efficacité de 25% par an pour l'IA Gen jusqu'en 2030 et 5% par an pour l'IA Gen Optimisée (toutes dates) et l'IA Gen à partir de 2030. L'évolution annuelle des usages est par la suite fortement variable d'un scénario à l'autre.

Le tableau ci-dessous rend compte des principales hypothèses prises sur l'IA :

	En 2024	Gain d'efficacité par an
Consommation (Wh/Workload) - IA Gen	4,00	25% /an jusqu'en 2029 5% /an à partir de 2030
Consommation (Wh/Workload) - IA Gen optimisée	0,40	5% /an
Evolution annuelle des usages	+100% entre 2023 et 2024, au-delà, variable selon le scénario	NA

5.3. Blockchain

Parmi les nouveaux usages récents, outre l'IA, la blockchain fait elle aussi l'objet d'hypothèses spécifiques. L'enjeu des scénarios prospectifs pour la Blockchain est d'identifier les tendances en termes d'usages et de gains en efficacité.

Le site CBECI ³⁴⁰ nous permet de calculer ces tendances depuis 2020 :

Tendance	2020	2021	2022	2023	CAGR
Consommation (TWh)	67,14	89	96	121	21,99%
Efficacité (kWh/kH)	17500	15833	9722	8611	-19,27%
# khashs	3836571	5621053	9874286	14051613	51,12%
	42,41%	46,51%	75,67%	42,31%	51,12%

³³⁹ <https://escholarship.org/uc/item/32d6m0d1>

³⁴⁰ <https://ccaf.io/cbnsi/cbeci/ghg/methodology>

Faute de données localisées, nous utilisons ces tendances à l'échelle mondiale pour le périmètre France.

En synthèse les 3 paramètres clefs pour les scénarios prospectifs sont :

1. Une répartition des usages par type de technologie : 80% PoW vs 20% PoS
2. Une croissance annuelle des usages (en "kHash équivalent") autour de 50% pour un usage similaire à la tendance actuelle
3. Un gain annuel en efficacité (kWh/kHeq) de 20%

Le tableau ci-dessous résume les hypothèses de scénarisation prises, par rapport à ces 3 paramètres de tendances, pour les 5 scénarios :

Tendances 2020-2023		Répartition des usages par type de technologie : 80% PoW vs 20% PoS	Croissance annuelle des usages de 50%	Gain annuel en efficacité (kWh/kHeq) de 20%
Scénarios prospectifs 2024 - 2035	Tendanciel	80% PoW vs 20% PoS	30% /an entre 2024-2030, puis baisse régulière de croissance annuelle jusqu'à 10% /an de croissance en 2040, puis entre 9 et 8% /an par la suite*	19% /an jusqu'en 2030, puis 5% /an par la suite
	Génération Frugale	Idem Tendanciel	De 10% /an à 6% /an entre 2025 et 2029, puis 5% / an jusqu'en 2035 et 4% /an par la suite*	Idem Tendanciel
	Coopérations Territoriales	80% PoW vs 20% PoS en 2023-2025, puis inversion progressive jusqu'à 20% PoS vs 80% PoW en 2031	5% /an jusqu'en 2046, 4% par la suite*	Idem Tendanciel
	Technologies Vertes	80% PoW vs 20% PoS en 2023-2025, puis évolution progressive vers 60% PoW vs 40% PoS en 2029 et années suivantes	25% /an entre 2024-2030, puis 20% /an entre 2031 et 2033, 15% /an entre 2034 et 2037, 10% /an en 2038 et 2039 et 8% /an par la suite*	Idem Tendanciel
	Pari Réparateur	Idem Tendanciel	45% /an entre 2024-2031 puis baisse progressive de la croissance dont 25% /an en 2025, puis 8% /an à partir de 2051*	Idem Tendanciel

*Pour ces scénarios, nous avons fait l'hypothèse d'une croissance annuelle des usages de la blockchain systématiquement inférieure à la croissance annuelle moyenne constatée dans les tendances 2020-2023. Deux raisons principales justifient ce choix :

1. Volatilité des tendances récentes : la croissance moyenne annuelle observée sur la période 2020-2023 présente de fortes disparités d'une année à l'autre, avec un pic à +75,67 % en 2022 (cf. données du site CBECI). Il est difficile de prévoir si ce niveau de croissance se reproduira, ni à quel horizon.

2. Effet cumulatif de la croissance annuelle : chaque pourcentage de croissance s'ajoute au volume de l'année précédente, entraînant une augmentation exponentielle en valeur absolue. Pour éviter un effet d'emballement, nous avons donc formulé des hypothèses de ralentissement progressif de la croissance, en particulier pour les scénarios les plus dynamiques. **Ainsi, le scénario Tendanciel envisage une croissance annuelle de 40% entre 2024-2030, puis 30% /an entre 2031-2033, puis 20% /an entre 2033 et 2035.** De manière similaire, les scénarios Technologies vertes et Pari Réparateur intègrent également une trajectoire de ralentissement progressif de la croissance des usages de la blockchain.

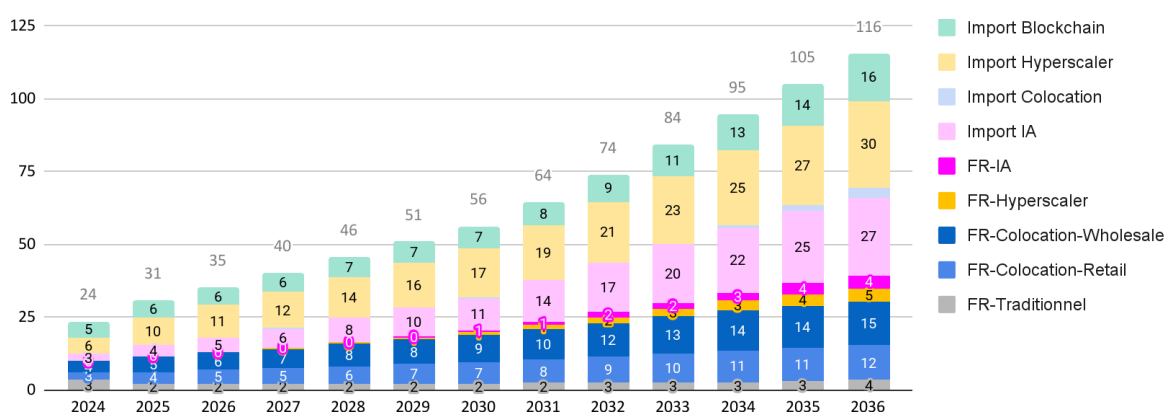
6. Résultats

6.1. Scénario tendanciel

6.1.1. Période 2024-2025

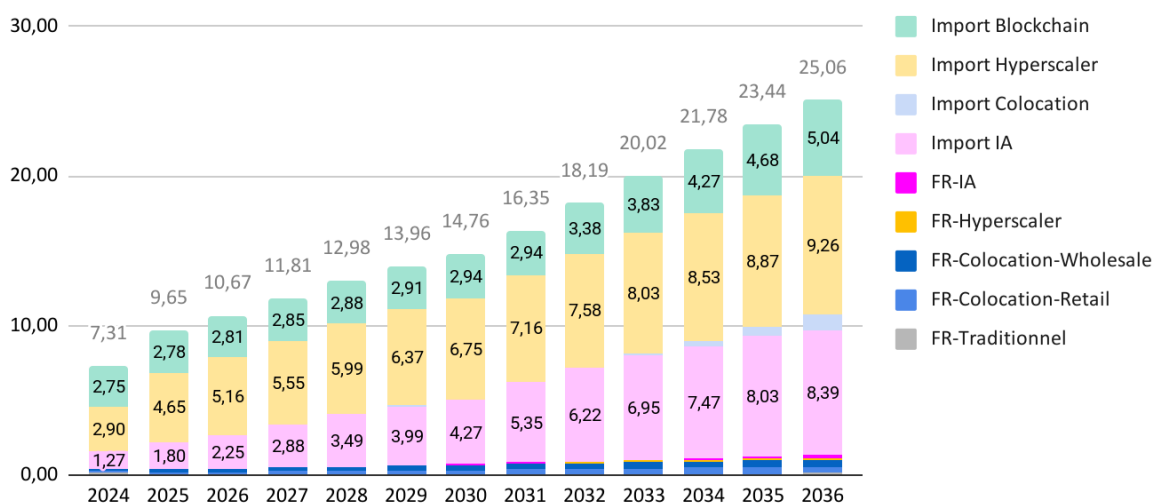
Le scénario Tendanciel aboutit à une **consommation électrique** totale de 37 TWh pour les centres de données en France en 2035, et de 105 TWh en incluant les consommations importées.

Scénario Tendanciel - Evolution des Consommations électriques des Centres de Données Usage Français (TWh)



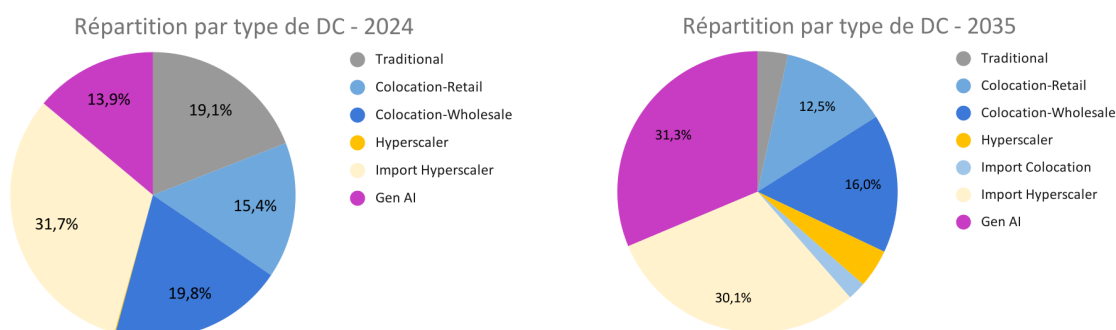
Le scénario Tendanciel aboutit à un total de 23,44 Mt CO₂eq. dont 1,25 Mt CO₂eq en France en 2035 pour **les émissions de gaz à effet de serre** liées aux usages des centres de données.

Scénario Tendanciel - Evolution Emissions GES Centres de Données Usage Français (MtCO₂eq)



6.1.2. Répartitions par type de centres de données en pourcentage des consommations

Notons ici la forte baisse des centres de données traditionnels proportionnellement dans la part des consommations électriques (3,5% en 2035), tandis que la part des centres de données dédiés à l'IA générative augmente très fortement, passant de presque 14% des consommations électriques à plus de 31%, soit plus qu'un doublement en proportion, mais une multiplication par plus de 11 en valeur absolue. D'autre part, nous observons le fort décollage des centres de données hyperscale encore inexistants en 2024 (4 TWh en 2035). La forte croissance des centres de données de colocation wholesale et retail en valeur absolue leur permet de se maintenir à un niveau de 28,5% des types de consommations électriques des centres de données.

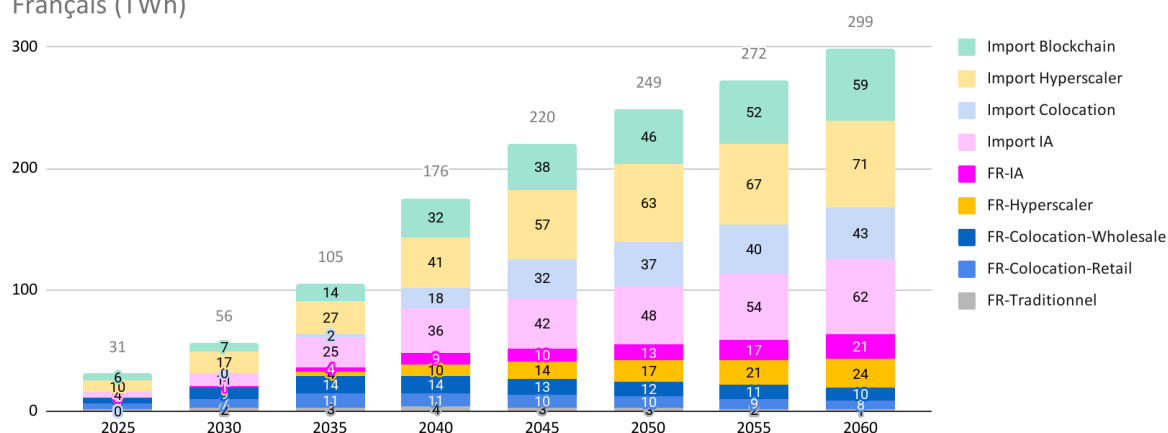


6.1.3. Modélisation post-2035

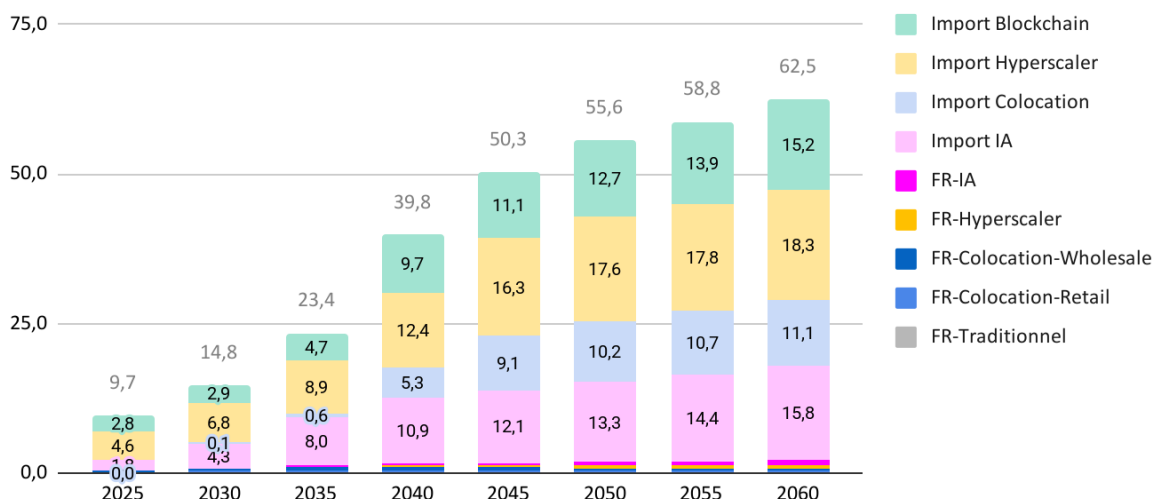
À l'horizon 2060, la consommation électrique liée aux usages numériques des Français pourrait atteindre 299 TWh, dont 64 TWh seraient imputables aux centres de données situés sur le territoire national. Dans le même temps, les émissions de gaz à effet de serre associées pourraient s'élever à 62,51 Mt CO₂eq, dont 2,13 Mt CO₂eq directement émises en France.

NB : A noter que cette prospective de très long terme est sujette à précautions, les prévisions sectorielles pour les centres de données s'arrêtant généralement à une perspective de 5 à 10 ans.

