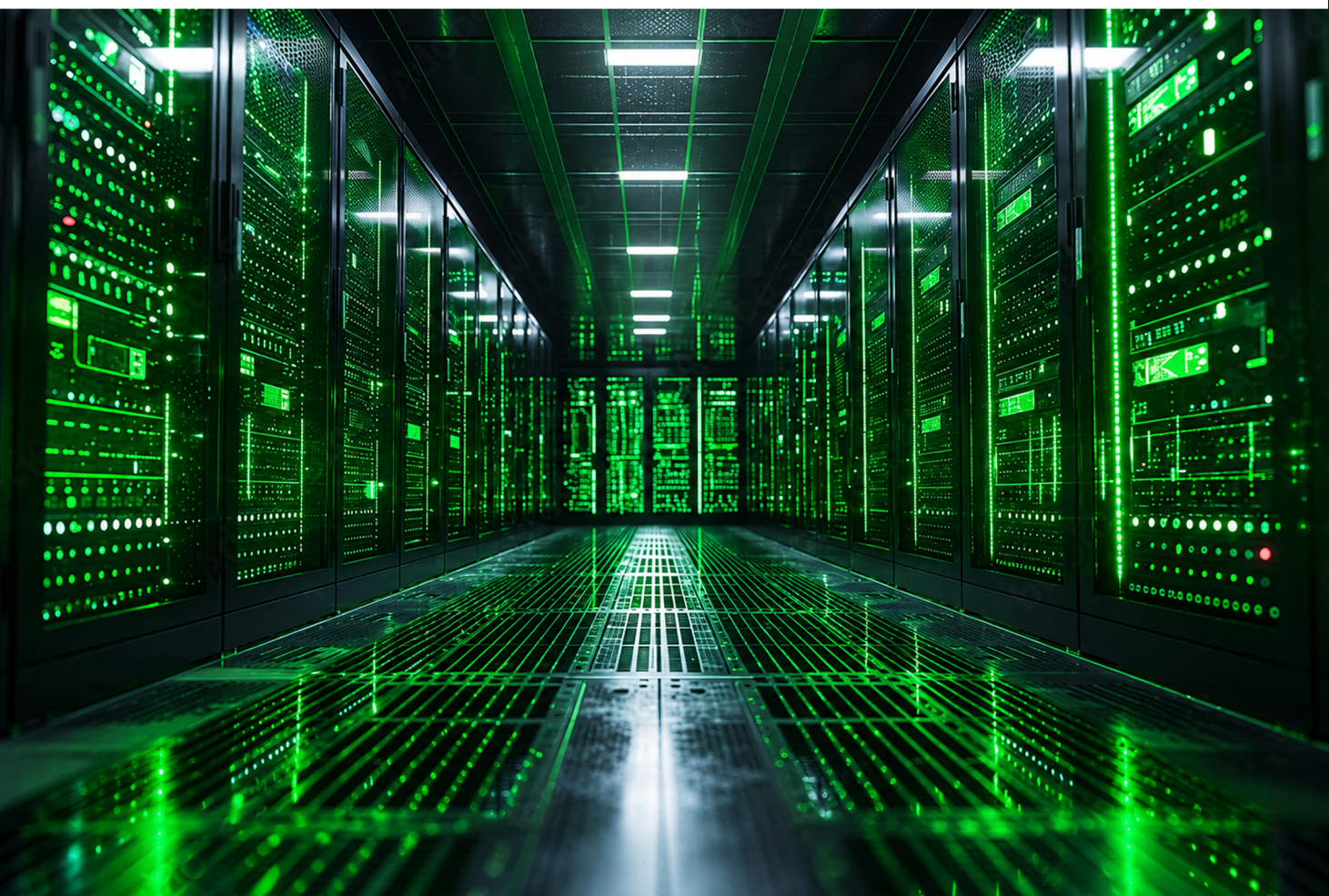




REFROIDISSEMENT DES DATA CENTERS

*Vers des technologies plus
responsables de l'environnement*

SOMMAIRE

Contexte	P 3
1 - Indicateurs de performance	P 5
1.1 - Power Usage Effectiveness	P 5
1.2 - Data Center Energy Management (DCEM) :.....	P 7
2 - Panorama des techniques de refroidissement	P 9
2.1 - Système classique : refroidissement par l’Air.....	P 9
2.2 - Free Cooling	P 12
2.2.1 - Utilisation du vecteur air	P 13
2.2.1.1 - Système ouvert	P 13
2.2.1.2 - Système isolé	P 14
2.2.2 - Utilisation du vecteur eau	P 15
2.3 - Refroidissement liquide	P 17
2.3.1 - Refroidissement liquide indirect	P 17
2.3.2 - Refroidissement liquide direct	P 19
2.3.2.1 - Par immersion	P 19
2.3.2.2 - Par Pulvérisation	P 24
2.3.2.3 - Conclusion refroidissement direct	P 25
3 - Valorisation de la chaleur fatale	P 28
4 - Déploiement des nouvelles technologies de refroidissement	P 29
Conclusion	P 30

CONTEXTE

L'explosion des usages numériques, du calcul haute performance (HPC) et plus récemment de l'intelligence artificielle (IA) entraîne une augmentation rapide des besoins en infrastructures de données. Derrière cette croissance se cache toutefois une réalité souvent méconnue : un data center moderne consomme presque autant d'énergie pour se refroidir que pour alimenter ses serveurs.

LE REFROIDISSEMENT AU CŒUR DES ENJEUX ÉNERGÉTIQUES DES DATA CENTERS

Avec 395 centres de données recensés à ce jour et une capacité installée qui atteignait 510 MW fin 2023, puis 714 MW fin 2024.

La France se positionne au 3^{ème} rang Européen derrière Le Royaume-Uni et L'Allemagne

En France entre 2022 et début 2026, le nombre de centres de données a progressé d'environ 25%. Depuis l'annonce du plan IA de 109 milliards d'euros début 2025, le développement des centres de données est considéré comme une infrastructure stratégique. Si l'on se projette jusqu'à 2030, on peut s'attendre à passer la barre des 500 centres de données pour une capacité estimée de 2,3 GW ^[2]. La maîtrise énergétique du refroidissement devient désormais un enjeu stratégique pour ces infrastructures.

Aujourd'hui, environ 50 % de l'énergie d'un data center est directement utilisée par les équipements informatiques (IT), tandis que près de 40 % est consacrée au refroidissement des installations. Autrement dit, pour 1 kW utilisé pour le calcul, il faut pratiquement 1 kW supplémentaire pour évacuer la chaleur produite ^[3]. Cette contrainte thermique est devenue l'un des principaux verrous énergétiques des infrastructures numériques. Avec l'augmentation de la densité de puissance des serveurs, notamment liée aux charges IA et HPC, les technologies historiques de refroidissement par air atteignent progressivement leurs limites. Les besoins en ventilation, en production de froid et parfois en consommation d'eau deviennent considérables.

Dans ce contexte, le refroidissement des data centers n'est plus seulement une problématique technique : il devient un enjeu énergétique, économique et environnemental majeur. De nouvelles approches émergent alors pour limiter la consommation des groupes froids, améliorer le PUE des installations et valoriser la chaleur produite : free cooling, refroidissement liquide, immersion cooling ou encore récupération de chaleur fatale.

Cette note présente les principales technologies de refroidissement actuellement utilisées ou en développement dans les data centers, ainsi que leurs avantages, limites et perspectives d'évolution.

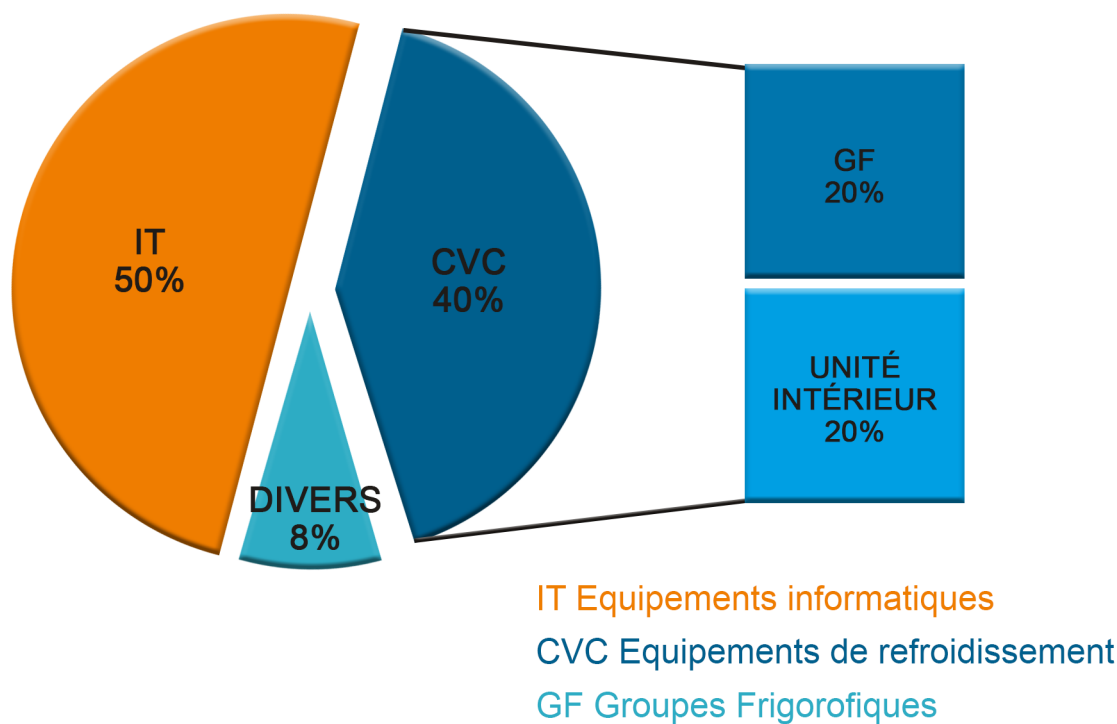


FIGURE 1: RÉPARTITION DES POSTES CONSOMMATEUR D'ÉNERGIE DANS UN DATA CENTER ^[1]



1. INDICATEURS DE PERFORMANCE

Il existe plusieurs indicateurs de performance afin de caractériser les data centers : le Power Usage Effectiveness (PUE), initialement créé en Amérique utilisé et le Data Center Energy Management (DCEM), créé en Europe. Le PUE est l'indicateur le plus couramment utilisé de nos jours.

