

Serveurs et baies

Un rack d'intelligence artificielle pèse aujourd'hui plus d'une tonne, coûte environ trois millions de dollars, et ne peut plus être refroidi par de l'air, un choix que son fabricant ne propose même plus. Le format qui a organisé le datacenter pendant trente ans vient de rencontrer sa limite physique.

Septembre 2026

Classification : Non sensible

Lecture : 17 min

Fondations informatiques : L'ordinateur et ses composants · Cartes graphiques · Systèmes d'exploitation & virtualisation · **Serveurs & baies** ← · NAS & SAN · Équipements réseau · Protocoles Internet · Cybersécurité, les bases · Bases de données · Algorithmique · Projets informatiques

L'ESSENTIEL

Un serveur est un ordinateur pensé pour tenir en rayonnage plutôt que sur un bureau, mesuré en unités de hauteur normalisées, l'unité U, quatre centimètres quatre vingt-quatre. Un modèle 1U tient dans la hauteur la plus fine, dense mais limité en emplacements d'extension ; un modèle 4U offre plus d'espace pour le stockage ou le refroidissement, au prix de l'encombrement. Un rack standard de datacenter empile jusqu'à quarante-deux de ces unités, un cadre resté stable depuis les débuts de l'informatique en réseau.

Ce cadre vient d'être pulvérisé par l'intelligence artificielle. La densité électrique moyenne d'un rack de datacenter a bondi de 69 % en un an, passant à environ 27 kilowatts en 2026. Un rack dédié à l'IA atteint déjà 80 à 120 kilowatts, huit à dix fois la norme d'entreprise classique. Et le système phare de Nvidia, le GB200 NVL72, qui regroupe soixante-douze puces graphiques et trente-six processeurs dans un seul rack intégré, consomme entre 120 et 132 kilowatts en régime nominal, pèse une tonne trois cent soixante, et coûte environ trois millions de dollars.

Ce niveau de puissance dépasse largement ce que l'air peut évacuer. Le refroidissement par air atteint son plafond pratique aux alentours de quarante à cinquante kilowatts par rack ; au-delà, le refroidissement liquide cesse d'être une option pour devenir la seule solution viable. Nvidia ne propose d'ailleurs plus de version refroidie par air pour son système le plus avancé : chaque puce, chaque processeur, chaque commutateur d'interconnexion est équipé de plaques froides raccordées directement à une boucle d'eau, une conception qui suppose que le bâtiment lui-même ait été pensé pour ça.

La trajectoire ne s'arrête pas là. Les architectures annoncées pour 2027 visent environ six cents kilowatts par rack, dix fois la densité déjà extrême du GB200 NVL72, avec une alimentation en huit cents volts continu plutôt que la norme actuelle. Un rack, autrefois une simple armoire métallique standardisée, est devenu un système industriel à part entière, vendu comme un seul produit intégré plutôt qu'assemblé pièce par pièce.

COMPRENDRE : À QUOI ÇA SERT, COMMENT ÇA MARCHE, OÙ ÇA S'IMBRIQUE

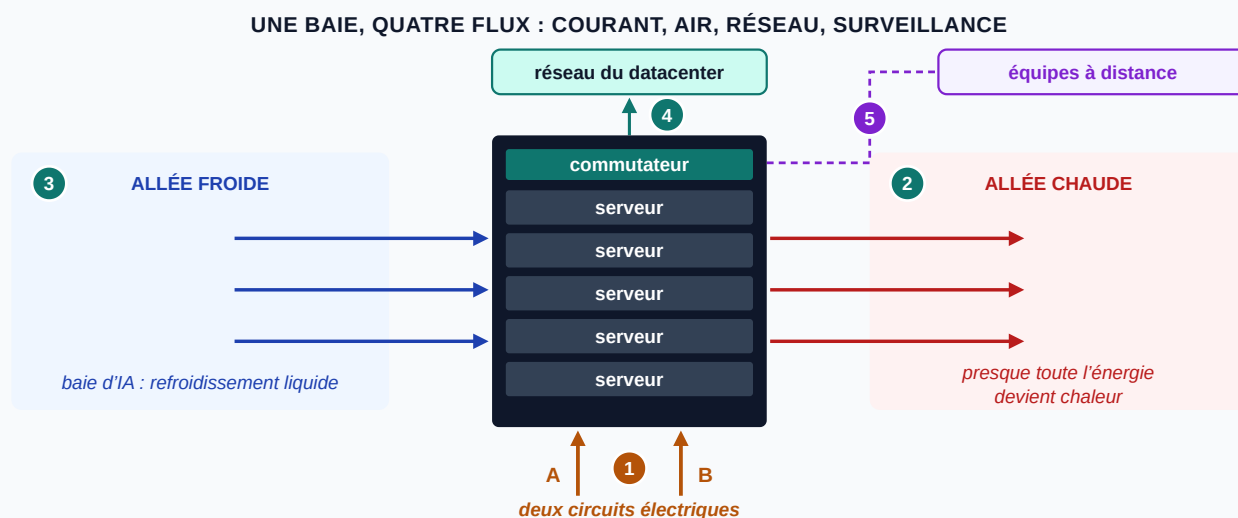
Pour le lecteur non spécialiste, environ quatre minutes. Initiés : passez directement aux chiffres qui suivent.