Scénario Tendanciel - Evolution des Consommations électriques des Centres de Données Usage Français (TWh)



Scénario Tendanciel - Evolution Emissions GES Centres de Données Usage Français (MtCO2eq)



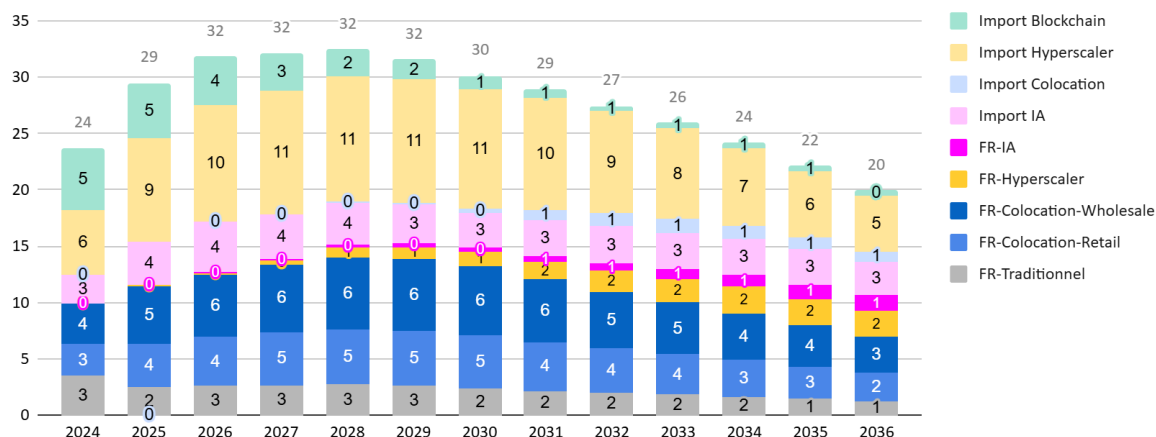
6.2. Scénario Génération Frugale

6.2.1. Période 2024-2025

Le scénario Génération Frugale aboutit à une **consommation électrique** totale de 12 TWh en France en 2035 et 22 TWh en incluant les consommations électriques importées.

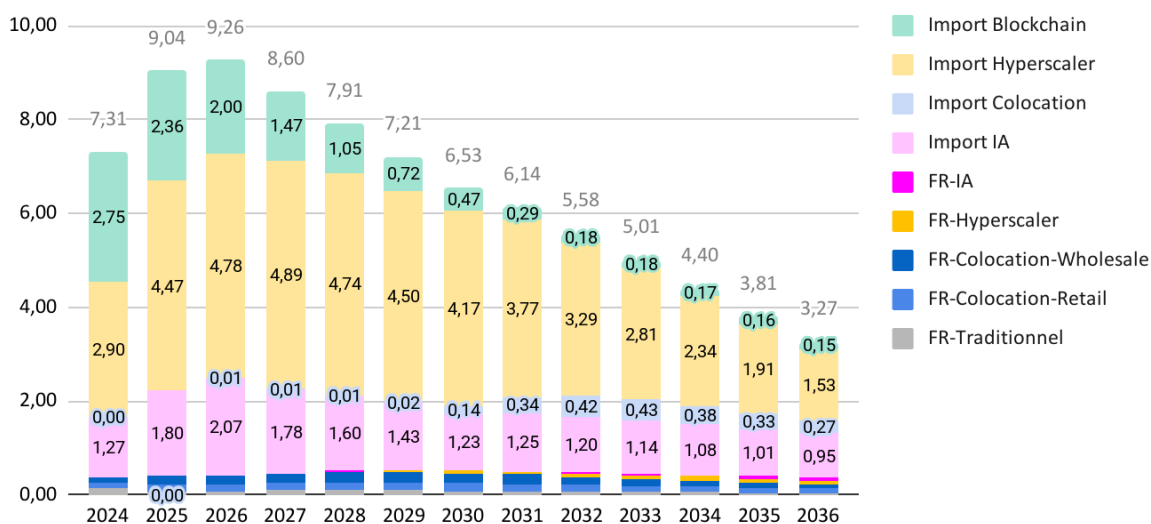
Dans ce scénario, les politiques de sobriété mises en place infléchissent progressivement les modes de vie des Français, qui réorientent leurs usages. L'effet de la mise en place d'un moratoire conduit à une construction rapide des centres de données dont la construction était déjà entamée, puis leur montée en puissance progressive (ramp up), tout en étant ralentie par la réorientation des usages. On constate ainsi un phénomène de dos rond entre 2024 et 2034, avant une baisse régulière et franche liée à la réorientation massive des modes de vie et d'usages du numérique. La part des consommations importées baisse plus rapidement que la consommation des centres de données présentes sur le sol français.

Scénario Génération Frugale - Evolution des Consommations électriques des Centres de Données Usage Français (TWh)



Le scénario Génération Frugale aboutit à un total de 3,81 Mt CO2eq dont 0,39 Mt CO2eq en France en 2035 d'émissions de gaz à effet de serre liées aux usages des centres de données.

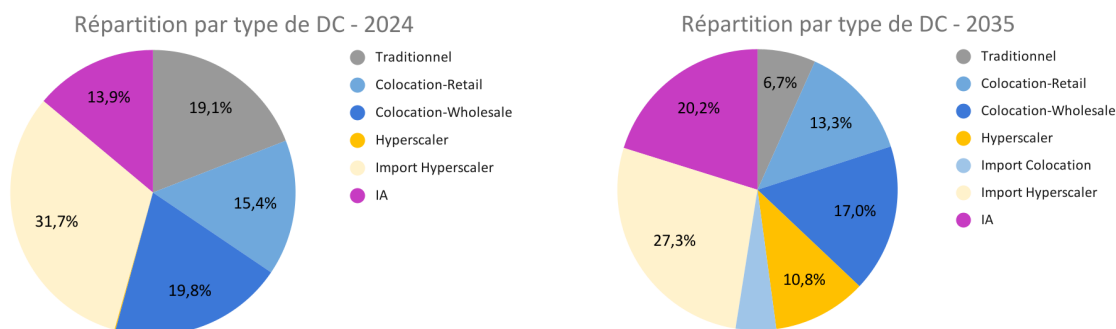
Scénario Génération Frugale - Evolution Emissions GES Centres de Données Usage Français (MtCO2eq)



6.2.2. Répartitions par type de centres de données en pourcentage des consommations

La part des centres de données traditionnels recule et celle des hyperscales présents sur le territoire français augmente (NB : ils partent de zéro). La croissance des parts des centres de données IA est contenue.

On note ici une relative stabilité de la proportion de consommation électrique allouée aux importations de données provenant des centres de données hyperscale, avec le développement de quelques centres de données hyperscale en France et la baisse des centres de données de colocation en France, retail comme wholesale.

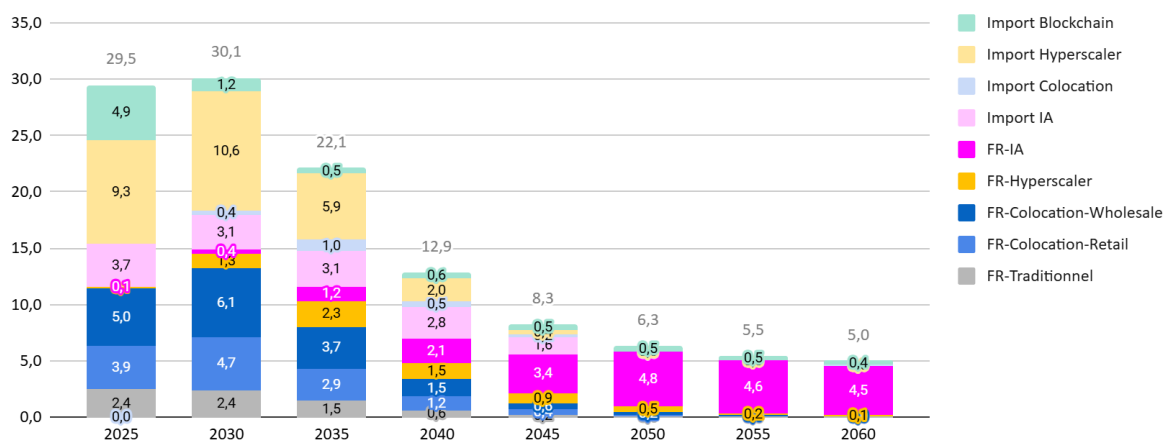


6.2.3. Modélisation post-2035

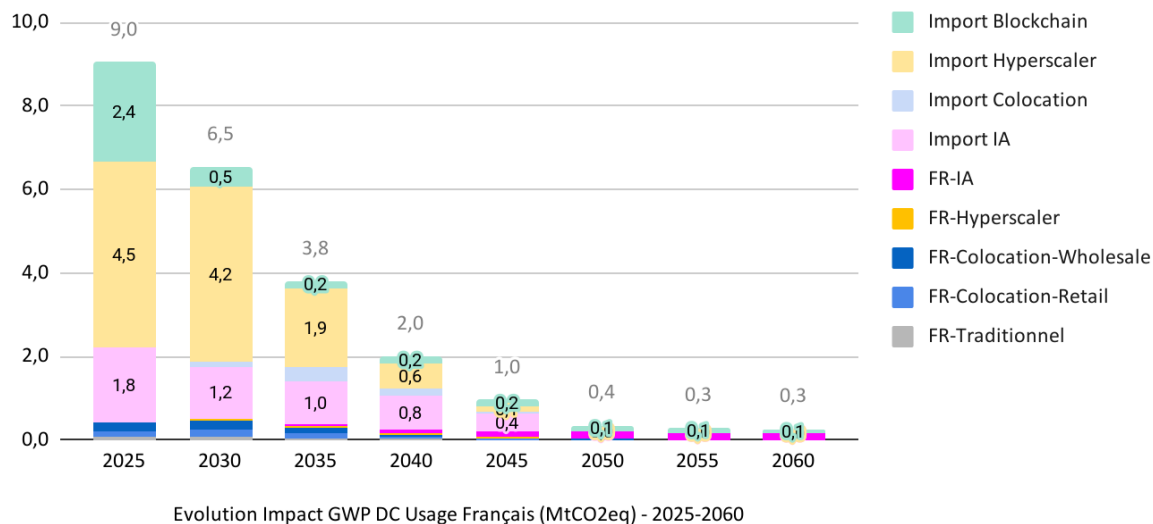
À long terme en 2060, la consommation électrique liée aux usages des Français pourrait baisser pour atteindre 5 TWh au total, dont 5 TWh liés aux centres de données présents sur le sol français ; tandis que les émissions de GES atteindraient 0,26 Mt CO₂eq au total dont 0,15 Mt CO₂eq pour les émissions depuis le sol français.

NB : A noter que cette prospective de très long terme est sujette à précautions, les prévisions sectorielles pour les centres de données s'arrêtant généralement à une perspective de 5 à 10 ans.

Scénario Génération Frugale - Evolution des Consommations électriques des Centres de Données Usage Français (TWh)



Scénario Génération Frugale - Evolution Emissions GES Centres de Données Usage Français (MtCO2eq)



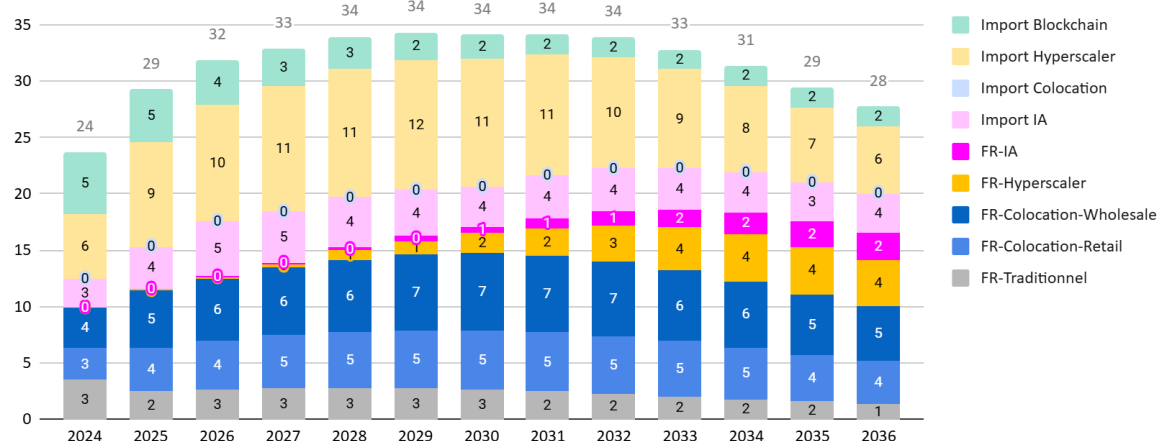
6.3. Scénario Coopérations Territoriales

6.3.1. Période 2024-2025

Le scénario Coopérations Territoriales aboutit à une **consommation électrique** totale de 18 TWh en France en 2035, et de 29 TWh en incluant les consommations importées.

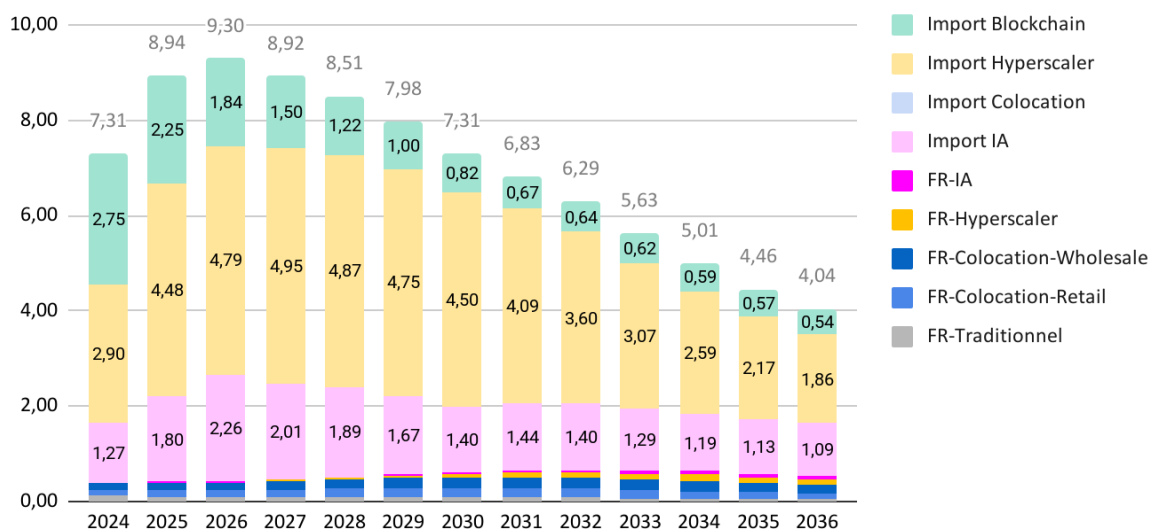
Dans ce scénario, les politiques mises en place freinent fortement la croissance des consommations électriques des centres de données en France et permettent de la contenir en dessous d'un seuil de 19 TWh (2033-2034), avant d'entamer une décrue progressive des consommations. Ainsi, en 2035, les consommations liées aux usages des Français des centres de données entame une dynamique à la baisse.