1.1 - POWER USAGE EFFECTIVENESS

Le PUE se définit de la façon suivante :

$$PUE = \frac{\text{Energie totale consommée par le data center}}{\text{Energie totale consommée par les équipements informatiques}}$$

Cet indicateur compare sur une période d'un an l'énergie utilisée uniquement par les baies (serveurs, stockage de données etc...) et l'énergie nécessaire au bon fonctionnement global de l'installation (équipements informatiques, refroidissement etc...). Plus cet indicateur est proche de 1 plus le data center est efficace.

Beaucoup de critiques sont faites à l'encontre de cet indicateur, critiques résumées dans l'article de Yuventi et al. ^[4] :

- L'évaluation de l'énergie totale consommée par le data center peut être sujette à discussion.
- Au moment de la conception du data center les plans sont faits de façon à atteindre une certaine valeur de PUE. Or, le dimensionnement des installations se fait généralement dans le pire des cas ce qui mène à des systèmes surdimensionnés. Ce surdimensionnement peut générer un fonctionnement du système à un rendement moins important pouvant ainsi augmenter le PUE final de l'installation.
- Il ne prend pas en compte la réutilisation de l'énergie des data centers dans le cas d'une installation revalorisant la chaleur.
- La non-prise en compte de la situation géographique. Un data center situé dans les régions froides pourrait être considéré plus efficace qu'un data center dans une région chaude sans que ce soit réellement le cas.

Une normalisation de l'utilisation du PUE a été effectuée dans la norme ISO/IEC 30134-2 afin d'en améliorer l'utilisation. Des valeurs cibles ont été définies dans la Fiche Valeurs absolues 2030 – « Salles serveurs et centres d'exploitation informatique » du ministère de la transition écologiques :

Sous-catégorie	Surface salle IT	PUE-climats «froids»	PUE-climats «chauds»
Local serveur	< 20m ²	2,00	2,00
Salle serveur	20m ² à 100 m ²	1,80	1,80
Mini data center	100m ² à 500 m ²	1,60	1,72
Petit data center	500m ² à 1000 m ²	1,60	1,72
Moyen data center	1000m ² à 5000 m ²	1,40	1,51
Grand data center	5000m ² à 10000 m ²	1,40	1,51
Très grand data center	> 10 000 m ²	1,20	1,29

D'autres indicateurs, découlant du PUE, peuvent néanmoins permettre de prendre en compte l'empreinte énergétique du data center. Comme par exemple, le CUE (Carbon Usage Effectiveness) qui inclut la mesure des effets de serre générés ou le GEC (Green Energy Coefficient) qui mesurent la part d'énergies renouvelables entrant en ligne de compte dans la consommation globale du data center.

Un autre indicateur pouvant être utilisé est le ERE (Energy Reuse Effectiveness).

L'ERE est identique au PUE mais permet de prendre en compte la revalorisation de l'énergie en soustrayant au numérateur la quantité d'énergie revalorisée des data center :

$$ERE = \frac{E_{IT} + E_{pertes} + E_{froid} + E_{eclairage} + \dots - E_{revalorisé}}{E_{IT}}$$

1.2 DATA CENTER ENERGY MANAGEMENT (DCEM)

Un autre indicateur de performance pouvant être utilisé est le DCME. Cet indicateur a été créé par l'ETSI (European Telecommunications Standards Institute) et est considéré comme plus complet que le PUE, notamment car il prend en compte la revalorisation de l'énergie : cet indicateur est le KPI_{DCEM} (key performance indicator). Cet indicateur est composé de deux indicateurs le DC_G (Data center Gauge) et le DC_p (Data center performance) [5].

Le DC_G est un indicateur de la limite de consommation énergétique d'un data center en fonction de sa taille, ce facteur dépend uniquement du KPI_{EC} correspondant à l'énergie totale consommée par le data center. Il permet de définir les performances attendues en termes de DC_p du data center.

FIGURE 2: «TAILLES» DE DATA CENTER ET INDICATEUR D'EFFICACITÉ ÉNERGÉTIQUE

DC _G	KPI _{EC} range
S	KPI _{EC} ≤ 1 GWh
M	1 GWh < KPI _{EC} ≤ 4 GWh
L	4 GWh < KPI _{EC} ≤ 20 GWh
XL	KPI _{EC} > 20 GWh

(a) "Tailles"

Table 2: Default Classes of DC_p

DC commissioning date	since 2005 (see note)		before 2005 (see note)	
	DC_p		DC_p	
Class	≥	<	≥	<
A		0,70		1,00
B	0,70	1,00	1,00	1,40
C	1,00	1,30	1,40	1,70
D	1,30	1,50	1,70	1,90
E	1,50	1,70	1,90	2,10
F	1,70	1,90	2,10	2,30
G	1,90	2,10	2,30	2,50
H	2,10	2,40	2,50	2,70
I	2,40		2,70	

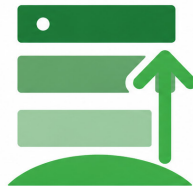
NOTE: Year of Kyoto Protocol entering into force.



ENERGY EFFICIENT



GREEN CERTIFICATE



GREEN DATA CENTER

Le DCp est calculé à partir de différents indicateurs d'efficacité énergétique (KPI) suivant l'équation :

$$DCP = KPI_{TE} * (1 - W_{REUSE} * KPI_{REUSE}) * (1 - W_{REN} * KPI_{REN})$$

Avec W_{reuse} et W_{REN} des facteurs de correction compris entre 0 et 1 et dépendant de la taille du data center. En général la valeur de 0.5 est prise par défaut.

Les facteurs KPI, exprimés en GWh, sont calculés sur une année et correspondent à : KPI_{TE} : idem que le PUE

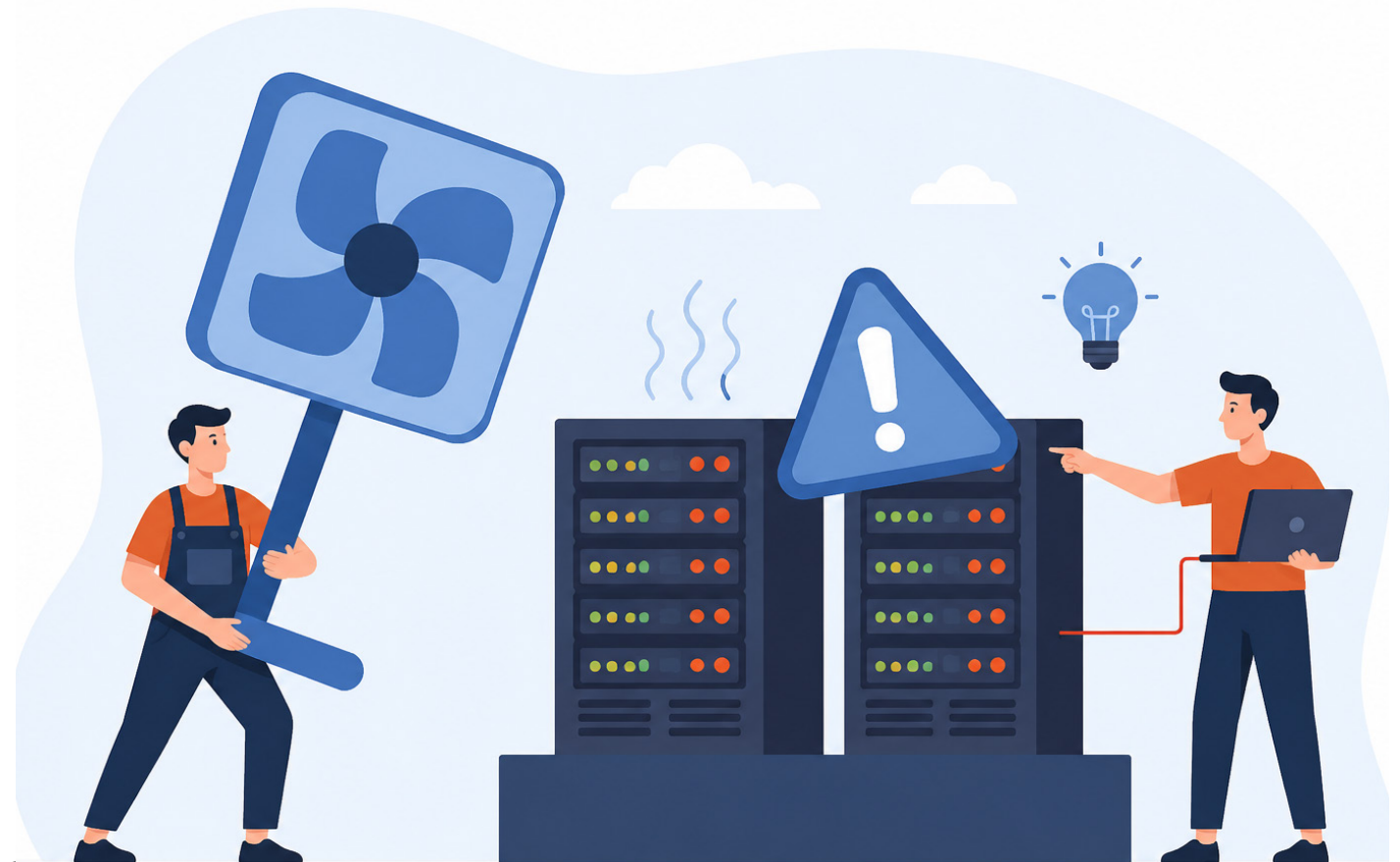
KPI_{REUSE} : fraction de l'énergie réutilisée par le data center sur l'énergie totale consommée

$$KPI_{REUSE} = \frac{EC_{REUSE}}{EC_{DC}}$$

KPI_{REN} : fraction de l'énergie provenant d'ENR sur l'énergie totale consommée par le data center

$$KPI_{REN} = \frac{EC_{REN}}{EC_{DC}}$$

L'avantage immédiat de cet indicateur est qu'il prend donc en compte la source d'énergie ainsi que la possible revalorisation d'énergie. Il ne corrige cependant pas tous les défauts du PUE en ne prenant pas en compte l'influence du climat par exemple.