À QUOI ÇA SERT

Un serveur est un ordinateur conçu pour fonctionner jour et nuit, empilé avec d'autres dans une baie, un rayonnage normalisé. La baie lui fournit ce qu'un ordinateur de bureau trouve tout seul : du courant fiable, de l'air frais, du réseau et une surveillance à distance. C'est ainsi qu'un datacenter concentre, dans un espace et une énergie maîtrisés, la puissance que consomment les services numériques.

Image mentale : une baie est un immeuble de bureaux vertical : chaque étage, un serveur, travaille, mais l'immeuble fournit l'électricité par deux colonnes montantes, la ventilation, le câblage vers l'extérieur, et un gardien qui surveille tout depuis sa loge.

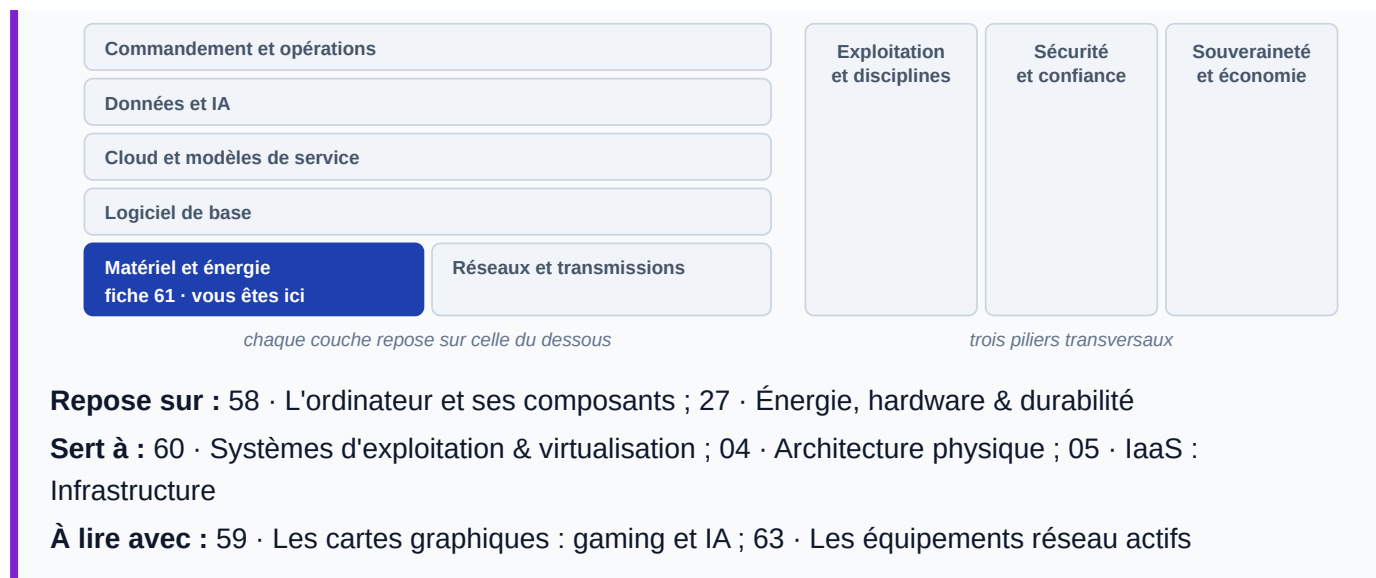
COMMENT ÇA MARCHE



- 1 L'électricité arrive par deux circuits indépendants, A et B, et chaque serveur a deux blocs d'alimentation : si un circuit tombe, l'autre prend le relais sans coupure.
- 2 Le serveur calcule, et transforme la quasi-totalité de cette électricité en chaleur : une baie d'IA en dégage autant que des dizaines de radiateurs.
- 3 L'air froid entre par la face avant, traverse les serveurs et ressort chaud à l'arrière, dans une allée isolée ; au-delà de quelques dizaines de kilowatts par baie, l'air ne suffit plus : un liquide vient refroidir directement les puces par des plaques froides, ou les serveurs sont plongés dans un fluide isolant, comme le détaille la section consacrée au refroidissement liquide.
- 4 Un commutateur, souvent placé en haut de la baie, relie chaque serveur au réseau du datacenter.