Scénario Coopérations Territoriales - Evolution des Consommations électriques des Centres de Données Usage Français (TWh)



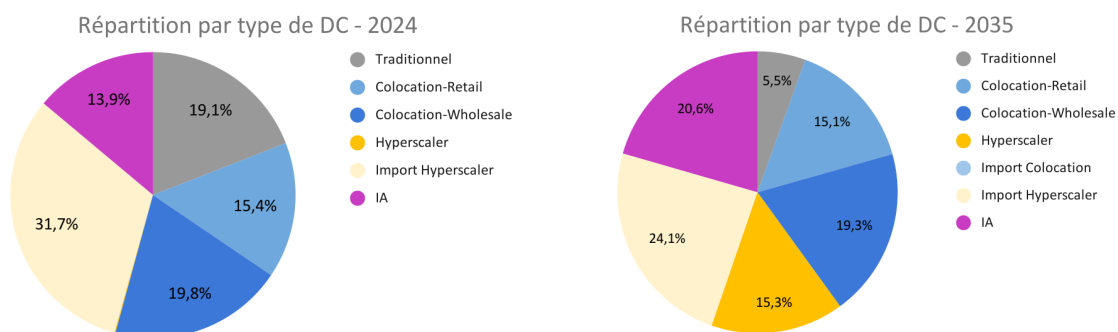
Le scénario Coopérations Territoriales aboutit à un total d'émissions de 4,46 Mt CO2eq, dont 0,59 Mt CO2eq en France en 2035 pour les **émissions de gaz à effet de serre** liées aux usages des centres de données.

Scénario Coopérations Territoriales - Evolution Emissions GES Centres de Données Usage Français (MtCO2eq)



6.3.2. Répartitions par type de centres de données en pourcentage des consommations

On note ici une forte diminution du nombre de centres de données traditionnels passant de 19,1% à 5,5% du total des consommations - alors que dans ce scénario, les consommations diminuent. On observe une croissance significative des centres de données hyperscales en France, entraînant proportionnellement un tassement des parts des centres de données de type colocation.

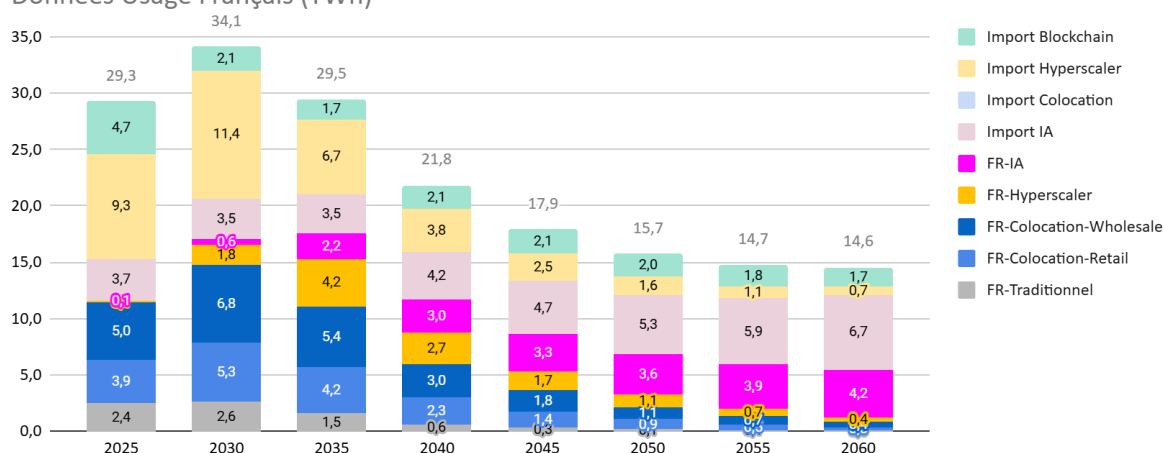


6.3.3. Modélisation post-2035

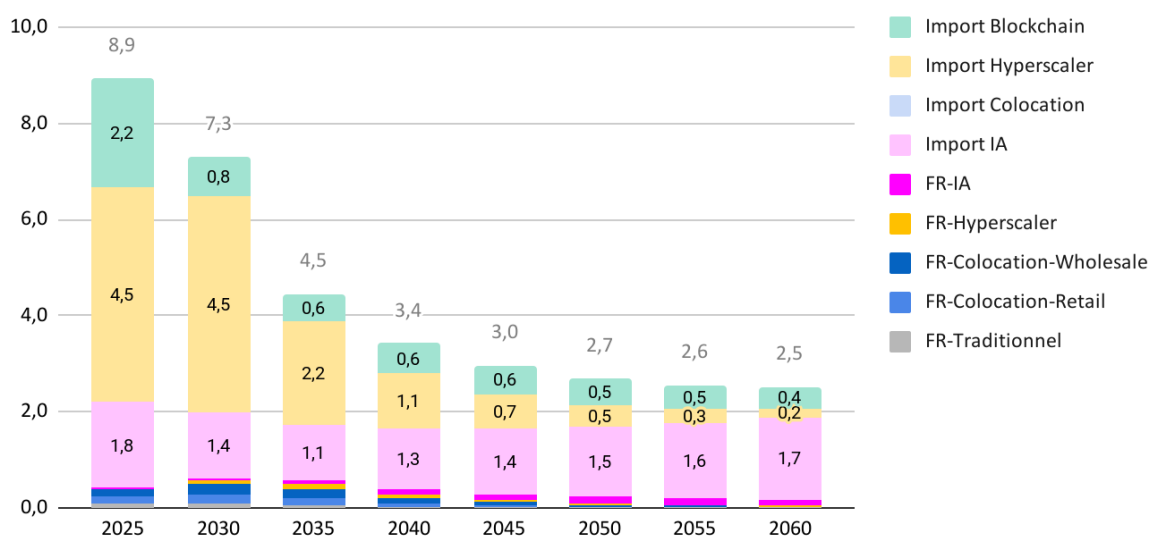
À long terme en 2060, la consommation électrique liée aux usages des Français des centres de données pourrait diminuer et atteindre 15 TWh au total, dont 5 liés aux centres de données présents sur le sol français ; tandis que les émissions de GES atteindraient 2,52 Mt CO2eq au total dont 0,18 Mt CO2eq pour les émissions sur le sol français.

NB : A noter que cette prospective de très long terme est sujette à précautions, les prévisions sectorielles pour les centres de données s'arrêtant généralement à une perspective de 5 à 10 ans.

Scénario Coopérations Territoriales - Evolution des Consommations électriques des Centres de Données Usage Français (TWh)



Scénario Coopérations Territoriales - Evolution Emissions GES Centres de Données Usage Français (MtCO2eq)



6.4. Scénario Technologies Vertes

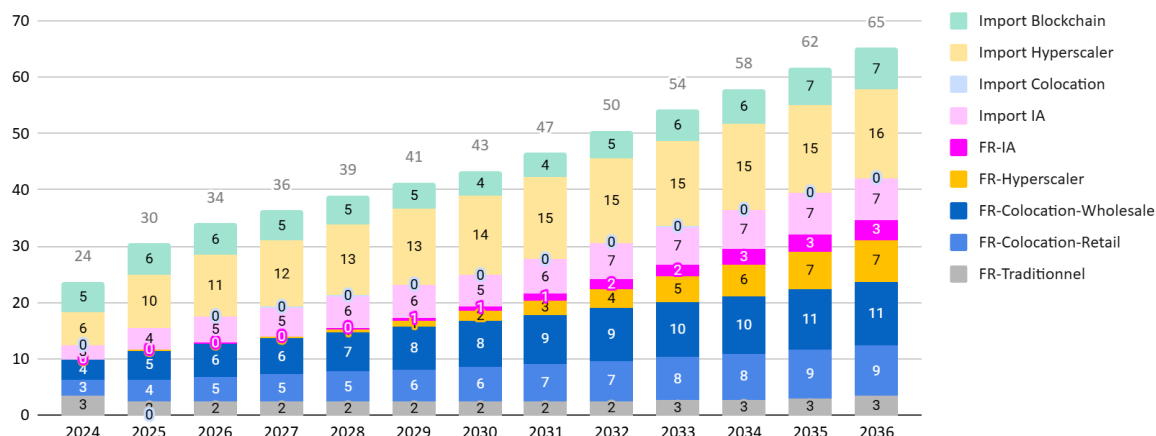
6.4.1. Période 2024-2025

Le scénario Technologies Vertes aboutit à une **consommation électrique** totale de 32 TWh en France en 2035 et 62 TWh en incluant les consommations importées.

Dans ce scénario, les politiques publiques interviennent avant tout pour inciter à l'innovation dans le domaine énergétique dans l'optique de minimiser l'impact du numérique sans en réduire les usages. On

constate donc une progression régulière des usages et de l'implantation des centres de données en France, notamment la concrétisation de certains projets présentés lors de l'AI Summit en février 2025.³⁴¹ L'augmentation des consommations électriques est tout d'abord portée par les centres de données spécialisés IA et les centres de données de type hyperscale (en particulier importés).³⁴² Concernant l'IA, ce scénario envisage que moins de la moitié des usages de l'IA par les Français serait traitée hors du territoire, notamment du fait de la politique Choose France. Par ailleurs, dans ce scénario, les usages de la blockchain suivent une croissance régulière qui ressort bien plus fortement qu'en TWh à l'observation des émissions de gaz à effet de serre en raison du mix énergétique mondial.

Scénario Technologies Vertes - Evolution des Consommations électriques des Centres de Données Usage Français (TWh)

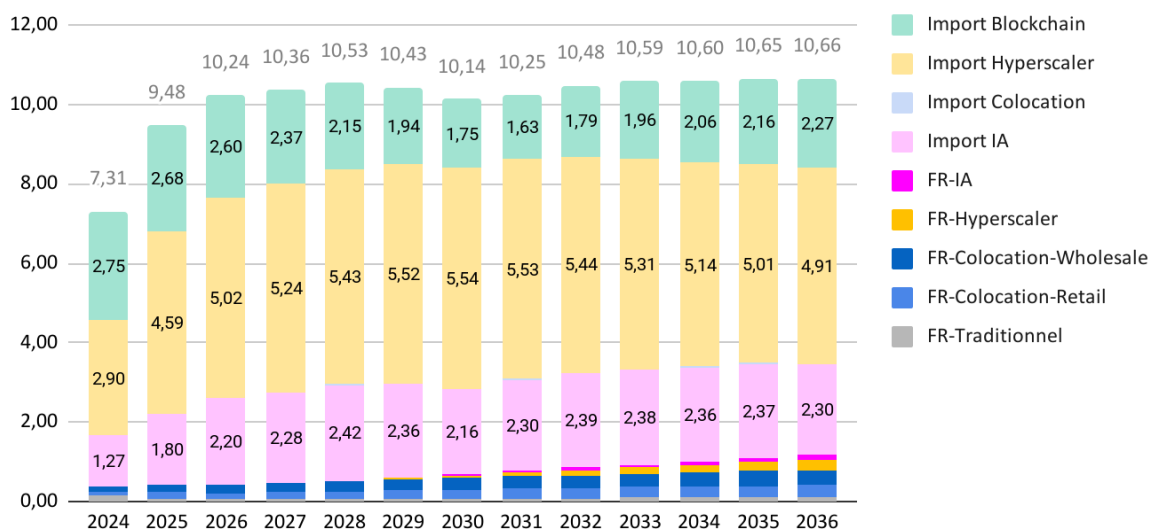


Le scénario Technologies Vertes aboutit à un total de 10,65 Mt CO₂eq dont 1,09 Mt CO₂eq en France en 2035 pour les **émissions de Gaz à Effet de Serre** liées aux usages des centres de données.

³⁴¹ <https://dcmag.fr/la-carte-regionale-des-35-sites-de-data-centers-dedies-a-lia-identifies-par-le-gouvernement/>

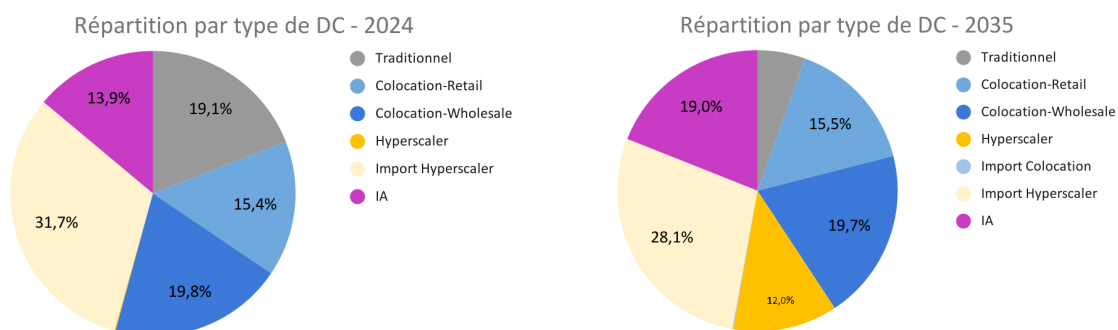
³⁴² Pour rappel, dans le modèle, pour permettre de distinguer la croissance de l'IA des autres types d'usages, nous avons distingué centres de données de type hyperscale et spécialisés IA. Dans la réalité, l'IA peut être hébergée dans des centres de données variés : hyperscale, colocation... Ici, afin de ne pas créer de doublon, nous avons poursuivi la distinction en considérant l'IA à part. De plus, nous faisons l'hypothèse que la croissance des usages de l'IA pourra avoir des effets stimulants sur la croissance d'autres usages hors IA, favorisant l'accroissement parallèlement des centres de données de colocation et hyperscale.

Scénario Technologies Vertes - Evolution Emissions GES Centres de Données Usage Français (MtCO2eq)



6.4.2. Répartitions par type de centres de données en pourcentage des consommations

Dans ce scénario Technologies Vertes, on note une forte croissance des parts de consommation électrique des centres de données IA, passant de 13,9% à 19%, ce qui représente une croissance d'autant plus importante que dans ce scénario, les consommations électriques des centres de données s'accroissent. La part des centres de données hyperscales en 2035 devient, elle aussi, structurante, avec plus de 39% du total des consommations électriques des centres de données répondant aux usages des Français provenant de centres de données hyperscales (dont 28,1% situés hors du territoire et 12% en France). La part des centres de données traditionnels se réduit proportionnellement de façon importante (et légèrement en valeur absolue). De même, la part de la colocation baisse légèrement, mais cet effet en proportion masque une croissance en valeur absolue importante de la colocation retail et wholesale, qui passe respectivement de 3 TWh et 4 TWh en 2024 à 9 TWh et 11 TWh en 2035.