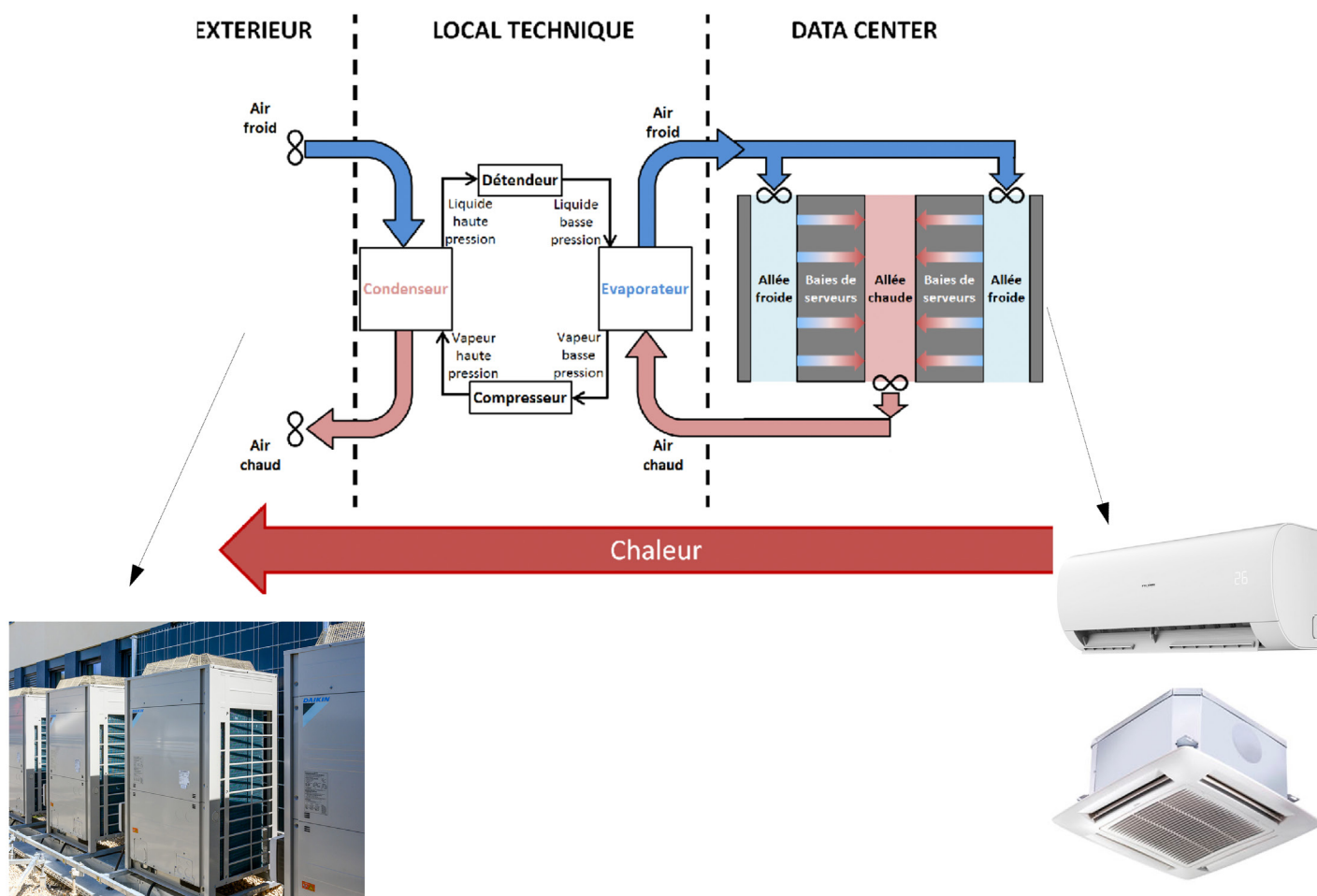
2. PANORAMA DES TECHNIQUES DE REFROIDISSEMENT

2.1 - SYSTÈME CLASSIQUE : REFROIDISSEMENT PAR L'AIR

La technique la plus répandue et la plus classique consiste à alimenter le local en air frais en utilisant un groupe froid. Deux systèmes peuvent être utilisés pour alimenter ce groupe froid.

- La détente directe, technologie principalement utilisée pour les data-center de petite à moyenne taille (max. 100/150 kW). Le principe de fonctionnement est détaillé dans le schéma ci-dessous.

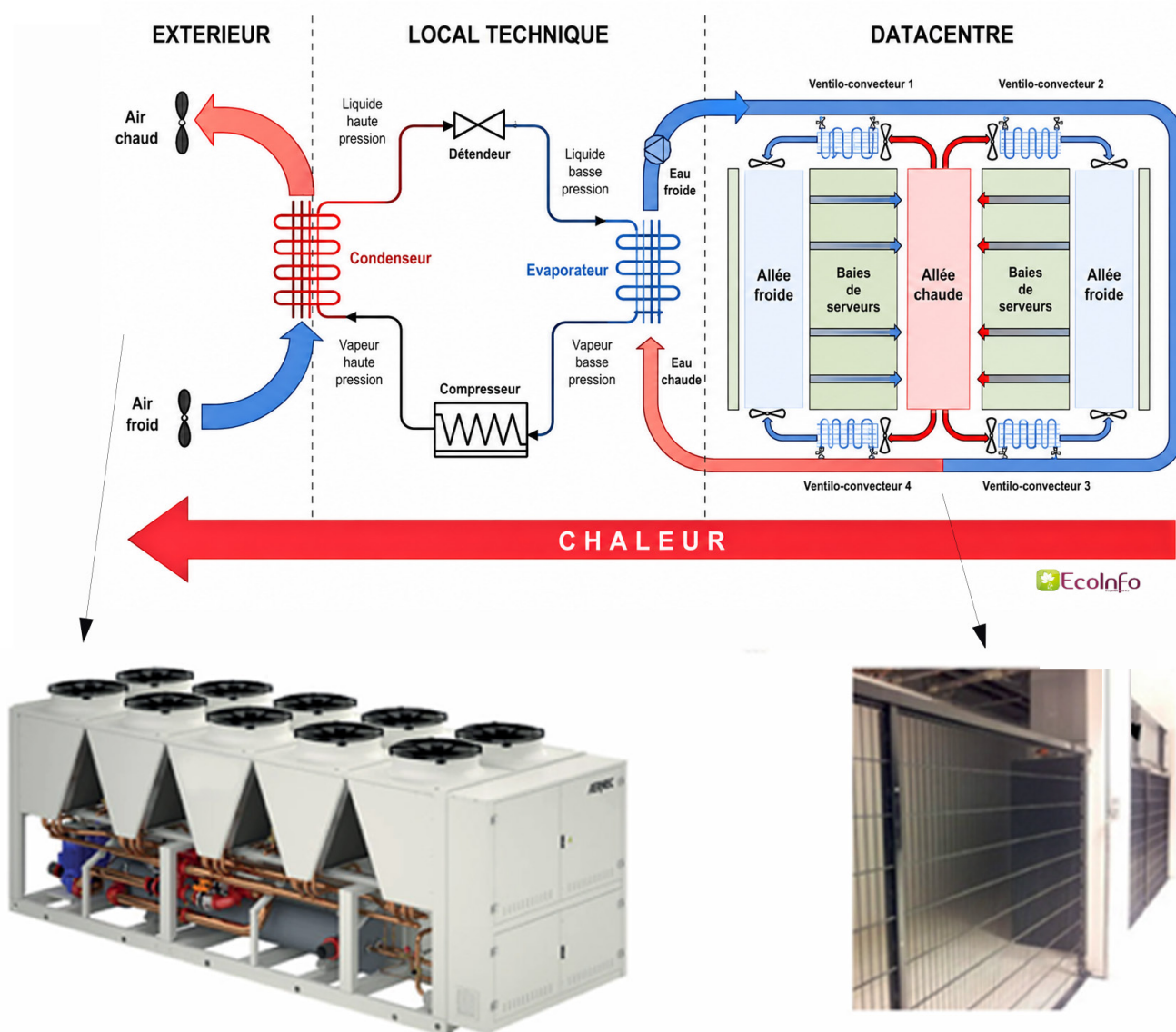
FIGURE 3 : SCHÉMA DE FONCTIONNEMENT D'UNE INSTALLATION À DÉTENTE DIRECTE AIR-AIR



Le système air-air est efficace dans les salles de petite taille type salles informatiques dans bâtiment tertiaire.

Système avec production d'eau glacée par condensation à air ou à eau

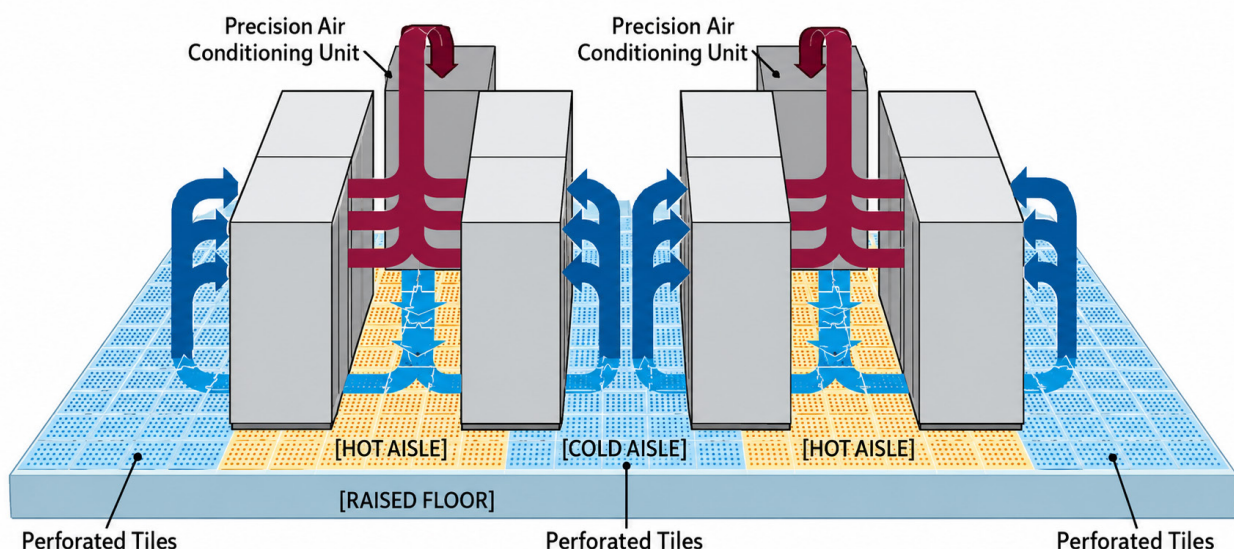
FIGURE 4 : SCHÉMA DE FONCTIONNEMENT AVEC PRODUCTION AIR-EAU



Dans les deux cas, le groupe froid va produire de l'eau glacée à environ 7°C. Le système air-eau est refroidi par l'air extérieur à partir de ventilateurs alors que pour le système eau-eau, il existe un réseau supplémentaire : le réseau de rejet de chaleur qui transfère l'eau chaude en sortie de condenseur vers le système de refroidissement que sont les aéroréfrigérants. Ainsi, un groupe froid eau-eau peut être installé dans un local technique fermé alors qu'un système air-eau doit obligatoirement se trouver en extérieur. Le plus souvent, ils sont installés sur les toits.