- 5 Une petite carte de gestion indépendante permet aux équipes de surveiller, d'éteindre ou de réinstaller un serveur à distance, même lorsqu'il ne répond plus.

OÙ ÇA S'IMBRIQUE DANS L'ENSEMBLE

Le serveur fait le lien entre le matériel et le cloud : c'est l'unité que l'hyperviseur découpe et que le cloud loue, et sa densité électrique relie directement cette fiche à l'énergie et à l'économie des datacenters.



LES CHIFFRES QUI RECADRENT LE DÉBAT (SEPTEMBRE 2026)

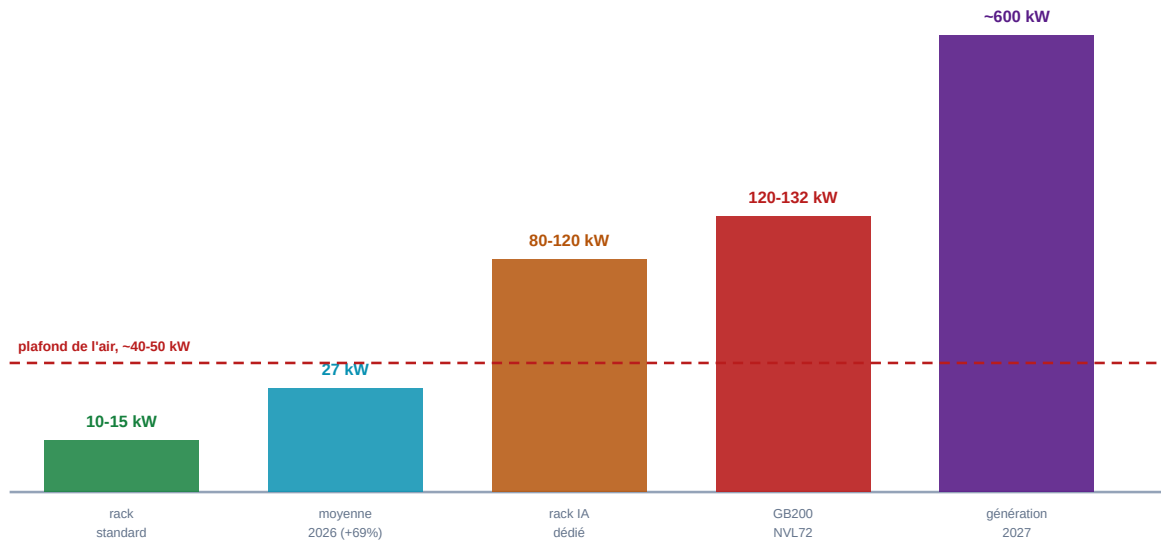
+69 % Croissance de la densité électrique moyenne d'un rack de datacenter en un an, atteignant environ 27 kilowatts en 2026, contre 7,6 kilowatts encore en 2025 selon l'Uptime Institute. <i>Network World, Uptime Institute, 2026</i>	120 à 132 kW Consommation nominale du rack Nvidia GB200 NVL72, jusqu'à 150 kW en pointe, soit seize à dix-sept fois la densité moyenne d'un rack traditionnel. <i>NVIDIA / HPE, 2025-2026</i>	1,36 tonne Poids d'un seul rack GB200 NVL72, pour un prix catalogue avoisinant les trois millions de dollars, soit environ 41 000 dollars par puce graphique. <i>Leviathan Systems, 2026</i>	~600 kW Densité visée par la génération de racks attendue en 2027, exigeant une alimentation en 800 volts continu, dix fois la densité déjà extrême du GB200 NVL72. <i>Feuille de route industrielle, 2026</i>
--	--	---	---

Lecture C-level : parler encore de racks en kilowatts à un chiffre, ou présumer que l'air suffit toujours à refroidir un datacenter moderne, revient à raisonner sur un monde qui a changé d'échelle en moins de deux ans.