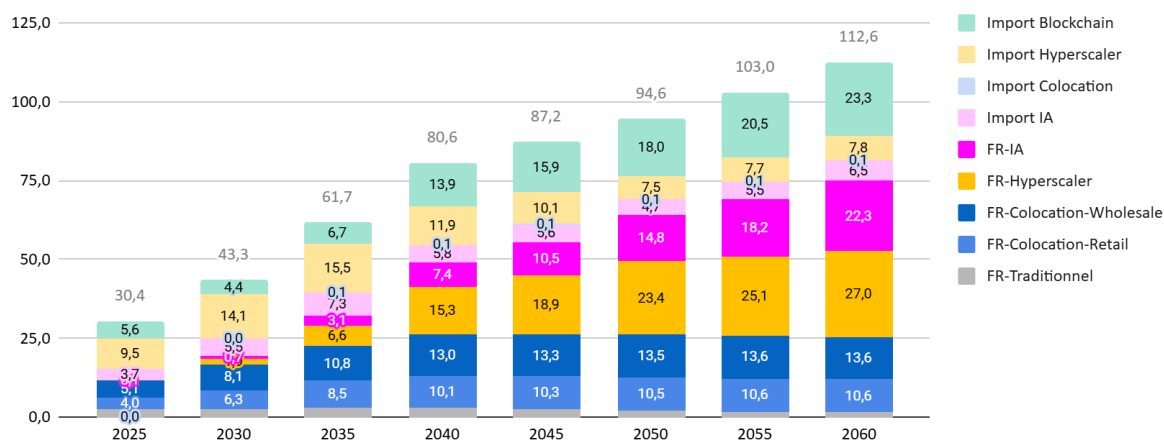
6.4.3. Modélisation post-2035

À long terme en 2060, la consommation électrique liée aux usages des Français pourrait atteindre 113 TWh au total, dont 75 TWh liés aux centres de données présents sur le sol français ; tandis que les

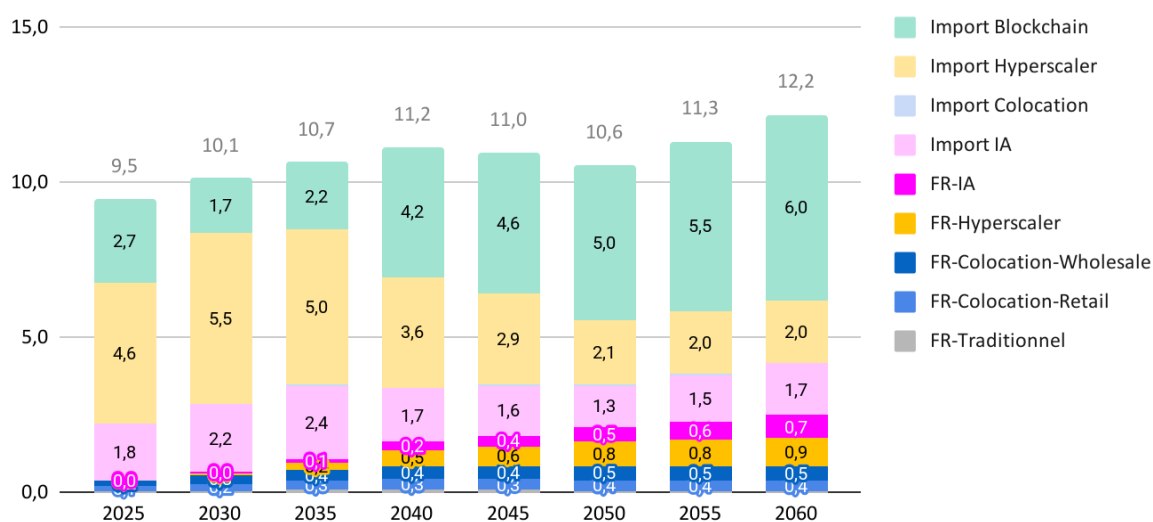
émissions carbone atteindraient 12,20 Mt CO₂eq au total dont 2,51 Mt CO₂eq pour les émissions sur le sol français.

NB : A noter que cette prospective de très long terme est sujette à précautions, les prévisions sectorielles pour les centres de données s'arrêtant généralement à une perspective de 5 à 10 ans.

Scénario Technologies Vertes - Evolution des Consommations électriques des Centres de Données Usage Français (TWh)



Scénario Technologies Vertes - Evolution Emissions GES Centres de Données Usage Français (MtCO₂eq)



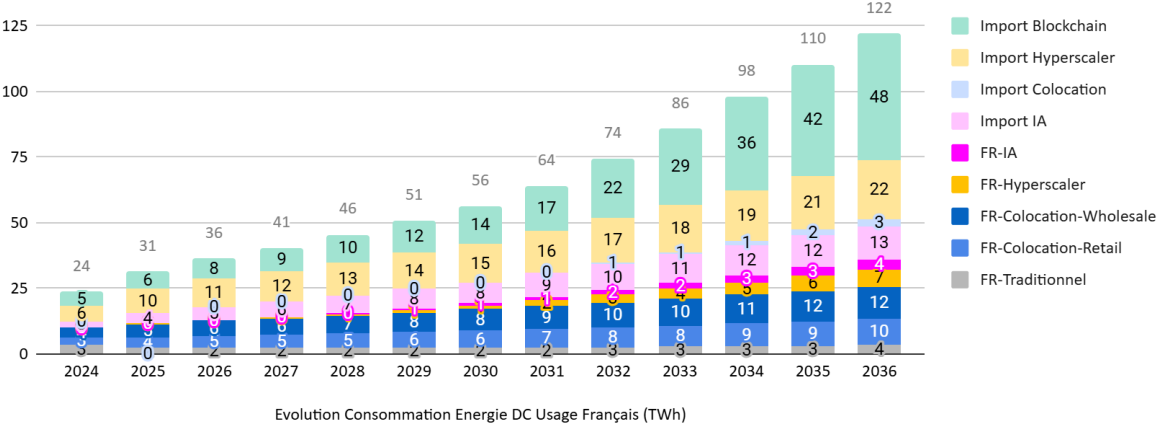
6.5. Scénario Pari Réparateur

6.5.1. Période 2024-2025

Le scénario Pari Réparateur aboutit à une **consommation électrique** totale de 33 TWh en France en 2035 et 110 TWh en incluant les consommations importées.

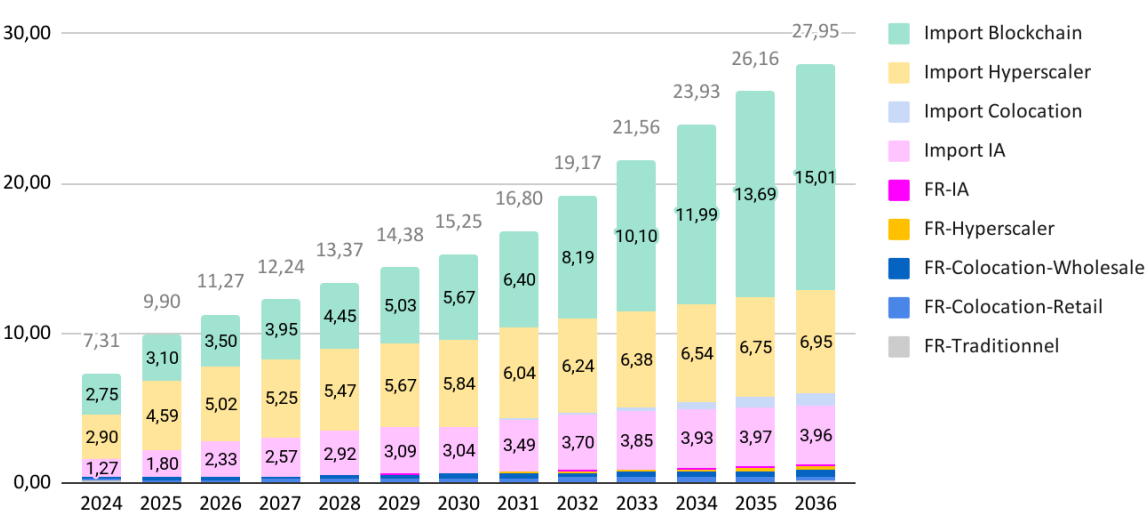
Dans ce scénario, les politiques publiques interviennent uniquement pour inciter à l'innovation technologique dans le but de monitorer et compenser les émissions de GES des autres secteurs d'activité. On constate une progression importante des usages et de l'implantation des centres de données en France, notamment la concrétisation de certains projets présentés lors de l'AI Summit en février 2025.³⁴³ L'augmentation des consommations électriques est tout d'abord portée par les centres de données spécialisés IA, encore plus fortement que dans le scénario précédent, Technologies Vertes, et les centres de données de type hyperscale (en particulier importés).³⁴⁴ Par ailleurs, dans ce scénario, les usages de la blockchain suivent une croissance très forte, observable nettement à la fois en termes de consommations électriques et en termes d'émissions de gaz à effet de serre.

Scénario Pari Réparateur - Evolution des Consommations électriques des Centres de Données Usage Français (TWh)



Le scénario Pari Réparateur aboutit à un total de 26,16 Mt CO2eq dont 1,12 Mt CO2eq en France en 2035 pour les **émissions de Gaz à Effet de Serre** liées aux usages des centres de données.

Scénario Pari Réparateur - Evolution Emissions GES Centres de Données Usage Français (MtCO2eq)

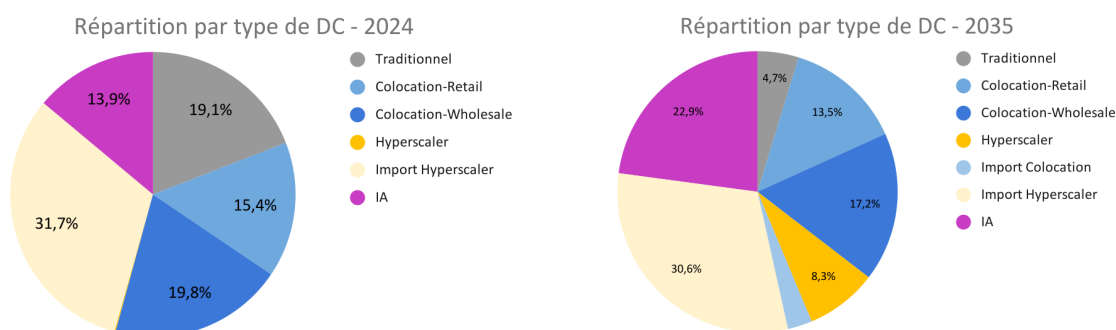


³⁴³ <https://dcmag.fr/la-carte-regionale-des-35-sites-de-data-centers-dedies-a-lia-identifies-par-le-gouvernement/>

³⁴⁴ Pour rappel, dans le modèle, pour permettre de distinguer la croissance de l'IA des autres types d'usages, nous avons distingué centres de données de type hyperscale et spécialisés IA. Dans la réalité, l'IA peut être hébergée dans des centres de données variés : hyperscale, colocation... Ici, afin de ne pas créer de doublon, nous avons poursuivi la distinction en considérant l'IA à part. De plus, nous faisons l'hypothèse que la croissance des usages de l'IA pourra avoir des effets stimulants sur la croissance d'autres usages hors IA, favorisant l'accroissement parallèlement des centres de données de colocation et hyperscale.

6.5.2. Répartitions par type de centres de données en pourcentage des consommations

On observe une très forte progression de la part des centres de données dédiés à l'IA et des importations liées aux centres de données hyperscales. On note l'apparition de besoins d'importations liés aux centres de données de colocation. En France, la part des centres de données de colocation, qu'ils soient retail ou wholesale, semble proportionnellement en recul, bien qu'en valeur absolue la colocation, retail et wholesale, augmentent respectivement de 6 TWh et de 8 TWh entre 2024 et 2035. Par ailleurs, les centres de données traditionnels enregistrent une baisse très significative dans la répartition, alors que leurs usages connaissent un tassement progressif en valeur absolue.

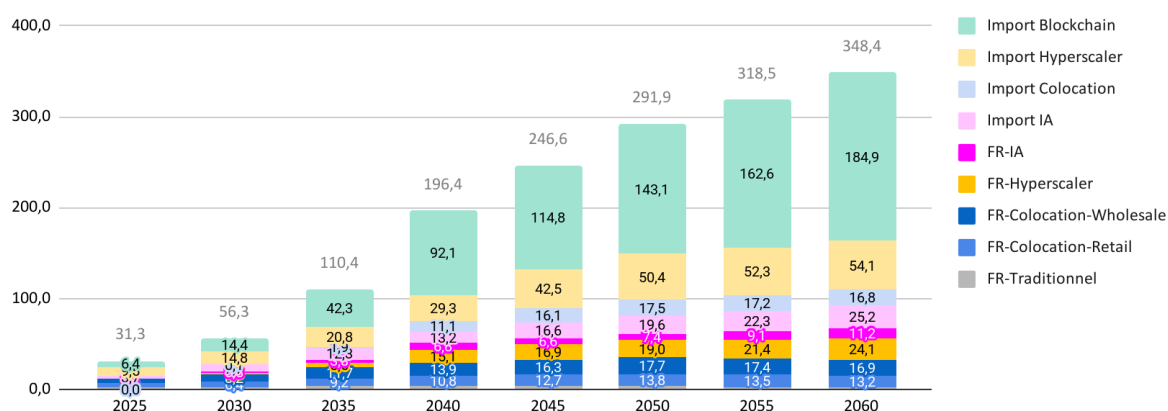


6.5.3. Modélisation post-2035

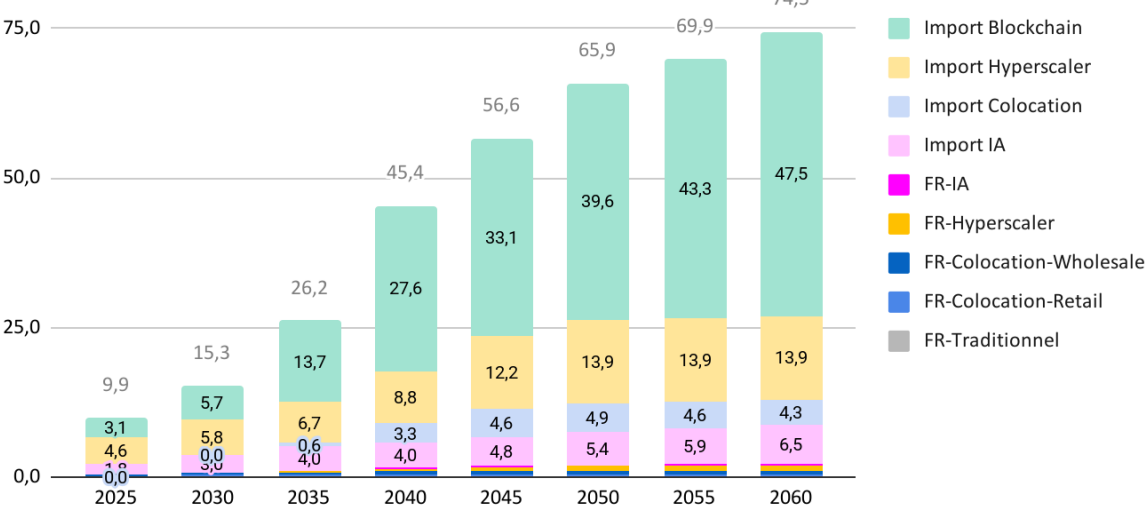
A long terme en 2060, la consommation électrique liée aux usages des Français pourrait atteindre 348 TWh au total dont 67 TWh liés aux centres de données présents sur le sol français ; tandis que les émissions de gaz à effet de serre atteindraient 74,46 Mt CO₂eq au total dont 2,26 Mt CO₂eq pour les émissions sur le sol français.