Dans les deux cas, l'air froid produit sera utilisé pour refroidir les salles de serveurs. L'arrangement des salles consiste à surélever le plancher et à faire circuler le flux d'air frais sous le sol pour le faire sortir par des ouvertures et le faire passer dans les racks de serveurs. L'air chaud (réchauffé par les serveurs) est ensuite prélevé en partie haute pour être refroidi à l'aide d'un groupe froid et réinjecté sous le plancher.

FIGURE 5 : SCHÉMA CLASSIQUE D'UNE SALLE CONTENANT DES DATA CENTERS ^[6]

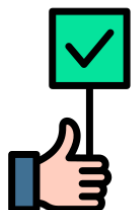


Le refroidissement actif par groupe froid nécessite de maintenir en place une bonne circulation d'air ainsi que de respecter les normes concernant la température ou l'humidité de l'air dictées par l'ASHRAE ou l'ETSI. L'ASHRAE impose une température d'air comprise entre 18 et 27°C avec une humidité relative maximale de 60% et une température de rosée équivalente à 15°C ^[7].

Cette température vise à être réévaluée pour diminuer les consommations. En effet, la plupart des composants fonctionnent à des températures allant jusqu'à 40°C. L'une des pistes pour réduire la consommation des data centers est donc d'augmenter cette température d'air dans les salles.

Ce système est assez simple à mettre en place et permet de conserver une bonne durée de vie des installations. Cependant, la présence du compresseur des groupes froids, des ventilateurs, des souffleurs et des pompes augmente considérablement la consommation énergétique des data centers, les nuisances sonores mais également la fiabilité de l'installation et peut augmenter les besoins en maintenance.

Pour pallier la consommation électrique importante du groupe froid, un autre système a été mis en place utilisant l'environnement extérieur afin de relarguer les calories vers l'extérieur : le free cooling



AVANTAGES :

- Technologie mature ;
- Faible coût initial ;
- Maintenance connue.



LIMITES :

- Efficacité réduite au-delà de 30 kW/rack ;
- Consommation importante des ventilateurs ;
- Difficultés pour les fortes densités thermiques dues aux charges IA.

2.2 - FREE COOLING

Une manière d'améliorer la consommation énergétique des data center est de limiter le fonctionnement des groupes de refroidissement, voire de s'en passer. Pour cela, il est intéressant de pouvoir se servir de l'environnement extérieur comme d'un réservoir où il est possible de relarguer les calories.

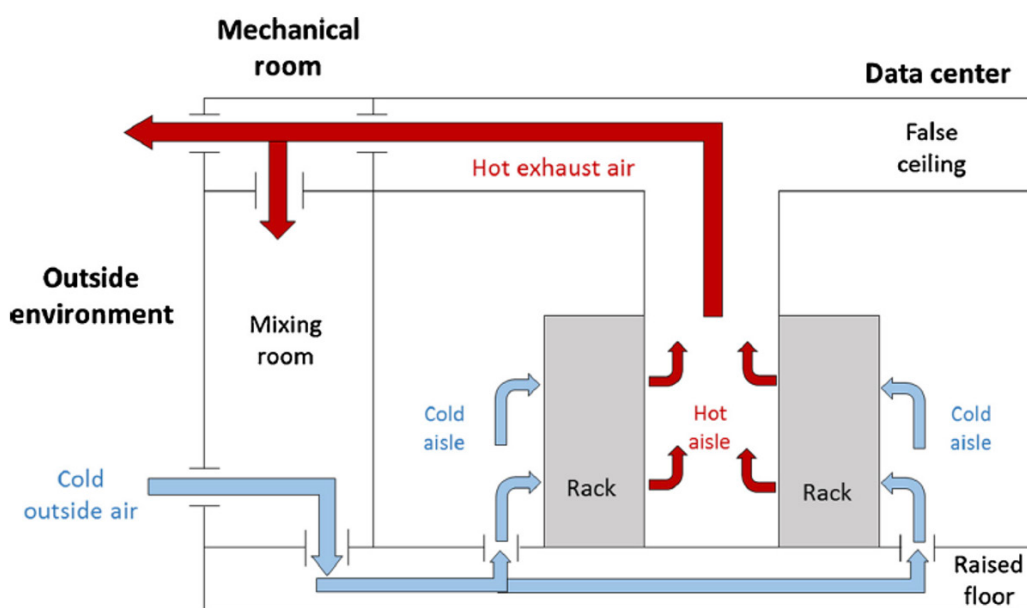
Deux solutions sont utilisées : utilisation de l'air extérieur ou bien l'utilisation d'une réserve de liquide.

2.2.1 - Utilisation du vecteur air

2.2.1.1 - Système ouvert

L'utilisation de l'air extérieur comme moyen de refroidissement consiste comme indiqué sur la Figure 4 à renvoyer une partie de l'air prélevé au data center vers l'extérieur pendant qu'une autre est dirigée vers une zone où elle sera mélangée avec l'air extérieur. La part de l'air recyclé et de l'air extérieur se fait de façon à atteindre la bonne température d'air en entrée de la salle de data center.

FIGURE 6 : FREE COOLING SUR AIR EXTÉRIEUR



Ce genre d'utilisation demande une attention particulière à la valeur d'humidité relative de l'air entrant de façon à éviter tout risque de corrosion. De plus, l'air extérieur peut éventuellement contenir des polluants également nocifs au système.

L'efficacité du free cooling sur l'air extérieur dépend intrinsèquement de la position géographique du data center. L'économie d'énergie dépendant évidemment du nombre d'heures par an où la température de l'air est suffisamment basse pour permettre d'extraire la puissance requise par les serveurs. En France, il a été calculé qu'en moyenne il était possible d'avoir recours au free cooling environ 300 jours /an. Le tout est de pouvoir maintenir une température de soufflage de maximum 27°C d'après les préconisations de l'ASHRAE ^[7].

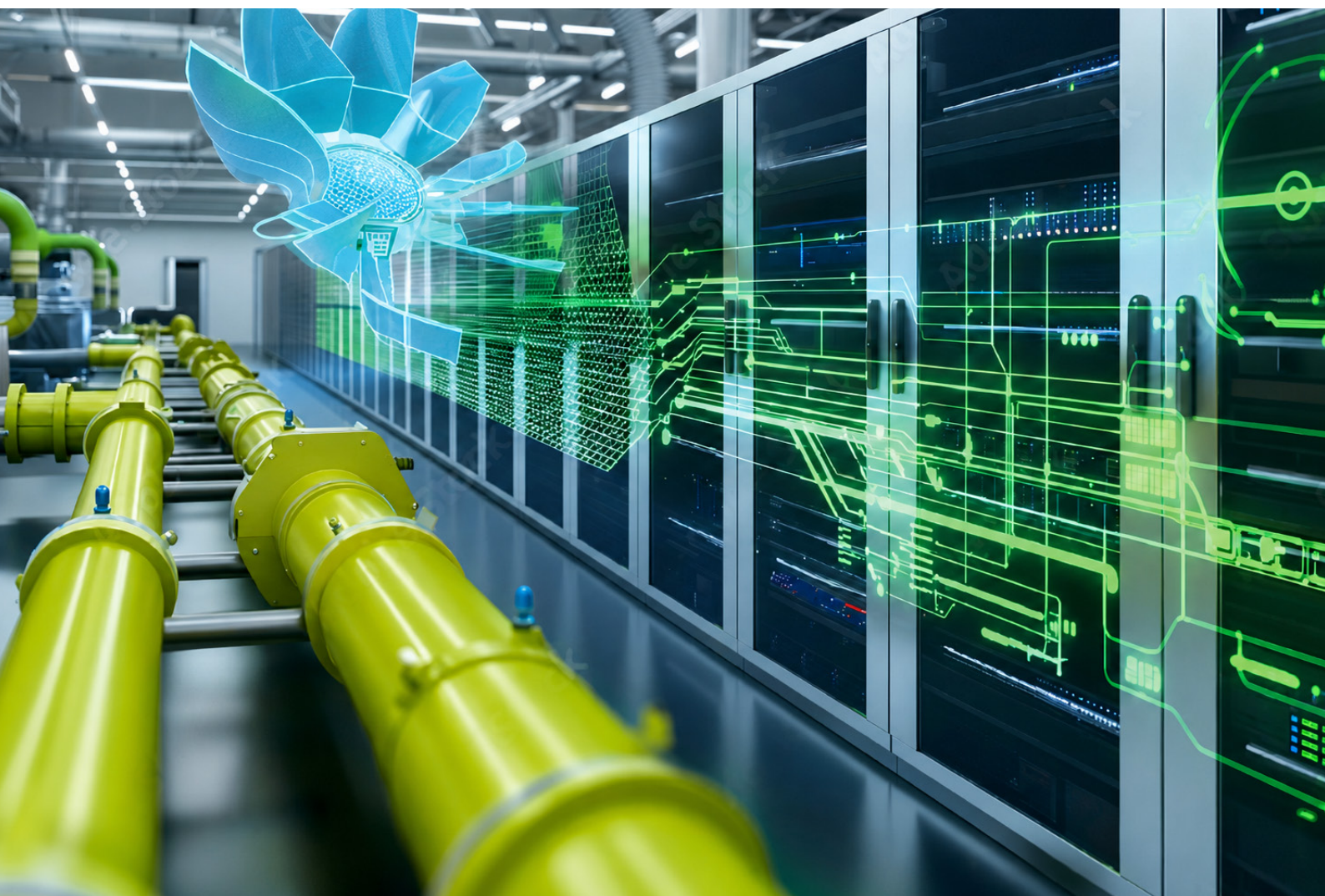
En 2017, un système de free cooling a été installé dans le data center de l'université de Perpignan pour assister le système de climatisation déjà en place. Les résultats ont montré que le free cooling permettait à lui seul d'alimenter en froid le data center pendant 8 mois pleins.