L'ESCALADE DE DENSITÉ, DE L'ARMOIRE STANDARD AU RACK-TONNE : LE SCHÉMA C-LEVEL

Cinq paliers, une même unité de mesure, une croissance qui n'a plus rien de linéaire.

LA DENSITÉ D'UN RACK, DE 10 À 600 KILOWATTS



AU-DELÀ DU SEUIL, LE REFROIDISSEMENT LIQUIDE N'EST PLUS UNE OPTION, C'EST LA SEULE PHYSIQUE DISPONIBLE

Cinq paliers de densité, une même unité, une croissance qui a franchi le plafond physique de l'air dès le rack IA dédié. Le GB200 NVL72 et la génération 2027 ne sont plus des évolutions du même objet, ce sont des systèmes industriels d'une autre catégorie.

LES FORMATS QUI ONT ORGANISÉ TRENTE ANS DE DATACENTER

Avant l'IA, cette normalisation suffisait à tout. Elle reste la base sur laquelle l'exception s'est construite.

1

L'unité de rack, une norme mécanique avant d'être une norme de puissance

Un serveur se mesure en unités U, chacune mesurant quatre centimètres quatre vingt-quatre de hauteur ; un modèle 1U est le plus fin et le plus dense, un modèle 2U ou 4U offre davantage d'espace pour le stockage interne, les cartes d'extension ou un refroidissement renforcé.

Un rack standard de datacenter compte quarante-deux de ces unités, un gabarit resté identique depuis les débuts de l'informatique en réseau, bien avant que la puissance électrique par rack ne devienne le facteur limitant.

2

La redondance, l'assurance-vie du matériel

Un serveur d'entreprise embarque le plus souvent deux blocs d'alimentation électrique indépendants, chacun capable de porter seul la charge complète, une conception qui protège contre la panne d'un seul composant sans jamais interrompre le service.

À l'échelle du datacenter, cette logique se prolonge par des architectures dites N+1 ou 2N, où une capacité de secours entière, parfois une infrastructure électrique et de refroidissement dupliquée à l'identique, reste disponible en cas de défaillance de la chaîne principale.

3

Le rack devient un produit vendu entier, plus une armoire assemblée pièce par pièce

Le GB200 NVL72 n'est jamais vendu comme un lot de serveurs à installer séparément dans une armoire vide ; c'est un système industriel unique, livré pré-assemblé, avec sa propre alimentation, son propre circuit de refroidissement liquide intégré et son propre prix catalogue de plusieurs millions de dollars.

Acheter un rack d'intelligence artificielle en 2026 ressemble moins à commander des composants informatiques qu'à commander une centrale industrielle miniature.

LE REFROIDISSEMENT LIQUIDE : QUAND L'AIR NE SUFFIT PLUS

L'air est un piètre transporteur de chaleur : à volume égal, l'eau en emporte plus de trois mille fois plus. Tant qu'une baie dégage une dizaine de kilowatts, souffler de l'air froid suffit ; au-delà de quelques dizaines, il faut amener un liquide au plus près des puces. Quatre techniques se partagent le terrain, de la plus simple à installer à la plus efficace.

TECHNIQUE	PRINCIPE	DENSITÉ PAR BAIE	ATOUT	LIMITE
Air confiné	Allées froides et chaudes séparées, climatisation de salle.	jusqu'à 20 à 40 kW	Savoir-faire universel.	Plafonne vite ; la climatisation consomme beaucoup.
Porte arrière refroidie	Un échangeur à eau remplace la porte arrière de la baie et capte la chaleur de l'air qui sort.	30 à 80 kW	Se greffe sur une salle existante.	L'air reste le premier transporteur de chaleur.
Direct sur puce	Des plaques froides parcourues de liquide sont posées sur processeurs et cartes graphiques.	40 à 200 kW	Norme de fait des baies d'IA en 2026.	Capte 70 à 90 % de la chaleur ; le reste repart par l'air.
Immersion	Les serveurs baignent dans un fluide qui ne conduit pas l'électricité.	100 à 250 kW	Capte presque toute la chaleur et facilite sa réutilisation.	Cuves, fluides et maintenance spécifiques.

La bascule est en cours. Selon le rapport AFCOM 2026 sur l'état des datacenters, 36 % des exploitants interrogés ont déjà déployé le refroidissement liquide et 28 % prévoient de le faire d'ici deux ans ; plus d'une baie sur quatre en est équipée. Le refroidissement direct sur puce en est la forme dominante, et celle que NVIDIA prescrit pour ses systèmes GB200.