NB : A noter que cette prospective de très long terme est sujette à précautions, les prévisions sectorielles pour les centres de données s'arrêtant généralement à une perspective de 5 à 10 ans.

Scénario Pari Réparateur - Evolution des Consommations électriques des Centres de Données Usage Français (TWh)



Scénario Pari Réparateur - Evolution Emissions GES Centres de Données Usage Français (MtCO2eq)



6.6. Comparaison des scénarios

Les deux tableaux ci-dessous permettent de comparer entre chaque scénario, leurs consommations (en TWh) et leurs émissions de GES (Mt CO2eq.) respectives.

Long terme - Évolution de la consommation électrique pour l'usage français (TWh / an)

	2024	2030	2035	2040	2045	2050	2055	2060
Tendanciel - en France	9,91	20,75	36,73	48,04	51,61	54,73	59,03	63,58
Tendanciel - import	13,79	35,61	68,53	127,66	168,68	194,26	213,30	234,97
Tendanciel - Total	23,71	56,36	105,25	175,70	220,29	248,99	272,33	298,55
Génération Frugale - en France	9,91	14,87	11,58	6,95	5,58	5,77	4,99	4,59
Génération Frugale - import	13,79	15,26	10,54	5,91	2,69	0,57	0,48	0,42
Génération Frugale - Total	23,71	30,13	22,12	12,86	8,27	6,34	5,47	5,01
Coopérations Territoriales - en France	9,91	17,10	17,54	11,69	8,58	6,81	5,88	5,46
Coopérations Territoriales - import	13,79	17,04	11,92	10,14	9,32	8,90	8,86	9,10
Coopérations Territoriales - Total	23,71	34,14	29,46	21,83	17,89	15,72	14,74	14,56
Technologies Vertes - en France	9,91	19,34	32,14	48,85	55,50	64,19	69,14	74,89
Technologies Vertes - import	13,79	24,01	29,53	31,77	31,67	30,39	33,85	37,70
Technologies Vertes - Total	23,71	43,34	61,67	80,62	87,17	94,58	102,99	112,59
Pari Réparateur - en France	9,91	19,28	33,04	50,65	56,47	61,31	64,07	67,46
Pari Réparateur - import	13,79	37,00	77,31	145,79	190,12	230,57	254,40	280,95
Pari Réparateur - Total	23,71	56,28	110,35	196,44	246,59	291,88	318,48	348,41

Long terme - Évolution des émissions de gaz à effet de serre (GES) pour l'usage français (Mega Tonnes CO2eq / an)

	2024	2030	2035	2040	2045	2050	2055	2060
Tendanciel - en France	0,38	0,73	1,25	1,61	1,73	1,83	1,98	2,13
Tendanciel - import	6,93	14,03	22,19	38,23	48,57	48,36	56,80	60,38
Tendanciel - Total	7,31	14,76	23,44	39,84	50,30	55,63	58,78	62,51
Génération Frugale - en France	0,38	0,52	0,39	0,23	0,19	0,19	0,17	0,15
Génération Frugale - import	6,93	6,01	3,41	1,77	0,77	0,16	0,13	0,11
Génération Frugale - Total	7,31	6,53	3,81	2,00	0,96	0,35	0,29	0,26
Coopérations Territoriales - in France	0,38	0,60	0,59	0,39	0,29	0,23	0,20	0,18
Coopérations Territoriales - import	6,93	6,71	3,86	3,04	2,68	2,47	2,36	2,34
Coopérations Territoriales - Total	7,31	7,31	4,46	3,43	2,97	2,69	2,56	2,52
Technologies Vertes - en France	0,38	0,68	1,09	1,64	1,86	2,15	2,32	2,51
Technologies Vertes - import	6,93	9,46	9,56	9,51	9,12	8,41	9,01	9,69
Technologies Vertes - Total	7,31	10,14	10,65	11,15	10,98	10,56	11,33	12,20
Pari Réparateur - en France	0,38	0,68	1,12	1,70	1,89	2,05	2,15	2,26
Pari Réparateur - import	6,93	14,58	25,04	43,65	54,74	63,85	67,75	72,20
Pari Réparateur - Total	7,31	15,25	26,16	45,35	56,64	65,90	69,89	74,46

À titre de comparaison, nous avons indiqué pour chacun des scénarios l'évolution que ce scénario subit en 11 ans, entre 2024 et 2035.

	Tendanciel	S1 Génération Frugale	S2 Coopérations Territoriales	S3 Technologies Vertes	S4 Pari Réparateur
Consommation (en France) en TWh en 2035	37	12	18	32	33
<i>Evolution par rapport à 2024</i>	x3,7	+17%	+77%	x3,2	x3,3
Consommation totale TWh en 2035*	105	22	29	62	196
<i>Evolution par rapport à 2024</i>	x4,4	-7%	+24%	x2,6	x4,7
Consommation totale TWh en 2060*	299	5	15	113	348
Émissions en Mt CO ₂ eq en 2035	23,44	3,81	4,46	10,65	26,16
<i>Evolution par rapport à 2024</i>	x3,2	-48%	-39%	x1,5	x3,6
Émissions en Mt CO ₂ eq en 2060	62,51	0,26	2,52	12,20	74,46

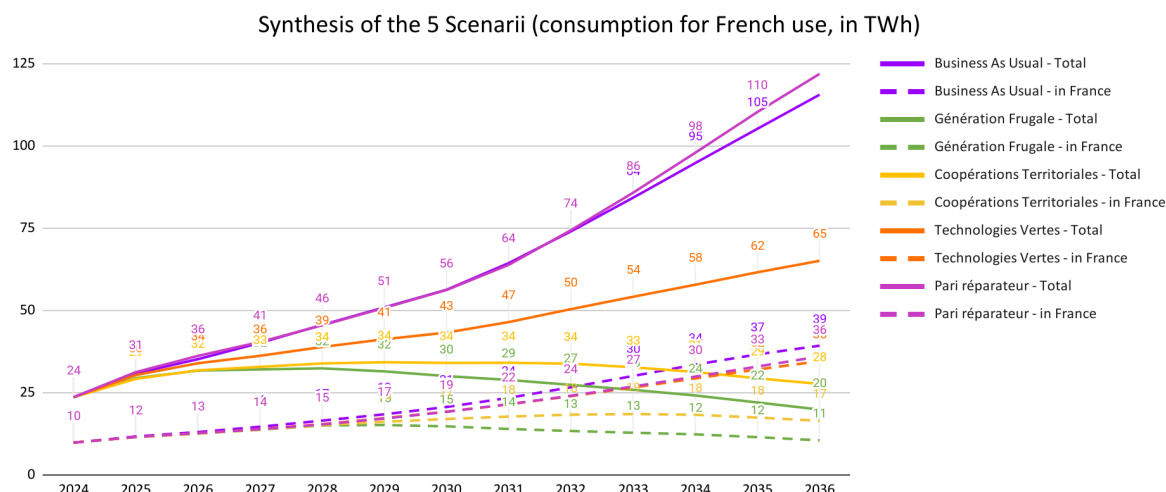
*La consommation électrique totale comprend la consommation électrique des centres de données répondant aux usages des Français, que cette consommation ait lieu sur le territoire ou à l'étranger (import).

On constate de très fortes divergences de perspectives d'évolution des consommations électriques et d'émissions de GES entre ces 5 scénarios.

Ainsi, le scénario Génération Frugale (S1) **permettrait une baisse potentielle des consommations électriques de -7% entre 2024 et 2035**, ainsi qu'une diminution potentielle des émissions de GES de -48%. À l'autre bout du spectre, le scénario Pari Réparateur atteindrait une multiplication par 4,7 de ses consommations électriques entre 2024 et 2035, pour une multiplication potentielle par 3,6 des émissions de GES.

6.6.1. Comparaison des consommations électriques 2024-2035

Le graphique ci-dessous synthétise les consommations comparées des 5 scénarios entre eux :



On constate ainsi une différence particulièrement claire entre deux types de scénarios :

1. Les scénarios permettant une réduction des consommations énergétiques des centres de données (Génération Frugale, Coopérations Territoriales).
2. Les scénarios qui rendent compte d'une augmentation importante des consommations énergétiques des centres de données (Tendancier, Technologies Vertes, Pari Réparateur).

Ce constat est valable à l'échelle des consommations des centres de données présents sur le sol français d'une part, et aussi, à une échelle plus importante encore, en incluant les émissions importées pour couvrir les usages des Français.

Parmi les scénarios étudiés, seuls ceux intégrant une décroissance des usages permettent une réduction de la consommation énergétique par rapport à la situation actuelle.

De manière générale, le principal facteur déterminant des consommations dans l'ensemble des scénarios est l'évolution des usages numériques, exprimés en workloads, et en workloads liées à l'IA. Plus ces usages sont élevés, plus la consommation d'énergie augmente.

La comparaison entre les scénarios Pari Réparateur et Technologies Vertes met en évidence deux éléments clés : le scénario Pari Réparateur est celui qui affiche la plus forte consommation énergétique globale, avec environ 39 TWh de plus que le scénario Technologies Vertes, pour répondre aux besoins numériques des Français. Toutefois, sur le sol français, ces deux scénarios génèrent une consommation très proche avant 2035 (un léger écart entre ces deux scénarios se creuse par la suite avec une consommation d'environ 7 TWh de plus pour le scénario Technologies Vertes en 2060 - ce qui est cohérent, ce dernier encourageant la relocalisation des centres de données en France pour tirer parti de la faible intensité carbone du mix électrique national).

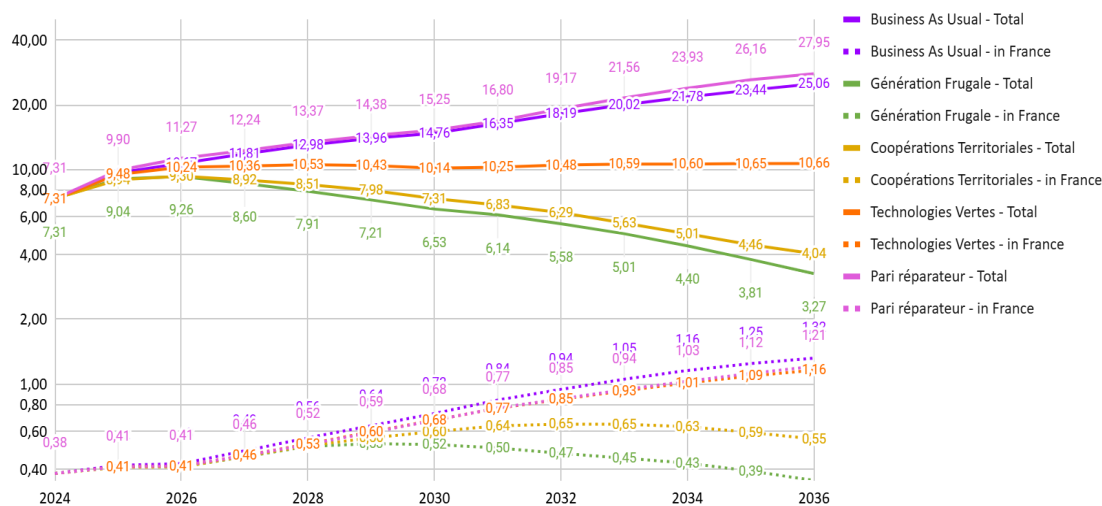
Bien que les dynamiques des scénarios Génération Frugale et Coopérations Territoriales convergent toutes les deux vers un plafond de consommations avant une décline de ces consommations, on constate que le scénario Génération Frugale permettrait une réduction plus importante des consommations électriques, à la fois en importé et en France (6 à 7 TWh de moins que pour le scénario Coopérations Territoriales pour l'année 2035, jusqu'à 9 TWh d'écart en 2060 entre les deux scénarios).

6.6.2. Comparaison des émissions de GES 2024-2035

Le graphique ci-dessous synthétise les émissions de GES comparées des 5 scénarios entre eux.

À noter que pour des besoins de lisibilité, l'axe vertical du graphique est exprimé sur une échelle logarithmique.

Synthesis of the 5 Scenarii (GWP for French use, in MtCO₂eq.)



L'écart entre les consommations en France et importées est encore plus sensible lorsqu'on observe la comparaison entre les émissions des 5 scénarios. Cette comparaison des émissions tient compte des différences de mix électrique entre la France, l'Europe, les Etats-Unis et la Chine (voir à ce titre le chapitre méthodologique correspondant).

Les scénarios Pari Réparateur et Tendanciel suivent ainsi une courbe croissante très proche, malgré des hypothèses un peu différentes. A partir de 2030, le léger écart de croissance des émissions se creuse entre ces deux scénarios et le scénario Pari Réparateur émet par la suite plus que le scénario Tendanciel au global (émissions en France et importées). Ceci est cohérent avec les hypothèses prises dans ce scénario Pari Réparateur où les centres de données et le numérique dans son ensemble sont développés prioritairement par rapport à d'autres secteurs pour favoriser et accélérer leur décarbonation.

Le scénario Technologies Vertes, figuré en orange sur le graphique, connaît une croissance de ses émissions globales avant 2026, puis atteint une zone de stabilité où l'équilibre entre les usages des Français depuis des centres de données en France et ceux importés permet de ne pas dépasser 11 Mt CO₂eq. d'émissions totales entre 2026 et 2036.

Par ailleurs, les scénarios Génération Frugale et Coopérations Territoriales sont les seuls scénarios permettant une réduction des émissions de GES par rapport à aujourd'hui. Ces deux scénarios permettent à la fois une réduction des émissions de GES émises par les centres de données depuis le territoire français (lignes en pointillés) mais également au global en tenant compte des émissions importées.