De plus, sur un an la consommation énergétique est passée de 12 MWh consommé par le climatiseur à une consommation totale de 2.2 MWh (répartie entre climatiseur et ventilateur utilisés pour le free cooling).

Dans les faits, le PUE de l'installation passait d'une valeur de 1.6 à une valeur de 1.2 ^[3].

2.2.1.2 - Système isolé

Afin de relarguer la chaleur vers l'extérieur en limitant l'apport de contaminant extérieur ou d'humidité il est également possible d'utiliser un échangeur de chaleur air-air de façon plutôt qu'une zone de mélange. Le risque est uniquement que les contaminants attaquent l'échangeur de chaleur et qu'il faille réaliser des opérations de maintenance supplémentaires.



2.2.2 - Utilisation du vecteur eau

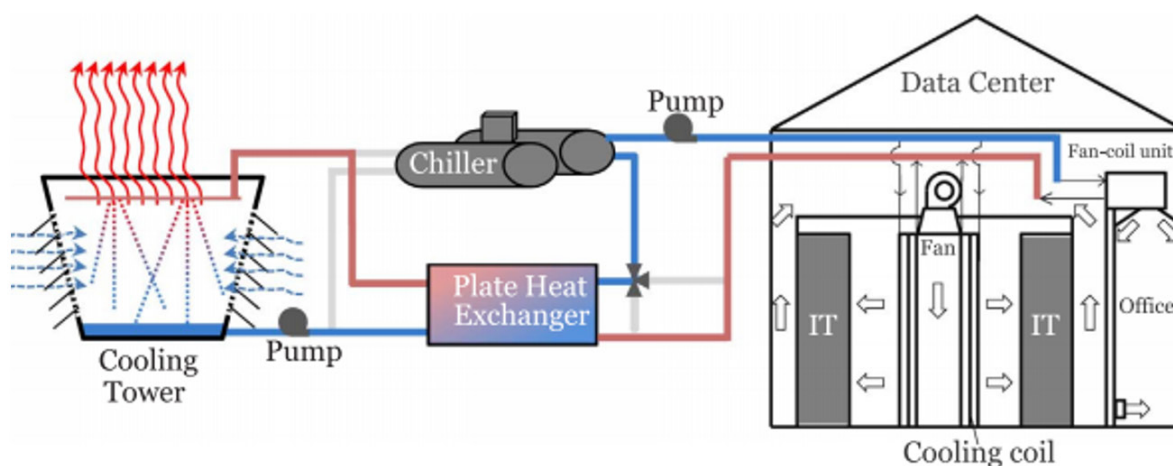
Le free cooling peut également être employé avec de l'eau. L'eau peut être celle d'un cours d'eau, d'un lac etc... et peut être utilisée directement si elle est assez fraîche. Cette eau peut être encore rafraîchie à l'aide d'un échangeur ou via une tour de refroidissement (Figure 7).

On distingue deux types de free cooling à eau :

- Le free cooling à eau direct : l'eau prélevée est directement distribuée dans le circuit d'eau glacée du data center
- Le free cooling à eau indirect : l'eau prélevée passe par un échangeur eau-eau et n'est donc pas utilisée directement dans le circuit d'eau du data center.

Dans les deux cas, le circuit d'eau glacée va servir à refroidir l'air ambiant dans le data center comme indiqué sur le schéma Figure 7.

FIGURE 7 : REFROIDISSEMENT D'UN DATA CENTER PAR FREE COOLING VIA COOLING TOWER



Pour une utilisation optimale du free cooling, il faut donc donner une attention particulière au lieu d'implantation du data center. Le géant du web Microsoft en collaboration avec Naval group a poussé ce concept encore plus loin en développant le projet NATICK. NATICK est une citerne de 12 mètres de long qui embarque 864 serveurs. La citerne est complètement étanche et sous atmosphère azotée afin de se prévenir des éventuels problèmes de corrosion liés au milieu marin. Plongé à 35 mètres sous le niveau de la mer au large des Orcades en Ecosse et alimenté uniquement via des énergies renouvelables, le concept de free cooling immergé propose un PUE de 1,07 pour le premier prototype.

Pour en savoir plus sur ce projet : <https://natick.azurewebsites.net/>

La technologie de free cooling a déjà fait ses preuves dans des domaines autres que les data centers. Par exemple, grâce aux installations de Climespace, société du groupe ENGIE, la climatisation des grands monuments parisiens est assurée depuis 2002 grâce à l'eau de la Seine ! ^[8]

Fulpagare et al. ^[9] indiquent cependant dans leur état de l'art sur la gestion thermique des data centers que le refroidissement par air arrive à ses limites et qu'il est impossible d'augmenter la densité des data centers avec ce système. Une nouvelle orientation pour le refroidissement des data centers est donc le refroidissement par un liquide.

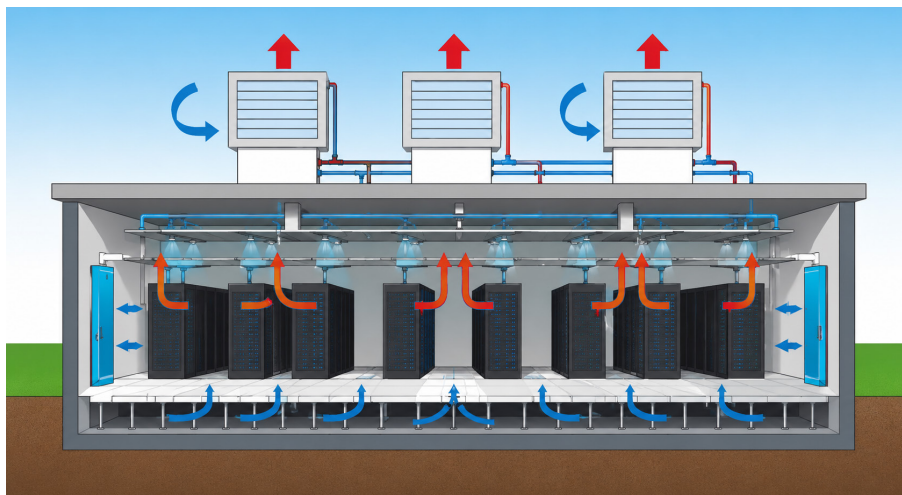
Le free cooling permet d'utiliser :

- L'air extérieur ;
- L'eau d'un cours d'eau ou d'un lac ;
- La fraîcheur du sous-sol.

Bénéfices observés :

- Réduction de 30 à 80 % de la consommation liée au refroidissement ;
- Amélioration significative du PUE ;
- Réduction de l'usage des groupes froids.

Ces solutions sont largement maîtrisées autant du point de vue de la conception que de la maintenance et sont les solutions les plus déployées aujourd'hui. Toutefois, de par leur faible densité et leur faible capacité d'extraction de chaleur, le vecteur air atteint une limite physique pour le refroidissement des data center. Ce refroidissement est d'autant plus limité qu'avec l'émergence des nouveaux microprocesseurs à haute performance, la chaleur à prélever pourrait atteindre 100 W/cm^2 . Or, dans leur étude Saini et al. [10,11] montrent que la limite haute pour l'air serait d'environ 40 W/cm^2 . Pour pallier cette limite technique, l'utilisation d'un fluide plus dense aux propriétés thermiques plus élevées a été étudié.



2.3 - REFROIDISSEMENT LIQUIDE

Il existe deux familles de refroidissement liquide, le refroidissement direct et indirect. Dans le refroidissement direct un fluide diélectrique (souvent de l'huile minérale) est directement mis en contact avec les cartes du serveur. Le refroidissement indirect lui met le fluide de transfert en contact avec le serveur via des tubulures ou plaques froides mais le fluide n'entre jamais en contact direct avec l'élément à refroidir.

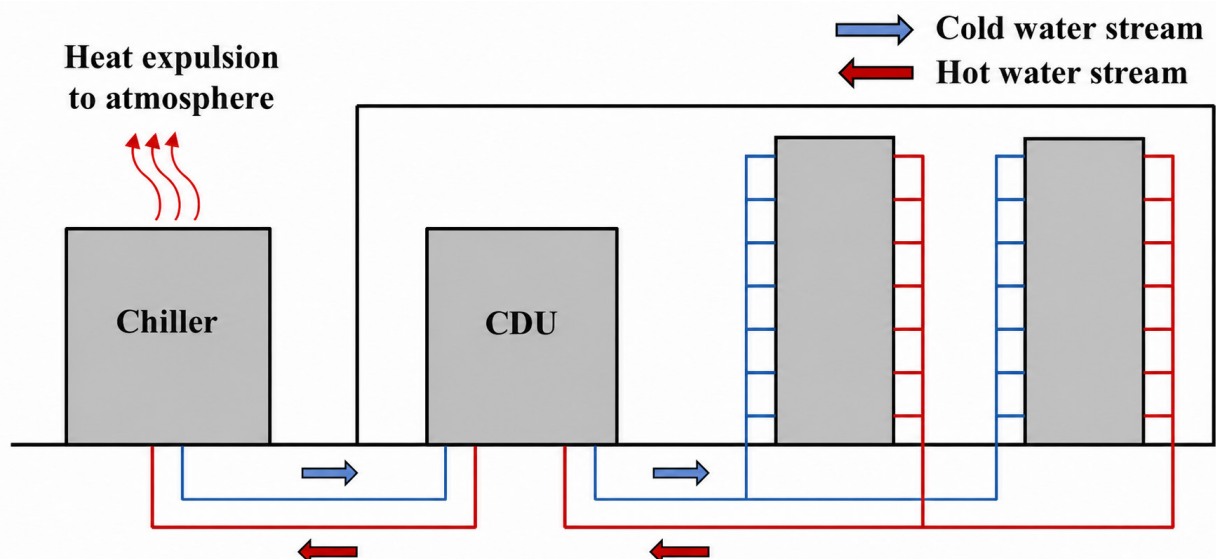
2.3.1 - Refroidissement liquide indirect

Le refroidissement indirect permet d'utiliser les propriétés de transfert thermique d'un fluide sans avoir de contact entre le fluide de transfert et les serveurs. Dans ce cas, les fluides utilisés peuvent être des fluides diélectriques, frigorigènes ou de l'eau. L'usage de l'eau entraînant cependant d'importants risques d'utilisation en cas de fuite.