Le piège des indicateurs. L'efficacité énergétique d'un datacenter se mesure par le PUE, rapport entre l'énergie totale et celle des seuls équipements informatiques. Il flatte pourtant les salles refroidies par liquide, parce que l'énergie des pompes bascule dans le périmètre informatique. D'où l'importance de publier aussi le WUE, l'eau consommée par kilowattheure : les tours de refroidissement par évaporation consomment beaucoup d'eau, les boucles fermées et l'immersion beaucoup moins. Un datacenter vertueux sur un indicateur peut être gourmand sur l'autre.

Une technique plus ancienne qu'on ne le croit. Le refroidissement à l'eau n'est pas né avec l'intelligence artificielle : un hébergeur français le pratique depuis 2003, avec des blocs au contact des processeurs, une boucle d'eau fermée et un refroidissement du reste des composants par l'air extérieur, sans climatisation. L'IA ne l'a pas inventé ; elle l'a rendu incontournable.

SANS COMPLAISANCE : LA DENSITÉ EXTRÊME A UN COÛT QUE PEU D'ACTEURS PEUVENT ASSUMER SEULS

POINT DE VIGILANCE	CE QU'IL FAUT SAVOIR
La majorité des datacenters existants ne peuvent pas accueillir cette densité	Un bâtiment conçu pour du refroidissement par air, la norme jusqu'à récemment, ne peut pas héberger un rack GB200 NVL72 sans une reconstruction quasi complète de son circuit électrique et de refroidissement.
Le coût d'entrée exclut de fait les acteurs de taille modeste	À environ trois millions de dollars le rack, l'accès direct à cette classe de matériel reste réservé à des budgets d'infrastructure d'une échelle que peu d'organisations, hors les plus grands opérateurs cloud, peuvent mobiliser seules.
La trajectoire ne ralentit pas, elle s'accélère	Passer de 120 à 600 kilowatts par rack en une seule génération, avec un changement de norme électrique à 800 volts continu, n'est pas un raffinement incrémental, c'est une rupture d'infrastructure que peu d'installations actuelles pourront absorber sans investissement massif.

DIMENSION	ENJEU 2026
La densité extrême concentre la capacité chez un nombre restreint d'opérateurs	– Seuls les acteurs disposant du capital et de l'accès électrique nécessaires peuvent construire des installations neuves calibrées pour cette classe de rack, renforçant la concentration déjà documentée ailleurs dans cette collection sur le marché du calcul d'IA.
Le refroidissement liquide reste un savoir-faire industriel à maîtriser, pas seulement à acheter	± Concevoir, exploiter et entretenir une boucle de refroidissement liquide à cette échelle exige des compétences d'ingénierie que peu d'opérateurs européens ont eu l'occasion de développer à cette densité avant 2025.
Des opérateurs européens documentent déjà des déploiements à cette échelle	+ Des installations en Europe, dont certaines françaises, figurent déjà parmi les références citées pour leurs performances de refroidissement liquide à haute densité, une preuve que la maîtrise européenne de cette technologie n'est pas hypothétique.

La thèse souveraine de cette fiche. La densité extrême de l'IA redistribue les cartes d'un jeu que la seule échelle financière ne suffit pas à gagner : construire un datacenter capable d'accueillir un rack de cent trente kilowatts exige une maîtrise d'ingénierie du refroidissement liquide que la simple capacité d'investissement ne remplace pas. Pour un acteur cloud souverain, c'est à la fois la contrainte la plus lourde de cette nouvelle génération d'infrastructure, et l'occasion de démontrer une compétence d'exécution que l'argent seul ne peut pas acheter du jour au lendemain.