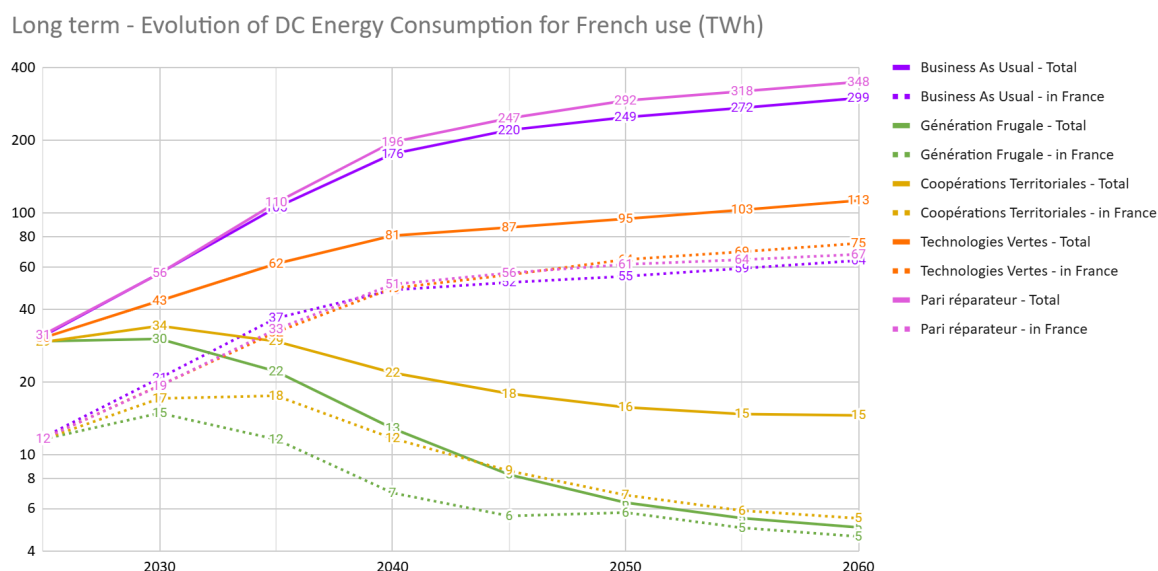
6.6.3. Comparaison des consommations électriques jusqu'à 2060

Le graphique ci-dessous fournit à titre indicatif, une comparaison de l'évolution des consommations électriques des 5 scénarios modélisés, pour la période 2024-2060.

NB : A noter que cette prospective de très long terme est sujette à précautions, les prévisions sectorielles pour les centres de données s'arrêtant généralement à une perspective de 5 à 10 ans.

Dans le modèle, nous avons poursuivi les tendances observées pour chacun des scénarios sur la période pré-2035. Pour ces raisons, nous ne fournissons pas d'analyse détaillée de ces résultats.

Pour des besoins de lisibilité, l'axe vertical du graphique est exprimé sur une échelle logarithmique.



Sur cette prospective long terme, la persistance post-20235 des tendances observées avant 2035 aboutit à plusieurs constats :

D'une part, on observe un dépassement des consommations électriques du scénario Pari réparateur par rapport au scénario Tendanciel. Bien que les hypothèses prises au sein de ces deux scénarios soient différentes, on observe toutefois une tendance qui, sur le graphique, peut sembler être sur une dynamique assez proche. C'est toutefois un effet visuel global qui masque les disparités entre les hypothèses prises sur ces deux scénarios, principalement concernant l'usage de la blockchain (hypothèse d'usage beaucoup plus forte dans le scénario Pari Réparateur que dans le scénario Tendanciel), et l'usage de l'IA (hypothèse inverse, plus forte dans le scénario Tendanciel que dans le scénario Pari Réparateur). Ce dépassement est valable au niveau des consommations électriques totales (en France et import pour couvrir les usages des Français), mais aussi au niveau des consommations électriques des centres de données sur le sol français, principalement du fait des hypothèses prises concernant une politique favorisant l'implantation des centres de données sur le sol européen et particulièrement en France. Cette politique favorisant l'implantation des centres de données sur le sol français est encore plus marquée dans le scénario Technologies Vertes, qui, à l'échelle des consommations électriques des centres de données en France, dépasse les scénarios Tendanciel et Pari Réparateur.

Inversement, le scénario Technologies Vertes permet de limiter la croissance des consommations électriques des centres de données importés pour répondre aux usages des Français, qui correspondent à entre 40% et 30% des consommations électriques totales pour répondre aux usages des Français, selon les années, soit entre 30 et 38 TWh.

Cependant, l'ensemble de ces trois scénarios, Tendanciel, Pari Réparateur et Technologies Vertes, tablent tous les trois sur une consommation de plus de 50 TWh en 2050 sur le sol français (respectivement, 55 TWh, 61 TWh, 64 TWh), ce qui correspondrait à **multiplier par entre 4,5 et 5,5 la consommation électrique actuelle des centres de données en France en 25 ans**. Or, si une consommation électrique pour le secteur des centres de données aux alentours de 40 TWh en 2050 reste soutenable, pour ne pas contraindre le développement des autres leviers d'électrification, **un dépassement de ce seuil de consommation contraindrait à prendre des mesures additionnelles pour réduire la consommation d'électricité dans d'autres secteurs dans le cadre de la Stratégie Nationale Bas Carbone**. Ce risque oblige ainsi à garder un œil particulièrement attentif à ces évolutions.

Concernant les scénarios Coopérations Territoriales et Génération Frugale, l'accent mis sur la sobriété produit une baisse importante de la consommation électrique des centres de données, année après année, tant au niveau des consommations électriques des centres de données en France qu'au niveau des consommations importées. Le scénario Coopérations Territorial permet de maintenir aussi à un niveau relativement faible les consommations importées, puisqu'après 2035, celles-ci se maintiennent à 10 TWh ou moins. Ce phénomène est encore plus prononcé pour le scénario Génération Frugale, pour lequel les consommations importées deviennent résiduelles, avec moins de 3 TWh en 2045 et par la suite.

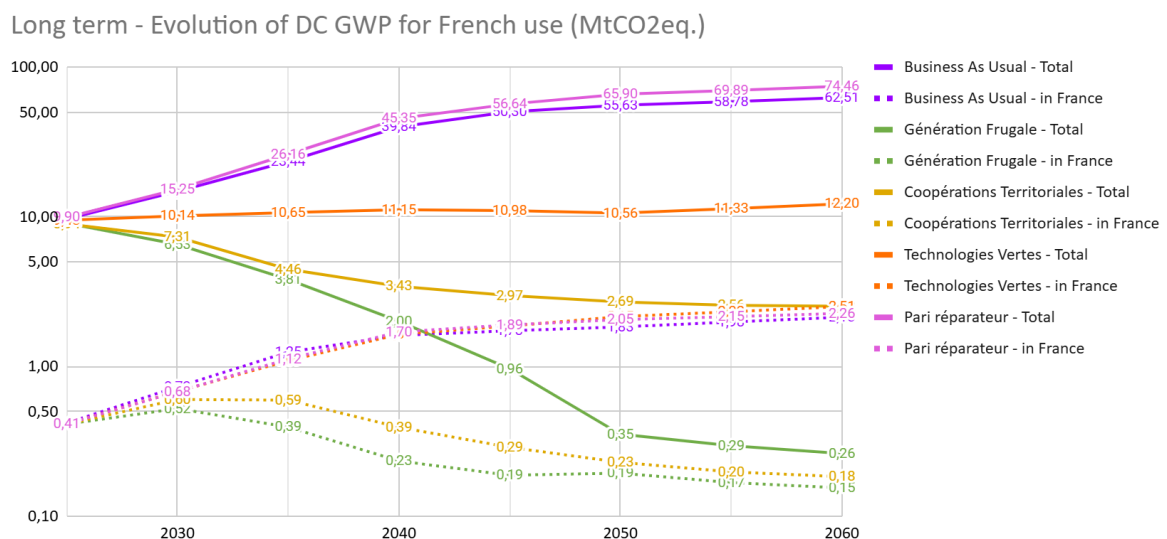
6.6.4. Comparaison des émissions de GES jusqu'à 2060

Le graphique ci-dessous fournit, à titre indicatif, une comparaison des émissions de GES des 5 scénarios modélisés, pour la période 2024-2060.

NB : A noter que cette prospective de très long terme est sujette à précautions, les prévisions sectorielles pour les centres de données s'arrêtant généralement à une perspective de 5 à 10 ans.

Dans le modèle, nous avons poursuivi les tendances observées pour chacun des scénarios sur la période pré-2035. Pour ces raisons, nous ne fournissons pas d'analyse détaillée de ces résultats.

Pour des besoins de lisibilité, l'axe vertical du graphique est exprimé sur une échelle logarithmique.



Les émissions totales (en France et importées pour répondre aux usages des Français) des scénarios Pari Réparateur et Tendanciel suivent une courbe assez proche, bien que les émissions du scénario Pari Réparateur dépassent celles du scénario Tendanciel. Les émissions totales du scénario Technologies Vertes connaissent une stabilité importante entre 2030 et 2050, avec la reprise d'une légère tendance haussière par la suite.

Les émissions totales du scénario Coopérations Territoriales connaissent une baisse importante et régulière, se rapprochant en 2060 des émissions des centres de données présents sur le sol français pour les scénarios Tendanciel, Pari Réparateur et Technologies Vertes.

Les émissions totales du scénario Génération Frugale marquent une baisse encore plus forte que celle de Coopérations Territoriales, avec l'atteinte d'un seuil en 2050.

Concernant les émissions provenant des centres de données présents uniquement sur le sol français, les scénarios Tendanciel, Pari Réparateur et Technologies Vertes, connaissent une évolution très proche, avec des émissions en augmentation qui pourraient atteindre respectivement autour de 2,13 Mt CO₂ eq., 2,26 et 2,51 Mt CO₂eq. en 2060.

Les scénarios Coopérations Territoriales et Génération Frugale quant à eux, marquent une baisse des émissions sur le sol français, et atteindraient respectivement 0,18 Mt CO₂eq. et 0,15 Mt CO₂ eq. en 2060. Ainsi, concernant les centres de données présents sur le sol français, l'écart d'émissions entre les scénarios en 2050 serait de 1 à 11 entre le scénario le plus bas (Génération Frugale) et le scénario le plus élevé (Technologies Vertes). Cet écart se creuse de façon impressionnante si l'on tient compte des émissions importées, celles-ci rendant compte d'un mix électrique plus émetteur de GES que le mix électrique français, avec un écart de 1 à 187 entre le scénario le plus bas (Génération Frugale) et le scénario le plus élevé (Pari Réparateur).

6.6.5. Modélisation du potentiel de récupération de chaleur fatale

Nous avons estimé le potentiel de récupération de la chaleur émise par les équipements IT des centres de données, communément appelée "chaleur fatale", en fonction des scénarios définis dans le modèle prospectif.

Le potentiel brut correspond à la quantité totale d'énergie émise sous forme de chaleur dans un centre de données et correspond à la consommation électrique de la partie IT. Le **potentiel net** de récupération de la chaleur fatale correspond au potentiel brut, pondéré par l'efficacité, actuelle et future, des différentes technologies de refroidissement présentes sur le marché, du point de vue de la récupération de chaleur et la capacité à trouver des débouchés. Cette efficacité augmente avec l'amélioration des techniques de captation de la chaleur et des débouchés plus atteignables, ce qui permet une amélioration du potentiel net de récupération. Ainsi l'utilisation du *liquid cooling* qui apparaît comme solution dans certaines salles informatiques permet de récupérer plus facilement la chaleur des serveurs et à une température plus élevée plus facilement valorisable. **Précisons également qu'il s'agit uniquement d'un potentiel réalisable d'un point de vue technique, qui ne correspond pas à la réalité à date puisqu'un nombre très restreint de centres de données français récupèrent réellement la chaleur fatale à ce jour.**

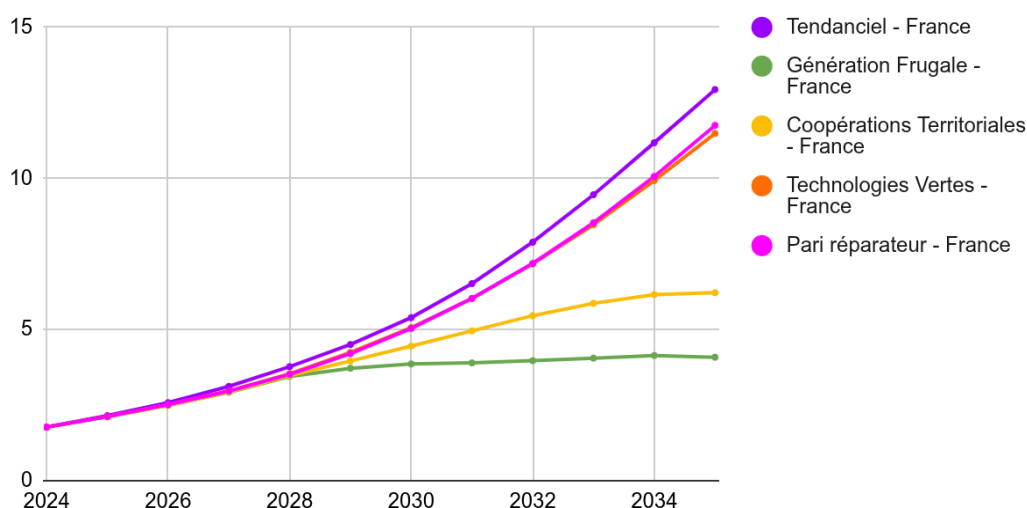
Ces projections ne tiennent pas compte de l'intérêt économique réel de la récupération de cette chaleur et des particularités locales propres à chaque centre de données, ces éléments réduisant la probabilité d'une telle récupération.³⁴⁵

Cette projection passe de 1,77 TWh de potentiel net de récupération de chaleur sur le territoire français pour 2024, aux valeurs suivantes en 2035 :

³⁴⁵ Pour plus de détails sur la méthodologie appliquée concernant l'estimation du potentiel de récupération de la chaleur fatale, voir la section correspondante dans le rapport

- 12,94 TWh pour le scénario Tendanciel
- 4,09 TWh pour le scénario Génération Frugale
- 6,22 TWh pour le scénario Coopérations Territoriales
- 11,49 TWh pour le scénario Technologies Vertes
- 11,76 TWh pour le scénario Pari Réparateur

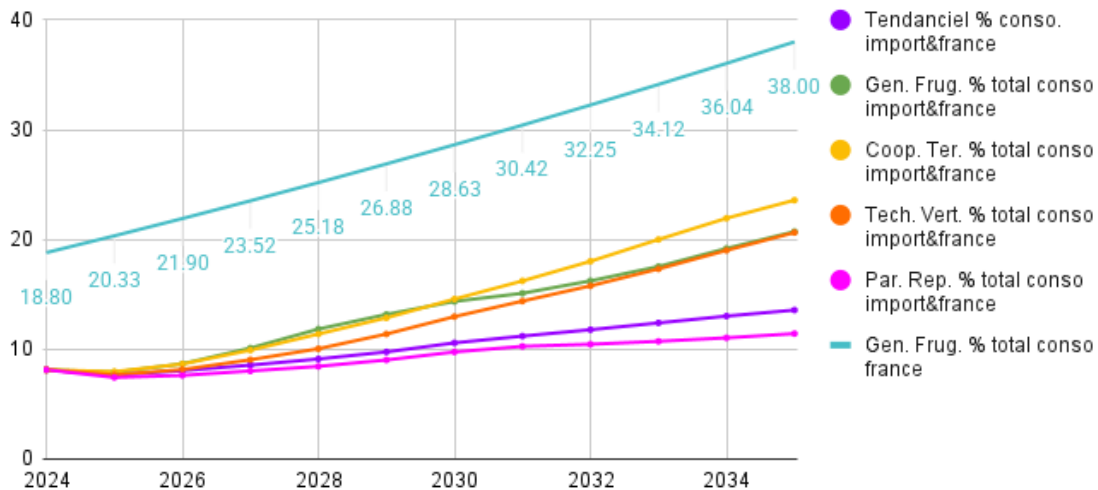
Projection de potentiel net de récupération de chaleur fatale en valeur absolue (TWh)



Ainsi, si la totalité du potentiel net de récupération de chaleur fatale était effectivement récupérée - hypothèse purement théorique et ne tenant pas compte de la faisabilité économique ni des réalités locales propres à chaque centre de données - cette récupération pourrait atteindre jusqu'à 23,6% de la consommation électrique des centres de données pour les usages des Français³⁴⁶ en 2035, dans le cas du scénario Coopérations Territoriales. *A contrario*, le scénario Pari Réparateur permettrait de récupérer en net jusqu'à 11,4% de la chaleur fatale de l'ensemble des centres de données du scénario, en 2035.