Le schéma classique de ce système est représenté sur la Figure 8. Le système envoie un fluide refroidi vers les data centers. Le fluide ressort plus chaud après avoir prélevé la chaleur et retourne à l'unité de refroidissement. L'unité de refroidissement consiste en une pompe et un échangeur de chaleur pour amener le fluide à la température de fonctionnement adéquate.

FIGURE 8 : REFROIDISSEMENT LIQUIDE INDIRECT DE SERVEUR

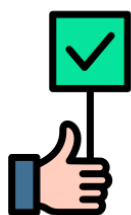


Le fluide dans l'autre partie de l'échangeur est soit un fluide frigorigène si l'installation utilise un groupe froid pour reconditionner le système, ou bien un autre fluide de transfert si une récupération de chaleur est réalisée.

La difficulté technique du refroidissement indirect est de créer les canaux qui vont refroidir les différents serveurs. Leur implantation doit être réalisée correctement sous peine de ne pas prélever correctement la chaleur et d'augmenter considérablement les pertes de charges du système.

Le fluide circulant dans les canaux peut être un fluide avec ou sans changement de phase. Dans le cas d'un fluide sans changement de phase, le refroidissement se manifestera par une élévation en température du fluide. Il est également possible d'utiliser un fluide à changement de phase où cette fois la température dans les canaux sera plus stable. Cette technologie a été employée par IBM dans l'un de ses data centers, où l'entreprise a comparé l'efficacité du refroidissement à air et par liquide indirect ^[12]. Leur système permettait également de récupérer la chaleur via le fluide de transfert. L'utilisation de l'eau permettait de monter la température de refroidissement en entrée de serveur à 60°C et donc de la revaloriser correctement. Le PUE du système était approximativement de 1.15 et l'ERE dépendait de la température d'entrée, cependant elle était toujours en dessous de 0.5.

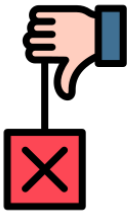
Cette technologie est utilisée depuis 2018 par Facebook qui utilise des membranes innovantes et parfaitement étanches pour refroidir ses serveurs à l'aide de l'eau. On peut également citer Lenovo, numéro 1 mondial du PC qui a choisi un système de liquid cooling, technologie qui achemine à l'eau de tubes de l'eau tiède au plus près des composants ^[13].



AVANTAGES :

- Réduction significative de la consommation énergétique du refroidissement
- Compatible avec les serveurs IA et HPC de très forte densité
- Diminution des besoins en ventilation
- Réduction du bruit
- Amélioration du PUE
- Récupération de chaleur facilitée
- Réduction de l'emprise au sol des infrastructures

LIMITES :



- Coût d'investissement plus élevé
- Complexité hydraulique accrue
- Risque de fuite (même s'il est aujourd'hui très maîtrisé)
- Maintenance nécessitant des compétences spécifiques
- Nécessité d'utiliser des serveurs compatibles

2.3.2 - Refroidissement liquide direct

2.3.2.1 - Par immersion

Le refroidissement par immersion (Immersion Cooling) constitue aujourd'hui l'une des technologies les plus innovantes dans le domaine du refroidissement des data centers. Son principe repose sur l'immersion complète des serveurs informatiques dans un fluide diélectrique, c'est-à-dire un liquide électriquement isolant qui ne provoque pas de court-circuit au contact des composants électroniques.

Contrairement aux technologies conventionnelles utilisant l'air ou aux systèmes Direct-to-Chip qui refroidissent uniquement certains composants spécifiques, le refroidissement par immersion permet d'évacuer directement la chaleur de l'ensemble des équipements électroniques : processeurs, mémoires, cartes réseau, alimentations et autres composants présents dans le serveur.

Cette approche offre des performances thermiques exceptionnelles et répond particulièrement bien aux besoins des infrastructures à très forte densité de puissance telles que les clusters d'intelligence artificielle, les supercalculateurs (HPC) et les applications scientifiques.

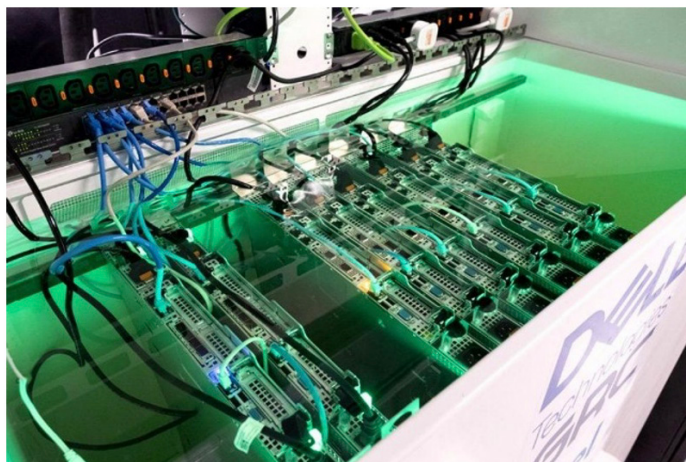
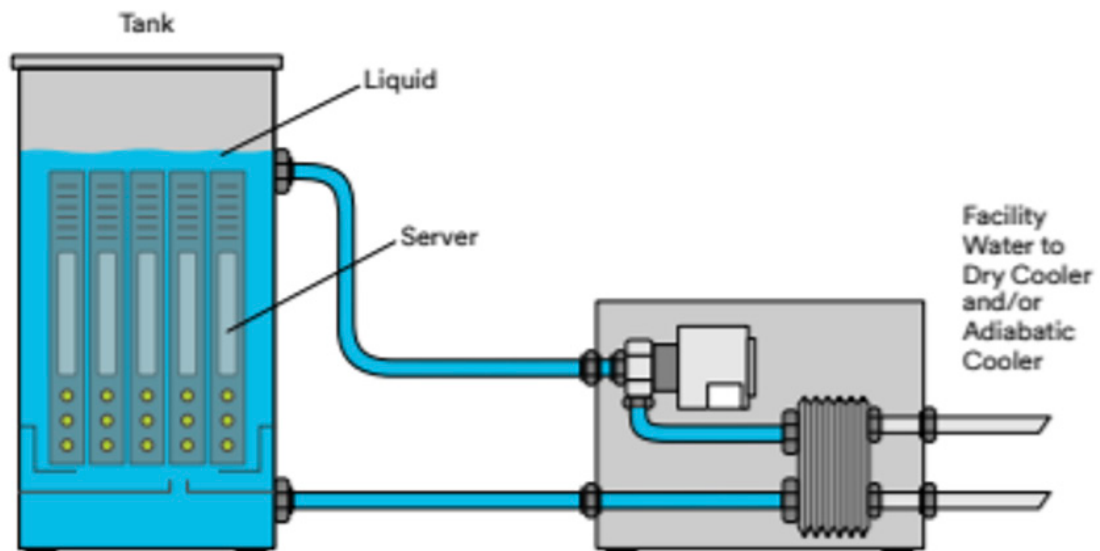
Ces systèmes sont plus simples à mettre en place que ceux du refroidissement indirect, car ici il ne s'agit que d'immerger directement les serveurs. Cette immersion rend le contact entre fluide et serveur optimal et génère moins de pertes de charge que l'indirect. Néanmoins, avec le temps un dépôt peut s'accumuler sur les serveurs et peut entraîner un mauvais vieillissement du câble. Des fluides sont proposés par divers fournisseurs pour se prémunir de ces risques. Cette solution impose l'utilisation de serveurs spéciaux compatibles avec l'immersion.

En plus d'améliorer le refroidissement, la solution par immersion présente d'autres avantages majeurs. En effet, elle permet de limiter le vieillissement des serveurs et de les protéger des contaminants environnementaux tels que la poussière et l'humidité. Les nuisances sonores sont aussi considérablement diminuées dans les data center. La maintenance de ces systèmes directs consiste essentiellement à vérifier qu'il n'y ait pas de pertes de fluide au cours du temps

Refroidissement sensible

Le refroidissement par immersion peut être réalisé de manière sensible. Pour cela, les systèmes sont immergés dans le fluide diélectrique qui est mis en circulation afin d'extraire les calories vers l'extérieur (voir Figure 9 ci-dessous).

FIGURE 9 : REFROIDISSEMENT LIQUIDE DIRECT : SENSIBLE (3M NOVEC)



La chaleur des serveurs augmente progressivement la température du bain. Une pompe assure ensuite la circulation du liquide vers un échangeur thermique externe où les calories sont transférées vers :

- Un circuit d'eau
- Un dry-cooler
- Une tour aéroréfrigérante
- Un réseau de valorisation thermique

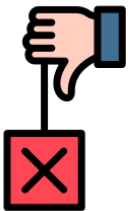
Le fluide refroidi retourne ensuite dans la cuve.