SIX CAS RÉELS À CITER, ET LEUR LEÇON

CAS	CE QUI S'EST PASSÉ	LEÇON POUR UN COMITÉ DE DIRECTION
Le GB200 NVL72, 120 à 132 kW, 1,36 tonne	Le système phare de Nvidia regroupe soixante-douze puces graphiques dans un seul rack intégré, ne proposant plus aucune variante refroidie par air.	Un fabricant peut désormais retirer purement et simplement une option technologique de son catalogue quand elle cesse d'être physiquement viable à l'échelle visée.
La densité moyenne en hausse de 69 % en un an	L'ensemble du parc de datacenters, pas seulement les installations dédiées à l'IA, voit sa densité électrique moyenne bondir sur une seule année.	Une tendance de fond ne se limite jamais aux seuls acteurs de pointe ; elle finit par redéfinir la norme de l'ensemble du secteur.
Le plafond physique de l'air, autour de 40 à 50 kW	Au-delà de ce seuil documenté, aucune configuration de refroidissement par air ne parvient plus à évacuer la chaleur produite, quelle que soit la sophistication de sa conception.	Certaines limites ne sont pas des choix d'ingénierie à optimiser, ce sont des plafonds physiques à respecter, un point que toute feuille de route doit intégrer sans négociation possible.
La feuille de route 2027, environ 600 kW par rack	La génération suivante de racks vise une densité dix fois supérieure au GB200 NVL72, avec un changement de norme d'alimentation électrique à 800 volts continu.	Une infrastructure construite aujourd'hui pour la densité actuelle risque d'être obsolète avant la fin de son amortissement comptable habituel.
Le rack vendu comme un système intégré à trois millions de dollars	Le GB200 NVL72 n'est plus acheté comme un ensemble de composants séparés, mais comme un produit industriel unique, préassemblé et testé en usine.	Le modèle d'achat de l'infrastructure informatique se rapproche de celui d'un équipement industriel lourd, avec ses propres logiques de financement et de cycle de vie.
Des opérateurs européens déjà référencés sur le refroidissement liquide haute densité	Plusieurs installations européennes, dont certaines françaises, figurent parmi les références de performance citées pour ce type de refroidissement.	La maîtrise de cette technologie critique n'est pas un monopole d'un seul continent, une réalité à documenter activement plutôt qu'à supposer acquise ou hors de portée.

1 · Faut-il construire neuf pour cette densité, ou tenter de rénover l'existant ?

Pour construire neuf : un bâtiment conçu dès l'origine pour le refroidissement liquide et cette classe de puissance électrique évite les compromis structurels d'une rénovation.

Pour rénover : l'investissement déjà consenti dans les installations existantes, et le délai de construction d'un site neuf, justifient d'explorer d'abord les limites réelles de l'existant.

Position défendable : réserver la construction neuve aux charges de travail les plus critiques et les plus pérennes, tout en explorant des solutions hybrides de rénovation pour les charges à densité intermédiaire, plutôt qu'un choix binaire appliqué uniformément à tout le parc.

2 · L'achat d'un rack intégré à trois millions de dollars est-il une simplification ou une nouvelle dépendance ?

Pour la simplification : un système préassemblé et testé en usine réduit le risque d'intégration et accélère la mise en service par rapport à un assemblage sur site.

Pour la dépendance : ce mode d'achat lie l'acheteur à un fournisseur unique pour l'ensemble du système, matériel, refroidissement et logiciel, bien plus étroitement qu'un assemblage de composants indépendants.

Position défendable : accepter cette dépendance pour les déploiements où la vitesse de mise en service prime, tout en maintenant en parallèle une compétence d'intégration indépendante pour ne jamais dépendre uniquement de ce mode d'achat.

3 · La course à la densité est-elle soutenable au rythme actuel ?

Pour la soutenabilité : chaque génération plus dense produit davantage de calcul par mètre carré, une efficacité globale qui compense la complexité d'infrastructure ajoutée.

Contre : un passage de 120 à 600 kilowatts en une seule génération dépasse la capacité de la plupart des réseaux électriques locaux à suivre, un problème documenté ailleurs dans cette collection sur l'énergie d'un cloudeur.

Position défendable : la densité continuera de croître tant que la demande de calcul d'IA le justifiera, mais son rythme réel sera arbitré autant par la disponibilité du réseau électrique que par la seule volonté des fabricants de puces.

MATRICE DES RISQUES DES SERVEURS ET DES BAIES

RISQUE	CRITICITÉ	MANIFESTATION
Sous-dimensionner un projet en supposant un refroidissement par air	Critique	Planifier l'accueil d'un rack d'IA dense dans une installation conçue pour un refroidissement par air, un plafond physique documenté à 40-50 kW que la densité visée dépasse largement.
Ignorer le coût de rénovation électrique nécessaire	Élevé	Ne pas intégrer le coût de mise à niveau de l'alimentation électrique du bâtiment dans le budget total d'un projet de densité élevée.
Confondre achat de composants et achat de système intégré	Moyen	Négocier un rack d'IA comme un ensemble de composants indépendants, alors qu'il s'agit d'un système industriel unique vendu et garanti comme tel.
Ne pas anticiper l'obsolescence accélérée de l'infrastructure	Élevé	Amortir un investissement de densité actuelle sur un cycle long, sans intégrer le risque de rupture technologique déjà annoncé pour 2027.
Sous-estimer la compétence d'ingénierie liquide nécessaire	Moyen	Traiter le refroidissement liquide comme un simple achat d'équipement, sans investir dans la compétence opérationnelle nécessaire à son exploitation durable.