³⁴⁶ On tient compte ici de la consommation d'électricité totale nécessaire pour répondre aux usages des Français (celle des centres de données sur le sol français et celle de la part des centres de données étrangers sollicités pour cet usage)

Projection de potentiel net de récupération de chaleur fatale (en % du total de consommation d'électricité des datacenters en France et à l'étranger pour l'usage



Dans tous les scénarios, à l'échelle de consommations des centres de données présents sur le sol français uniquement, ce potentiel net de récupération de la chaleur fatale passerait d'un plafond de 18,8% en 2024 à un plafond de 38% en 2035. Ainsi, si le potentiel de récupération de chaleur fatale est un levier intéressant à mobiliser pour optimiser l'usage de l'énergie, son intérêt est à mettre en regard du niveau de sobriété, du total de la consommation d'énergie en valeur absolue et de sa faisabilité.

7. Limites de l'étude

7.1. Limites liées au périmètre de l'étude

7.1.1. Modèle limité à la phase d'usage

Notre modèle est, pour cette première version, limité à la phase d'usage. Des recherches complémentaires, et la mise à disposition publique de nouvelles données, permettront à l'avenir de venir compléter le modèle sur ce point, afin de proposer une modélisation des émissions sur l'ensemble du cycle de vie des centres de données. Cette limite, à date, est d'autant plus importante que la phase de fabrication représente couramment, en France, la majeure partie des émissions de GES du cycle de vie d'un centre de données.

Alors que le contexte actuel voit une forte croissance de la construction de bâtiments ex-nihilo pour les centres de données, notamment en raison de l'essor de l'IA, et le renouvellement rapide des équipements IT, cette limite est d'autant plus dommageable qu'une approche uniquement basée sur la phase d'usage favorise, du fait de l'invisibilisation de la fabrication, le renouvellement des équipements et le gain en efficacité. Elle peut ainsi contribuer involontairement à encourager le renouvellement des équipements et la recherche de gains d'efficacité, qui sont les recommandations les plus visibles à ce jour, alors qu'une analyse du cycle de vie complet des centres de données et des matériels plaide pour une augmentation de la durée de vie et une rationalisation des volumes, en priorité.

7.2. Limites liées à la maille géographique définie

7.2.1. Limites de l'application de données mondiales à une échelle nationale

La définition d'une maille géographique apporte un certain nombre de limites. En effet, beaucoup de données ou d'estimations sont mondiales : cette échelle, très influencée par une typologie de centres de données différente de ceux présents à l'échelle française (principalement états-unienne, du fait de leur première place au rang du nombre et de la taille des centres de données), entraîne des difficultés à adapter ces données à l'usage France, par manque de données spécifiques. Nous avons cherché à réduire cette limite par la réalisation d'un certain nombre d'hypothèses spécifiques à la France.

Le manque de visibilité des phénomènes de congestion locaux

Par ailleurs, la maille nationale ne permet pas non plus de rendre compte de certains phénomènes locaux de congestion liés à une forte concentration régionale de centres de données : limites de capacité de transport du réseau électrique, consommation électrique, concentration foncière, consommation d'eau, etc. L'invisibilisation de ces points de congestion ne permet pas de lever les points d'attention sur des phénomènes de tension locaux qui pourraient advenir bien plus tôt à l'échelle régionale dans certains territoires qu'ils ne se manifesteront à l'échelle nationale (le phénomène étant lissé).

7.3. Limites liées à la disponibilité des données

7.3.1. Données disponibles sur le périmètre des “centres de données traditionnels”

Bien que nous ayons collecté des données pour les centres de données privés, du secteur public et HPC, pour établir la baseline 2024, nous sommes partis de l'évaluation des besoins en surface d'hébergement des entreprises, initiée par les co-auteurs de l'étude ADEME/ARCEP 2022, pour définir le point de départ de la catégorie “centres de données traditionnels”, dans notre modèle prospectif. Les “centres de données traditionnels” concernent les centres de données hors colocation (compris dans notre collecte), mais aussi les salles informatiques comprises dans des bâtiments tertiaires, que nous ne pouvions pas identifier de la même manière. La catégorie établie par l'étude ADEME/ARCEP nous a permis de partir d'un périmètre plus complet en incluant ces salles informatiques, mais ceci nous a contraint de ne pas inclure une partie de notre collecte dans le modèle prospectif. Ce sujet mériterait de prolonger l'effort de collecte, avec d'autres méthodes, en l'étendant aux salles informatiques qui ne se trouvent pas dans un bâtiment dédié à cet usage (condition nécessaire à une collecte par image satellite ou données de cadastre), pour affiner l'estimation les concernant.

7.3.2. Usage du Cloud et tendances move-to-cloud

La répartition entre les usages des services Cloud externalisés et les usages d'infrastructures gérées par les organisations bénéficiaires est une donnée inaccessible. Notre veille n'a pas non plus permis d'identifier de données robustes pour l'estimer. De la même manière, l'intensité de la tendance de move-to-cloud est à questionner, ainsi que celle du retour à des infrastructures on-premise de la part de certaines organisations.

7.3.3. Consommations d'énergie des centres de données

Les consommations d'énergie estimées pour 2024 sont basées sur des données de surface et des données d'intensité énergétique par m² de surface, issus de la base de données européenne alimentée dans le cadre de la directive d'efficacité énergétique. Celles produites par le modèle prospectif s'appuient sur un inventaire IT projeté et des tendances d'usage. Dans un cas comme dans l'autre il s'agit de modélisation, ce qui présente l'avantage de fournir de premiers enseignements malgré le manque de données primaires. L'obtention de données primaires serait un atout majeur pour une meilleure compréhension, plus exhaustive de ce sujet. La base de données européenne devrait fournir une amélioration sur ce sujet, une fois la directive d'efficacité énergétique retranscrite en droit français et une fois que la collecte des données couvrira plus largement le parc de centres de données français.

7.3.4. Surfaces IT et surface de plancher

Notre méthode de collecte pour établir la baseline 2024 est très dépendante des surfaces obtenues à partir, soit des données de cadastre, soit de données satellite. Il est possible qu'une partie de ces données soient erronées, soit parce que le cadastre n'est pas complètement à jour, soit parce que les images satellite ne sont pas tout à fait à jour (nous avons déjà identifié des différences sur des images provenant de deux outils différents pour la même année), soit parce que le ratio de 1/2 pour inférer la surface IT à partir de la surface de plancher sont inadaptés à un site particulier (c'est une moyenne, avec tous les biais que cela comporte).

7.3.5. Taux de charge et de remplissage

Nous avons pris en compte un taux de remplissage des m² de surface IT disponibles, dans notre décompte de la baseline 2024. Ce taux de remplissage, prenant en compte la probabilité que l'opérateur d'un centre de données ne déploie pas dès le départ l'ensemble des salles IT et de la surface disponible pour l'hébergement. C'est un premier niveau de modélisation de l'occupation réelle du centre de données. Dans le cadre de cette étude il était impossible de fournir un niveau de modélisation plus détaillé. Des travaux ultérieurs en ce sens auraient beaucoup de sens pour affiner les estimations.

7.3.6. Courbes de ramp-up

Le ramp-up désigne la montée en charge IT d'un centre de données, ce qui peut se traduire à la fois en une augmentation progressive de la puissance électrique installée et de la surface IT et totale construite du bâtiment. Un projet de centre de données s'accompagne généralement d'annonces d'installation de puissance électrique d'une ampleur plus importante que celle qui est réellement installée pour l'inauguration. Cette puissance électrique cible est atteinte progressivement, suivant plusieurs paliers de raccordement. De la même manière, la surface IT disponible annoncée au départ est souvent supérieure à la surface initialement déployée. Des salles IT supplémentaires sont installées au fur et à mesure de la vie du centre de données. Enfin, la densité d'équipements IT installés dans le centre de données varie au fil du temps. Cette densité augmente jusqu'à s'approcher d'une saturation éventuelle, soit en surface, soit en puissance, ce qui peut entraîner l'extension du bâtiment, la fabrication d'un second bâtiment, ou bien une extension par d'autres moyens (voir [Évolution bâtiment : les centres de données modulaires](#)). Une modélisation fine de ces différents niveaux et nuances de montée en charge est nécessaire pour refléter plus fidèlement les évolutions de consommation d'énergie et d'eau des centres de données, en fonction de leur date de mise en service et de leur catégorie. L'évaluation des consommations 2024 effectuée dans cette étude comprend une modélisation de l'augmentation progressive des surface IT effectivement disponibles. La projection à 2026 intègre un ratio global de ramp-up qui vise à rassembler ces différentes nuances. Une modélisation plus fine de ces différentes dynamiques fait partie des améliorations à apporter au modèle de cette étude pour la suite.

7.3.7. Consommations d'eau

L'accès aux données de consommation d'eau a été l'une des difficultés les plus marquées de cette étude. En France, l'ARCEP dispose de données sur la consommation d'eau, en 2023, d'un échantillon de 150 centres de données étudiés pour son enquête annuelle "pour un numérique soutenable"³⁴⁷. Il nous a été impossible de consulter les données de consommation de ces centres de données, que ce soit par centre de données, ou agrégées par catégorie de centres de données, car elles sont sous protection pour secret des affaires. L'autre piste se trouvait dans les données collectées dans le cadre de la directive d'efficacité énergétique de l'Union Européenne. Si cette source de données deviendra vraisemblablement une source de référence dans les prochaines années, l'état des données collectées à date n'a pas permis à cette étude de progresser sur le sujet des consommations d'eau. Une mise à jour de notre modèle avec des données provenant de cette base de données, une fois l'obligation de reporting retranscrite en droit français, constituerait une amélioration significative.

7.3.8. Intensité de renouvellement du matériel IT

L'étude se focalise sur la partie usage des centres de données et ne prend pas en compte la phase de Fabrication du matériel IT.

³⁴⁷ [Enquête annuelle "pour un numérique soutenable", ARCEP, 2025](#)

La prise en compte de la phase de Fabrication nous permettrait également de quantifier l'impact du renouvellement du matériel pour améliorer la consommation énergétique. C'est un sujet important, par exemple dans le cadre de la prise en compte du développement de l'IA qui va entraîner un déploiement potentiellement massif de nouveaux matériels.

7.3.9. Destination et provenance de l'usage des centres de données

L'objet de l'étude est d'évaluer la consommation des centres de données pour un usage français. Or, il n'existe pas de données estimant le volume et l'impact :

- Des centres de données extérieurs dédiés à un usage français
- Des centres de données opérés en France pour un usage à l'étranger.

Nous avons proposé une approximation du volume des services importés basé sur un modèle à l'échelle mondiale. Il présente forcément des limites. Il serait opportun d'étudier plus précisément la balance import/export des centres de données pour un usage français.

7.4. Limites liées aux indicateurs

Le rapport se base sur deux indicateurs : consommation électrique et potentiel de dérèglement climatique. Une approche ACV (Analyse de Cycle de Vie) permettrait de traiter l'ensemble des indicateurs d'impacts environnementaux. Par exemple, la prise en compte de l'indicateur ADPe (potentiel d'épuisement des ressources en minéraux et métaux) apparaît essentiel dans le traitement du numérique (si prise en compte de la phase de fabrication).

7.5. Limites liées à l'exercice prospectif

La réalisation d'un exercice prospectif est un exercice d'analyse des tendances (passées, présentes, pour anticiper les futures), qui demande une vision large, de la prise de la hauteur, mais aussi de considérer une multitude de paramètres dont certains sont porteurs de conséquences potentiellement systémiques. Un exercice prospectif n'est de ce fait, **pas un exercice de prédiction** ou de probabilité, mais **plutôt l'étude des futurs possibles** ; la liberté, la volonté et la responsabilité collective des acteurs pouvant à tout moment infléchir dans une direction ou une autre l'avenir.

Dans le cadre de cet exercice prospectif sur l'évolution des consommations des centres de données à l'échelle française, plusieurs limites sont identifiées, dont principalement :

1. Le besoin de données macroéconomiques :
Le développement de la numérisation des autres secteurs d'activités, l'adoption et le périmètre du déploiement de l'IA dans les années et décennies à venir nécessite des connaissances et des données en macro-économie que nous n'avons pas obtenues au moment de l'étude. Les scénarios autour de la place que prendra l'IA de nouvelle génération sont balbutiants et manquent de la prise de recul nécessaire.
Ainsi, on note des changements de stratégie soudains et brutaux de la part de nombreux acteurs de premier plan qui soulignent le haut niveau d'incertitude de l'exercice prospectif.
2. La difficulté de viser une cible temporelle si éloignée (2060) : à l'échelle du développement des centres de données et l'évolution du numérique, l'an 2060 apparaît comme un horizon extrêmement lointain, sujet à de nombreuses fluctuations difficilement anticipables, tant sur les évolutions pouvant parfois être disruptives en termes d'innovations, technologiques ou d'usages, que sur des retournements de marché (bulles, disponibilité ou tension sur les ressources, ...). Si de

tels paramètres peuvent être imaginés, il est en revanche bien délicat de les anticiper au point de faire apparaître sur une courbe quantitative un seuil et une date de rupture.