AVANTAGES

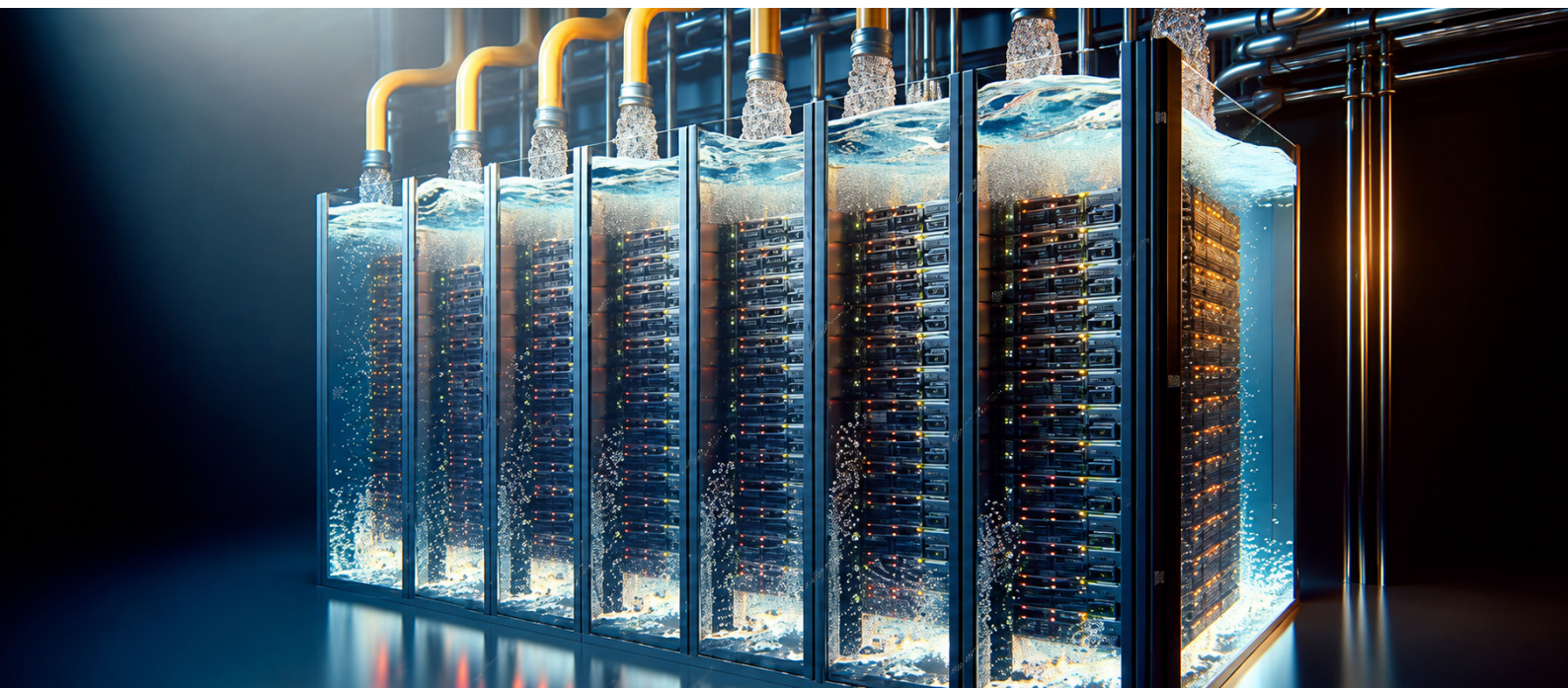


- Technologie relativement simple
- Maintenance facilitée
- Excellente fiabilité
- Faible consommation énergétique
- Absence de changement de phase à gérer
- Moins de corrosion
- Evacuation de chaleur plus facile car niveau de température plus élevé

LIMITES :



- Coût élevé du fluide diélectrique
- Nécessité d'adapté les serveurs
- Maintenance plus complexe (extraction des équipements du bain avant intervention, manipulation de composants recouverts de fluide, temps d'intervention potentiellement plus long)
- Etude préalable sur la compatibilité matériaux et sur la durée de vie
- Gestion des fuites et de la contamination
- Recyclage et fin de vie du fluide

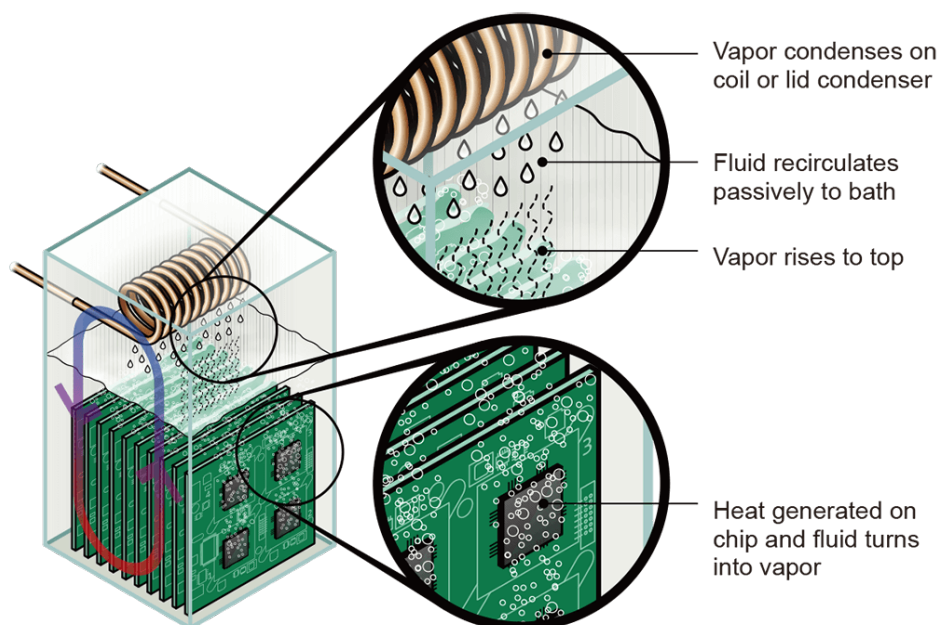


Fluide à changement de phase

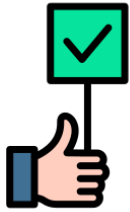
Il est également possible d'utiliser un fluide à changement de phase dans l'immersion cooling. Dans ce cas, le fluide en contact direct avec le serveur monte en température jusqu'à atteindre sa température de changement de phase. Là, il passe sous forme de vapeur et monte dans le haut de la cuve. Un condenseur est positionné en haut de cuve afin de faire se liquéfier à nouveau le fluide qui retombe dans la cuve et ainsi de suite (voir schéma de principe figure 9). Ce type de refroidissement est très efficace car, le changement de phase du fluide permet de mettre la cuve en ébullition entraînant une augmentation des transferts thermiques.

Un paramètre important de ce genre de système est le choix du fluide à changement de phase. En effet, la température de changement de phase doit être bien adaptée à l'utilisation. La liquéfaction du fluide dans le condenseur doit également être correctement dimensionnée afin de ne pas devoir être effectuée par un fluide avec une température trop faible

FIGURE 10 : REFROIDISSEMENT DIRECT PAR FLUIDE À CHANGEMENT DE PHASE [9]



L'immersion cooling est déjà très utilisée dans le domaine des data centers. Le Crédit Agricole, par exemple, utilise la technologie du hollandais Asperistas basée sur le refroidissement liquide par immersion. Un projet pilote a été déployé sur le site de Chartres. Les performances du site pilote sont très prometteuses comme en témoigne l'indicateur d'efficacité énergétique : le PUE est passé de 1,4 à 1,03 (en hiver) et 1,04 (en cas de canicule). Fort de ce succès, la prochaine étape du projet est de remplacer 6000 m² d'espaces blancs refroidis à l'air par deux data centers occupant 150 m² [14] !

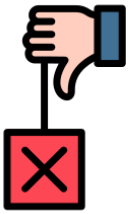


AVANTAGES

Le refroidissement diphasique par immersion présente de nombreux avantages :

- Excellente capacité de refroidissement grâce à l'utilisation de la chaleur latente de vaporisation
- Coefficients d'échange thermique très élevés permettant de refroidir les processeurs et accélérateurs les plus puissants
- Température très homogène des composants, limitant les points chauds ;
- Suppression des ventilateurs des serveurs, réduisant la consommation électrique, le bruit et les opérations de maintenance
- Très faible consommation d'énergie des systèmes auxiliaires, le mouvement du fluide étant principalement assuré par convection naturelle
- Possibilité d'utiliser une eau de refroidissement à température élevée, limitant voire, supprimant le recours à des groupes frigorifiques
- Récupération de chaleur facilitée pour le chauffage de bâtiments ou la production d'eau chaude
- Augmentation importante de la densité informatique, pouvant dépasser 100 kW par rack.





LIMITES :

Cette technologie présente néanmoins plusieurs contraintes :

- Coût d'investissement élevé des cuves, des condenseurs et des fluides diélectriques
- Disponibilité limitée de certains fluides à changement de phase, notamment depuis l'évolution des réglementations environnementales concernant les substances fluorées
- Nécessité de sélectionner un fluide présentant des propriétés thermiques, chimiques et environnementales adaptées
- Maintenance plus complexe nécessitant des procédures spécifiques lors des interventions sur les équipements immergés
- Compatibilité des matériaux (joints, plastiques, connecteurs) avec le fluide utilisé devant être vérifiée
- Nombre encore limité de fournisseurs proposant des solutions industrielles complètes
- Nécessité d'une formation spécifique des équipes d'exploitation.
- Gestion des fuites et de la contamination
- Recyclage et fin de vie du fluide

2.3.2.2 - Par Pulvérisation

Deux types de refroidissement projetés existent aujourd'hui : le refroidissement par spray et projection de jet. Les deux technologies ont le même fonctionnement : un fluide est injecté dans une buse et finement pulvérisé sur le serveur à une plus ou moins forte pression ou via l'ajout d'un gaz porteur. La différence entre ces deux technologies est la vitesse du fluide : la vitesse est plus importante par projection ce qui peut entraîner des problèmes d'érosions de surface et/ou de pertes de pressions. Le refroidissement par spray permet un transfert de chaleur plus uniforme et limite les problèmes d'érosions ou de pression. ^[15,16]

Cette technologie est à ce jour peu employée dans le cas du refroidissement de data centers. Les principaux freins concernent la complexité de réalisation des systèmes de pulvérisation, leur fiabilité à long terme ainsi que les contraintes de maintenance.

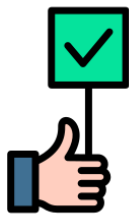
2.3.2.3 - Conclusion refroidissement direct

L'augmentation continue de la puissance des processeurs, accélérateurs graphiques (GPU) et serveurs dédiés à l'intelligence artificielle remet progressivement en question les limites des systèmes de refroidissement à air traditionnels. Les densités thermiques atteintes dans les data centers modernes, pouvant dépasser 100 kW par baie, nécessitent désormais des solutions capables d'évacuer efficacement de très fortes puissances tout en limitant la consommation énergétique des infrastructures.

Le refroidissement liquide direct constitue aujourd'hui l'une des réponses les plus prometteuses à ces nouveaux enjeux. Grâce aux propriétés thermiques des liquides, dont la capacité calorifique et la conductivité sont nettement supérieures à celles de l'air, ces technologies permettent d'évacuer la chaleur au plus près de sa source avec une efficacité bien supérieure aux systèmes conventionnels. Qu'il s'agisse du refroidissement par plaques froides (Direct-to-Chip), de l'immersion monophasique, de l'immersion diphasique ou encore du refroidissement par pulvérisation, chacune de ces architectures répond à des besoins spécifiques en matière de densité de puissance, de coût et de complexité d'exploitation.