- Q1** **Nos installations actuelles connaissent-elles leur plafond réel de densité par rack ?**
Documenter précisément cette limite avant tout engagement sur un projet de densité élevée.
-
- Q2** **Avons-nous chiffré le coût complet d'accueil d'un rack haute densité, électricité comprise ?**
Intégrer la rénovation électrique et de refroidissement dans tout budget, pas seulement le prix catalogue du matériel.
-
- Q3** **Distinguons-nous achat de composants et achat de système intégré dans nos négociations ? Adapter la stratégie contractuelle à la nature réelle du produit acquis.**
-
- Q4** **Notre horizon d'amortissement intègre-t-il le risque de rupture 2027 ?** Ajuster la durée d'amortissement prévue à la lumière de la trajectoire de densité déjà annoncée.
-
- Q5** **Investissons-nous dans la compétence d'ingénierie liquide, pas seulement dans l'équipement ?** Former ou recruter les compétences nécessaires à l'exploitation durable de cette technologie.
-
- Q6** **Suivons-nous les références européennes de refroidissement liquide haute densité ?**
Intégrer cette veille dans toute décision de positionnement ou de partenariat technique.

Q1 Qu'est-ce qu'une unité de rack (U), et pourquoi ce format a-t-il duré trente ans ?

Attendu : une unité de hauteur normalisée de quatre centimètres quatre vingt-quatre, permettant d'empiler jusqu'à quarante-deux serveurs dans un rack standard, un gabarit mécanique stable bien avant que la densité électrique ne devienne le facteur limitant.

Q2 Pourquoi le refroidissement liquide devient-il obligatoire au-delà d'un certain seuil ?

Attendu : parce que le refroidissement par air atteint son plafond physique autour de 40 à 50 kilowatts par rack, un seuil que les racks d'IA dédiés dépassent largement, rendant le liquide seul capable d'évacuer la chaleur produite.

Q3 Qu'est-ce que le GB200 NVL72, concrètement ?

Attendu : un rack intégré de Nvidia regroupant soixante-douze puces graphiques et trente-six processeurs, consommant 120 à 132 kilowatts, pesant 1,36 tonne, vendu comme système unique préassemblé à environ trois millions de dollars.

Q4 Qu'est-ce que la redondance N+1 ou 2N à l'échelle d'un datacenter ?

Attendu : une architecture où une capacité électrique ou de refroidissement de secours, parfois dupliquée à l'identique, reste disponible en cas de défaillance de la chaîne principale, prolongeant à l'échelle du bâtiment la logique de double alimentation d'un serveur individuel.

Q5 La densité des racks va-t-elle continuer à augmenter au même rythme ?

Attendu : la feuille de route industrielle vise environ 600 kilowatts par rack dès 2027, une progression que la disponibilité réelle du réseau électrique local pourrait toutefois freiner davantage que la seule volonté des fabricants.

« Notre datacenter peut accueillir n'importe quel rack d'IA moderne »

Sans refroidissement liquide et sans alimentation électrique renforcée, aucune installation conçue pour l'air ne peut héberger un rack dépassant 40 à 50 kilowatts.

« Un rack d'IA s'achète comme n'importe quel serveur, composant par composant »

Les systèmes les plus avancés se vendent comme des unités industrielles intégrées, avec leur propre garantie et leur propre chaîne de responsabilité, jamais assemblés sur site à partir de pièces séparées.

« Le refroidissement liquide est une option pour les cas les plus extrêmes seulement »

Au-delà du seuil physique de l'air, ce n'est plus une option parmi d'autres, c'est devenu la seule solution techniquement viable, un fait physique documenté, pas un choix de conception.

« La densité actuelle restera stable encore plusieurs années »

La feuille de route publique vise déjà une densité dix fois supérieure dès 2027, une trajectoire annoncée, pas une hypothèse.

« L'Europe n'a pas la maîtrise du refroidissement liquide à cette échelle »

Plusieurs installations européennes figurent déjà parmi les références documentées de performance sur cette technologie précise.