3. La pertinence de considérer un ou plusieurs scénarios “crunch” : en lien avec le point précédent, l’advenue sur le moyen ou long terme d’un scénario crunch, en raison de l’atteinte d’une ou plusieurs limites (rupture de flux ayant des effets sur les ressources disponibles par exemple), pourrait être considérée. Dans le cadre de cette étude, les scénarios analysés et envisagés devaient, sauf le tendanciel (qui comme son nom l’indique suit cette trajectoire théorique), tous s’inscrire dans les scénarios de transition de l’ADEME. Un scénario crunch n’entrant pas dans la définition d’un scénario de transition (mais de rupture), un tel scénario n’a pas fait l’objet de modélisation, bien que certaines études analysées dans le cadre de notre recherche aient parfois modélisé un tel scénario.³⁴⁸
4. Enfin, du fait de l’effet de lissage mentionné dans les limites liées à la maille géographique, mais aussi du fait du manque d’études prospectives permettant de combiner les risques environnementaux systémiques et locaux, notre étude ne tient pas compte dans ces projections des besoins d’adaptation aux effets des polycrises environnementales sur les centres de données. Par exemple, des tensions sur la disponibilité d’une ressource, telle que celle sur l’eau, pour la fabrication des microprocesseurs à Taïwan en 2021, pourraient avoir des répercussions en cascade, pouvant ralentir la montée en puissance de certains centres de données. L’addition de plusieurs crises pouvant produire des effets de rupture de flux aux impacts systémiques, cette limite est une limite importante, qui sera à prendre en compte de manière de plus en plus importante prochainement, en raison des effets combinés du réchauffement climatiques (tensions sur l’eau, la disponibilité de l’électricité, les risques d’inondation, de feux, ...) et les autres facteurs de risques environnementaux ou sociétaux.

³⁴⁸ Voir à ce titre l’étude de Schneider Electric, *Artificial Intelligence and Electricity*, Sustainable Research Institute, Dec 2024 : https://download.schneider-electric.com/files?p_Doc_Ref=TLA_System_Dynamics_Approach&p_enDocType=Thought+Leadership+article&p_File_Name=SRI_Artificial_Intelligence_A_System_Dynamics_Approach_December_2024_V20250414.pdf

Conclusion

Au terme de ce rapport, nous avons pu voir que l'évolution tendancielle des consommations des centres de données connaît actuellement une hausse majeure, du fait d'une croissance importante des usages, particulièrement portée par l'essor de l'IA générative, des services numériques captant données et temps d'attention de façon croissante, mais aussi du fait de l'essor des projets d'implantation de nouveaux centres de données et des services proposés. Ces projets amplifient la demande IA et cloud par la facilité d'accès aux infrastructures. Dans le même temps, les services visibles des utilisateurs mettent les fonctions basées sur l'IA générative³⁴⁹ au centre des interfaces et poussent les utilisateurs à s'en servir indépendamment d'une demande explicite.

Nous avons pu voir que la "migration" vers le cloud s'apparente plutôt à la mise en œuvre de nouveaux usages et services complémentaires aux précédents, dont la substitution n'est pas démontrable. L'accumulation des infrastructures dites "on-premise" et des infrastructures Cloud contribuent donc à la croissance ininterrompue des consommations d'énergie et des impacts environnementaux des centres de données.

Ces évolutions sont soutenues par des évolutions techniques en termes de capacité de calcul (notamment sur les GPU et TPU pour l'IA en particulier générative) : bien que les capacités de calcul continue de gagner en efficacité énergétique, ces gains sont cependant compensés par l'augmentation des capacités des composants d'une part, de leur volume de production de l'autre, en plus de l'augmentation encouragée des usages, créant un effet rebond (cf. à ce titre la section [Évolutions matériel IT](#)).

Les évolutions récentes et en cours des capacités des composants (principalement destinés à l'IA générative), stimulent la construction de centres de données de puissances IT et comprenant des surfaces bien plus importantes que ceux construits précédemment en France, allant de pair avec une évolution importante en cours des techniques de refroidissement et de la densité énergétique croissante des baies informatiques. Ces nouvelles techniques refroidissement (DLC, Immersion Cooling, voir la section [Techniques de refroidissement](#)) apportent une capacité de refroidissement bien plus importante par unité de surface et pourraient contribuer à une mutualisation et une optimisation plus forte des infrastructures IT, ainsi qu'à une meilleure récupération de la chaleur fatale. Cependant, l'essor de ces nouvelles techniques de refroidissement est ralenti par des limites techniques importantes : précautions de manipulation, conditionnement voire disposition des machines, urbanisme des salles IT différent, ou même dans certains cas refonte complète du centre de données ou ouverture d'un centre de données dont l'architecture prévoit nativement ces contraintes. De plus, les usages couramment associés à ces technologies sont de nature à contribuer à l'augmentation significative des consommations globales, ce qui est particulièrement visible à l'échelle de la puissance demandée dans une baie informatique moyenne prévue à cet effet, par rapport aux standards précédents³⁵⁰.

Il faut ajouter que les consommations électriques des centres de données et les émissions de GES associées à ces consommations ne représentent pas une vue complète des impacts et enjeux environnementaux associés aux centres de données : les impacts sur les consommations de foncier, la fabrication des bâtiments (et leur impact sur la biodiversité), les consommations d'eau (ou l'eutrophisation quand l'eau est rejetée à une température trop élevée), mais aussi liés à la fabrication des équipements (en particulier IT, avec une durée de vie courte) contribuant à l'épuisement des minéraux et métaux, sont autant d'enjeux environnementaux à considérer, potentiellement plus importants que les enjeux liés à la phase d'usage. Ces pôles d'impact ne sont pas quantifiés dans le modèle développé par le consortium CLIK dans le cadre de cette étude.

³⁴⁹ Voir la section [IA Générative](#)

³⁵⁰ Voir la section [Augmentation de l'intensité énergétique par Rack](#)

Cette étude a nécessité la réalisation d'un inventaire des centres de données présents sur le sol français métropolitain, nous permettant de disposer d'une modélisation de la situation initiale pour l'année de référence 2024. Les résultats de cette collecte nous ont permis d'identifier **352 centres de données en exploitation**, dont 42% en Île-de-France, représentant 67% de la puissance installée pour l'ensemble des centres de données à l'échelle nationale. La puissance installée totale cumulée de ces centres en exploitation est de **2,46 GW en 2024 en France**. Ce travail a été complété par la comptabilisation partielle des projets nouveaux documentés à date de l'inventaire jusqu'à 2035, dont nous n'avons pas pu établir la probabilité d'aboutissement. Ces projets annoncés totalisent à eux-seuls une puissance de 6,5 GW, à horizon 2035, soit une puissance installée totale du parc de centres de données de près de 9 GW si l'ensemble de ces projets venaient à se réaliser dans les conditions annoncées et à leur pleine capacité. La volatilité et le peu de fiabilité des annonces nous a confortés à ne pas utiliser ces projections de puissances installées pour nos projections.

Le modèle CLIK, développé dans le cadre de cette étude, et basé sur une méthodologie complète expliquée au chapitre [3. Méthodologie du modèle](#), nous a permis de modéliser une situation initiale des consommations d'électricité des centres de données,³⁵¹ de 10 TWh en 2024 en France, et de 14 TWh concernant les consommations d'électricité des centres de données situés en dehors du territoire français pour répondre aux usages des Français (dits, centres de données "importés"), soit un total de **24 TWh en 2024** pour répondre aux usages des français.

À partir de cette situation initiale, nous avons modélisé 5 scénarios (Tendanciel, Génération Frugale, Coopérations Territoriales, Technologies Vertes, Pari Réparateur)³⁵² année après année, avec un premier jalon méthodologique à moyen terme en 2035 puis une seconde méthodologie post-2035 jusqu'à 2060. Les résultats de cette modélisation montrent une variabilité très forte des consommations électriques et des émissions de GES selon les scénarios, un écart qui s'accroît au fur et à mesure du temps en fonction des choix pris.

Ainsi, une forte expansion des consommations des centres de données entraînerait un accroissement important des émissions du secteur, incompatible avec l'Accord de Paris et la trajectoire SBTi du secteur du numérique. Concernant les scénarios Transition(s) 2050,³⁵³ l'atteinte du Zéro Émissions Net en 2050 se fait à travers des mesures diverses et complémentaires, auquel chaque scénario a recours de manière variée.³⁵⁴ Dans le cadre de cette étude sur le périmètre des centres de données, nous avons pu voir que pour les cas des scénarios Technologies Vertes et Pari Réparateur, l'augmentation de la consommation électrique doit être anticipée et planifiée pour satisfaire des usages toujours croissants qui, passé un certain seuil, pourraient être concurrents de l'électrification demandée pour la transition écologique d'autres secteurs d'activité. De même, **si les scénarios Technologies Vertes et Pari Réparateur soutiennent aussi la réduction des émissions d'autres secteurs d'activité, les émissions des centres de données dans ces deux scénarios sont toutefois très significatives et nécessitent une compensation importante des émissions**. Ainsi, à l'échelle des consommations des centres de données, les optimisations technologiques telles que l'augmentation de l'efficacité et la récupération de la chaleur fatale ne suffiront pas à stabiliser la croissance des consommations électriques des centres de données en France sans la

³⁵¹ Incluant les salles informatiques des bâtiments tertiaires

³⁵² Voir à ce titre les [chapitres 4. Scénarios : approche de révision des scénarios Transition 2050 et scénario tendanciel](#) et [6. Résultats](#)

³⁵³ L'ensemble de ces scénarios visent à viser à atteindre le Zéro Émissions Net en 2050 sur les émissions du périmètre France, c'est-à-dire, les émissions produites depuis le territoire français, ceci en tenant compte du fait que les centres de données ne sont qu'une des composantes parmi les secteurs calculés pour atteindre le Zéro Émissions Net dans une approche globale tous secteurs

³⁵⁴ Ces mesures (sobriété, efficacité, compensation par d'autres secteurs d'activités ou par captation, puits technologiques ou naturels...) n'ont pas été évaluées à l'échelle inter-sectorielle dans cette étude : nous nous focalisons ici sur les scénarios Transition(s) 2050 sur le périmètre des centres de données sans évaluation de leur contribution relative à la baisse des émissions sur d'autres secteurs (hors périmètre de cette étude).

mise en place d'une politique active de sobriété pour les services numériques et les usages. Par ailleurs, l'empreinte liée aux émissions importées sur ces scénarios est proportionnellement très importante au regard des émissions des centres de données en France.

Les cas des scénarios Génération Frugal et Coopérations Territoriales nous ont permis d'explorer que **seule une politique de sobriété très volontariste, modifiant en profondeur nos modes de vie, permettrait un inflexion permettant de réduire les consommations futures des centres de données**, permettant aux centres de données de contribuer moins fortement aux besoins de compensation pour atteindre le Zéro Émissions Net pour chacun des scénarios de Transition(s), et de réduire au global de façon importante leur empreinte. En raison de leur dimension systémique, de telles politiques ne peuvent réussir qu'à travers la mobilisation conjointe des pouvoirs publics, des entreprises et des citoyens. La modélisation des scénarios nous montre que leurs effets sur les consommations ne sont pas immédiats mais graduels sur plusieurs années.

Enfin, la réalisation de cette étude, d'une base et données des centres de données en France, et d'un modèle prospectif des évolutions des consommations électriques en France, ouvre la voie à un étayage futur de ces travaux pour rendre compte de façon plus complète et systématique des impacts environnementaux des centres de données, notamment :

- mieux répertorier les consommations, au-delà des consommations électriques : eau, foncier, matériel informatique, ...
- tenir compte aussi des autres phases du cycle de vie des centres de données en complément de la phase d'utilisation : construction des centres, fabrication des équipements, transport, fin de vie. La phase de fabrication (bâtiment et matériels IT) - non prise en compte dans cette étude - représente 74% de l'ensemble de l'empreinte en potentiel de dérèglement climatique des datacenters sur le territoire français selon le rapport ADEME/ARCEP. Le poids de la phase de fabrication pourrait même augmenter si la course à l'efficacité énergétique augmente les taux de renouvellement du matériel IT
- tenir compte d'autres indicateurs d'impacts environnementaux en complément de celui sur les émissions de GES, etc.

Cet étayage futur nécessite une plus grande disponibilité et transparence sur les données, et des travaux complémentaires, qui pourraient permettre une version plus aboutie et complète du modèle. C'est dans ce but que l'ensemble des modèles de l'étude sont disponibles en open data.

L'ADEME EN BREF

Au cœur des missions qui lui sont confiées par le ministère de la Transition écologique, le ministère en charge de l'Energie et le ministère en charge de la Recherche, l'ADEME - l'Agence de la transition écologique - partage ses expertises, coordonne le financement et la mise en œuvre de projets de transformation dans plusieurs domaines : énergie, économie circulaire, décarbonation, industrie, mobilité, alimentation, adaptation et sols.

Elle mobilise les citoyens, les acteurs économiques et les territoires, et leur donne les moyens de progresser vers une société économe en ressources, plus sobre en carbone, plus juste et harmonieuse. Résolument engagée dans la lutte contre le changement climatique et la dégradation des ressources, l'ADEME conseille, facilite et aide au financement de nombreux projets, de la recherche jusqu'au partage des solutions. Elle met ses capacités d'expertise et de prospective au service des politiques publiques.

L'ADEME est un Établissement public à caractère industriel et commercial (EPIC).

LES COLLECTIONS DE L'ADEME



FAITS ET CHIFFRES

L'ADEME référent : Elle fournit des analyses objectives à partir d'indicateurs chiffrés régulièrement mis à jour.



CLÉS POUR AGIR

L'ADEME facilitateur : Elle élabore des guides pratiques pour aider les acteurs à mettre en œuvre leurs projets de façon méthodique et/ou en conformité avec la réglementation.



ILS L'ONT FAIT

L'ADEME catalyseur : Les acteurs témoignent de leurs expériences et partagent leur savoir-faire.



EXPERTISES

L'ADEME expert : Elle rend compte des résultats de recherches, études et réalisations collectives menées sous son regard.



HORIZONS

L'ADEME tournée vers l'avenir : Elle propose une vision prospective et réaliste des enjeux de la transition énergétique et écologique, pour un futur désirable à construire ensemble.



EXPERTISES

PROSPECTIVE D'EVOLUTION DES CONSOMMATIONS DES CENTRES DE DONNEES EN FRANCE DE 2024 A 2060

Cette étude de l'ADEME répond à 3 objectifs :

1. Dresser l'état des lieux de la consommation électrique actuelle des centres de données en France ;
2. Proposer un modèle prospectif détaillé permettant de modéliser des scénarios d'évolution des consommations des centres de données dans le temps ;
3. Modéliser et analyser 5 scénarios prospectifs jusqu'à 2060 : un scénario tendanciel et 4 scénarios envisageant les 4 chemins possibles de transition écologique.

Cette étude a considéré les émissions de gaz à effet de serre engendrées, la pression sur les ressources en eau, l'artificialisation des sols, la concurrence avec les d'autres usages de l'électricité, ou encore les enjeux de souveraineté numérique et distingue les consommations en France des consommations à l'étranger pour un usage français.

Les prospectives à long terme de consommation énergétique des data centers pour un usage en France sont très contrastées en fonction des choix de société.

Les projets déjà engagés entraîneront une consommation multipliée par 4 d'ici 10 ans pour 37 TWh en France.

Pour un scénario pariant sur les technologies, les consommations nécessaires de près de 350 TWh se feront pour 80% à l'étranger, interrogeant notre souveraineté numérique.

Le modèle prospectif sera accessible et mis en open source par l'ADEME.