Unité de rack (U)	Unité de hauteur normalisée valant quatre centimètres quatre vingt-quatre, servant à désigner la taille verticale d'un serveur (1U, 2U, 4U).
Rack	Armoire normalisée destinée à recevoir des serveurs empilés, comptant traditionnellement jusqu'à quarante-deux unités U.
Redondance N+1	Architecture où une unité de capacité supplémentaire est disponible au-delà du strict nécessaire, garantissant la continuité en cas de défaillance d'une unité.
Redondance 2N	Architecture où l'intégralité de la capacité est dupliquée à l'identique, offrant le niveau de continuité le plus élevé, au prix d'un coût d'infrastructure doublé.
PDU	<i>Power Distribution Unit</i> . Unité de distribution électrique répartissant l'alimentation depuis la source vers les différents équipements d'un rack.
Refroidissement par air	Méthode d'évacuation de la chaleur par circulation d'air, atteignant son plafond pratique autour de 40 à 50 kilowatts par rack.
Refroidissement liquide direct sur puce (direct-to-chip)	Technique où des plaques froides raccordées à une boucle d'eau sont fixées directement sur les composants les plus chauds, la solution dominante au-delà de 50 kilowatts par rack.
Refroidissement par immersion	Technique où les composants sont directement immergés dans un liquide diélectrique, capturant la quasi-totalité de la chaleur produite.

PUE	<i>Power Usage Effectiveness.</i> Indicateur d'efficacité énergétique d'un datacenter, rapportant l'énergie totale consommée à celle réellement utilisée par le calcul.
GB200 NVL72	Système rack-échelle de Nvidia regroupant soixante-douze puces graphiques Blackwell et trente-six processeurs Grace, vendu comme unité industrielle intégrée à refroidissement liquide obligatoire.
800 volts continu (800VDC)	Norme d'alimentation électrique annoncée pour la prochaine génération de racks à très haute densité, remplaçant les standards de tension actuels.
Refroidissement direct sur puce	Plaques froides parcourues de liquide posées sur les composants les plus chauds, processeurs et cartes graphiques ; norme de fait des baies d'IA.
Refroidissement par immersion	Serveurs plongés dans un fluide qui ne conduit pas l'électricité et capte presque toute leur chaleur.
WUE	<i>Water Usage Effectiveness.</i> Litres d'eau consommés par kilowattheure d'électricité, complément indispensable du PUE.

La seule question vraiment stratégique à poser à votre COMEX :

« Connaissons-nous le plafond de densité réel de chacune de nos installations, et savons-nous exactement ce qu'il faudrait investir pour le dépasser ? *Le rack qui organisait sagement trente ans de datacenter pèse aujourd'hui plus d'une tonne et ne se refroidit plus qu'à l'eau. Ignorer cette bascule physique, ce n'est pas prendre un risque technique mineur, c'est planifier un investissement sur des hypothèses déjà dépassées par le marché.* »

Du calcul au stockage partagé

Cette fiche a posé le serveur et sa baie, l'unité physique du calcul. La fiche suivante, Le stockage en réseau, NAS et baies SAN, traite l'autre grande famille d'équipement de datacenter, celle qui ne calcule pas mais conserve, partagée entre plusieurs serveurs plutôt qu'enfermée dans un seul châssis.

Sources principales (septembre 2026)

1. GPU Smith, GPU Rack Power Requirements: Data Center Planning Guide : densité GB200 NVL72, feuille de route 2027, seuil de l'air.
2. ModulEdge, Nvidia Blackwell Explained: Data Center Impact : poids 1,36 tonne, absence de variante air, spécifications comparées.
3. Gainam, Liquid Cooling for GPU Clusters: The Direct-to-Chip Decision Framework : densité moyenne +69 %, comparatif des méthodes de refroidissement.
4. How To Store Electricity, Nvidia GB200 NVL72 Liquid Cooling Data Centre 2026 : prix par rack, références européennes de refroidissement liquide.
5. ServNet UK, GPU Rack Power Density 2026 : trajectoire 2027, norme 800VDC.
6. Fiches Économie d'un cloudeur, Énergie et Les cartes graphiques de la collection.
7. Upsite, The State of Data Center Cooling in 2026 (rapport AFCOM) : taux d'adoption du refroidissement liquide.
8. Obelinf, Direct to Chip vs Rear Door vs Immersion Cooling (2026) et Budlong, Data Center Cooling Systems Compared (2026) : techniques, densités, limites du PUE.
9. Data Center Cooling Economics 2026 : consommation d'eau selon la technique.
10. DCD et Next (2023).

Briefing synthétisé pour audience C-level · Informations arrêtées au 19 septembre 2026, à revérifier · Ce document est une note de synthèse, non un conseil technique ou opérationnel.