Au-delà de leurs performances thermiques, ces technologies ouvrent également de nouvelles perspectives pour améliorer l'efficacité énergétique globale des data centers. En réduisant, voire en supprimant, le recours aux groupes frigorifiques et en facilitant la récupération de chaleur à des températures plus élevées, elles participent directement à la diminution de l'empreinte carbone des infrastructures numériques.

Néanmoins, le choix d'une solution de refroidissement liquide ne peut se limiter aux seules performances thermiques. Les coûts d'investissement, les contraintes de maintenance, la compatibilité des équipements informatiques, le choix des fluides utilisés ainsi que la maturité des différentes technologies constituent autant de critères déterminants. Les solutions les plus performantes, telles que l'immersion diphasique, demeurent aujourd'hui plus coûteuses et nécessitent encore une évolution de l'écosystème industriel avant une adoption massive.



AVANTAGES

PERFORMANCES OBSERVÉES :

- PUE proche de 1 ;
- Suppression quasi totale des ventilateurs ;
- Forte réduction du bruit ;
- Densité de puissance très élevée.
- Favorise le free cooling ou la valorisation de chaleur



Cette technologie est particulièrement adaptée au HPC et à l'IA.

CAS D'USAGE PRIVILÉGIÉS

L'immersion cooling est particulièrement adaptée aux :

- Infrastructures d'intelligence artificielle ;
- Supercalculateurs ;
- Centres de calcul scientifique ;
- Calcul haute performance industriel ;
- Applications blockchain à forte densité énergétique ;
- Data centers modulaires ou conteneurisés.



Pour des centres de données traditionnels de faible densité, le retour sur investissement peut être moins favorable.



LIMITES

CONTRAINTES TECHNIQUES

Malgré ses performances, l'immersion présente plusieurs défis.

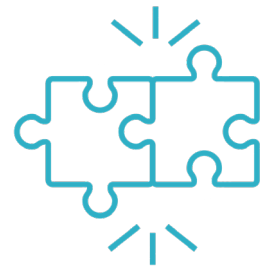


COMPATIBILITÉ MATÉRIELLE

Tous les équipements ne sont pas initialement conçus pour fonctionner immergés.

Certaines adaptations sont nécessaires :

- Retrait des ventilateurs ;
- Modification de certains matériaux ;
- Validation des connecteurs.



MAINTENANCE

Les opérations de maintenance sont différentes de celles d'un data center conventionnel :

- extraction des serveurs hors du bain ;
- manipulation du fluide ;
- nettoyage éventuel des équipements.



COÛT DES FLUIDES

Les fluides diélectriques spécialisés peuvent représenter un investissement important, notamment pour les systèmes diphasiques.



STANDARDISATION

Le marché est encore relativement jeune et les standards industriels continuent d'évoluer.

3. VALORISATION DE LA CHALEUR FATALE

L'un des principaux avantages du refroidissement liquide direct ou indirect est la température relativement élevée du fluide en sortie permise par l'augmentation de la température de fonctionnement.

Dans un système à air classique la température de récupération est souvent comprise entre 25 et 35°C. tandis que dans un système de refroidissement liquide direct ou indirect la température peut atteindre 50 à 65°C.

Cette chaleur produite à des niveaux de température plus intéressants peut donc être utilisée pour :

- chauffer des bureaux ;
- produire de l'eau chaude sanitaire ;
- alimenter un réseau de chaleur urbain ;
- servir de source à une pompe à chaleur.

Le refroidissement liquide devient alors un levier de décarbonation et non plus uniquement un poste de consommation énergétique.



4. DÉPLOIEMENT DES NOUVELLES TECHNOLOGIES DE REFROIDISSEMENT

Un réel besoin freiné par de vrais défis

Comme pour la plupart des problématiques industrielles, aucune technologie de refroidissement des data centers ne constitue aujourd'hui une solution universelle. Chaque architecture présente ses avantages, ses limites et ses contraintes de déploiement.

Les systèmes conventionnels basés sur le refroidissement par air ou le free cooling restent aujourd'hui largement majoritaires grâce à leur maturité technologique, leur simplicité de mise en œuvre et la maîtrise des opérations de maintenance. Toutefois, l'augmentation rapide de la densité énergétique des serveurs, notamment avec les applications IA et HPC, pousse progressivement ces technologies vers leurs limites physiques. Les besoins en ventilation, en échange thermique et en puissance frigorifique, ainsi que les coûts d'exploitation associés, deviennent alors considérables.

Dans ce contexte, les technologies de refroidissement liquide direct ou indirect, et plus particulièrement l'immersive cooling, apparaissent comme des alternatives particulièrement prometteuses. Grâce aux propriétés thermiques des fluides utilisés, ces solutions permettent d'extraire des densités de chaleur beaucoup plus importantes que les architectures traditionnelles à air.

Cependant, malgré leurs performances thermiques, le déploiement à grande échelle de ces technologies est relativement récent. Le marché est aujourd'hui principalement constitué de projets pilotes ou de déploiements ciblés sur des applications spécifiques à forte densité énergétique. Le passage à des infrastructures de grande taille nécessite des investissements importants et les retours d'expérience sur l'exploitation à grande échelle restent encore relativement récents.

Au-delà des aspects technologiques, ces nouvelles architectures soulèvent également des problématiques d'exploitation et de maintenance. Leur mise en œuvre nécessite des compétences spécifiques encore peu répandues. D'autres défis concernent également la compatibilité des équipements, la standardisation des architectures, la durabilité des matériaux au contact des fluides diélectriques ou encore la disponibilité et le coût de certains fluides spécialisés.

Enfin, l'amélioration du retour sur investissement de ces technologies passe souvent par la valorisation de la chaleur fatale produite par les data centers. Cette approche permet de réduire l'impact énergétique global des installations en utilisant la chaleur récupérée pour alimenter des réseaux de chaleur, produire de l'eau chaude sanitaire ou encore servir de source à des pompes à chaleur.

La puissance électrique consommée par un data center étant globalement du même ordre de grandeur que la puissance thermique rejetée, la récupération de cette chaleur représente un potentiel énergétique considérable. Plusieurs installations exploitent déjà cette approche. Le data center du site de Val d'Europe à Marne-la-Vallée alimente par exemple un réseau de chaleur urbain couvrant l'équivalent de 600 000 m². À Paris, la piscine de la Butte-aux-Cailles est partiellement chauffée grâce à la chaleur récupérée sur les serveurs du studio d'animation TeamTO.

Toutefois, le déploiement de ces solutions reste fortement dépendant du contexte local : proximité des besoins thermiques, disponibilité des réseaux, implantation géographique, climat ou encore adéquation entre production et besoin de chaleur.

Ces contraintes doivent souvent être intégrées dès la conception du projet, ce qui limite fortement les possibilités de valorisation sur les infrastructures existantes. La valorisation de la chaleur fatale apparaît ainsi comme un levier particulièrement prometteur pour améliorer l'efficacité énergétique globale des data centers, à condition d'intégrer très tôt cette problématique dans la conception des infrastructures.

CONCLUSION



La croissance rapide des usages numériques, du calcul haute performance et de l'intelligence artificielle place désormais le refroidissement au cœur des enjeux énergétiques des data centers. Longtemps considéré comme un besoin auxiliaire, il devient aujourd'hui un facteur majeur de consommation énergétique, de coût d'exploitation et de complexité technique.

Face à l'augmentation des densités énergétiques, les architectures historiques basées uniquement sur le refroidissement par air atteignent progressivement leurs limites. Les technologies de free cooling, de refroidissement liquide ou d'immersion apparaissent ainsi comme des leviers essentiels pour améliorer les performances thermiques et limiter l'impact environnemental des infrastructures numériques.

Cependant, les performances thermiques seules ne suffisent pas à garantir le déploiement massif de ces solutions. Les contraintes de coût, de maintenance, de disponibilité des compétences, de standardisation ou encore de valorisation de la chaleur jouent également un rôle déterminant dans le choix des architectures de refroidissement.

Le refroidissement des data centers ne constitue donc plus uniquement un enjeu technique. Il devient un sujet global mêlant énergie, exploitation, contraintes industrielles et impact environnemental, appelé à prendre une place croissante dans le développement des infrastructures numériques des prochaines années.

[1] Cécile Digue, 2019, L'IMPACT SPATIAL ET ENERGETIQUE DES DATA CENTERS SUR LES TERRITOIRES, ADEME

[2] [Combien de data centers en France 2026 : 350 sites et 714 MW](#)

[3] Olivier Fruchier, Philippe Egea, 2017 Valorisation énergétique d'un local informatique par free

[4] Jumie Yuventi, Roshan Mehdizadeh, A Critical Analysis of Power Usage Effectiveness and Its Use as Data Center Energy Sustainability Metrics 2013

[5] Operational energy Efficiency for Users (OEU); Global KPIs for ICT Sites

[6] Chayan Nadjahi, Hasna Louahlia, 2018, Stéphane Lemasson A review of thermal management and innovative cooling strategies for data center

[7] <https://www.ashrae.org/news/hvacindustry/2019-update-to-standard-90-4>

[8] <https://www.fraicheurdeparis.fr/reseau-de-froid-urbain>

[9] Yogesh Fulpagare ,Atul Bhargav, 2015, Advances in data center thermal management

[10] Ralph L. Webb, 2005, Next Generation Devices for Electronic Cooling With Heat Rejection to Air

[11] M Saini, Ralph L. Webb, 2003, Heat rejection limits of air cooled plane fin heat sinks for computer cooling

[12] Severin Zimmermann, Ingmar Meijer, Manish K. Tiwari, Stephan Paredes, Bruno Michel, Dimos Poulikakos, Aquasar: A hot water cooled data center with direct energy reuse, 2012, Energy

[13] https://www.lexpress.fr/actualite/refroidir-nos-datacenters-un-vrai-casse-tete_2028689.html

[14] <https://datacenter-magazine.fr/le-credit-agricole-a-choisi-la-technologie-de-refroidissement-par-immersion-dasperitas/>

[15] Ali Habibi Khalaja, Saman K. Halgamuge, 2017, A Review on efficient thermal management of air- and liquid-cooled data centers: From chip to the cooling system

[16] https://www.sft.asso.fr/Local/sft/dir/user-775/documents/actes/Congres_2016/resumes_toulouse/132.pdf



49 avenue René Cassin
22100 DINAN



02 96 87 20 00



contact@pole-cristal.tm.fr



www.pole-cristal.fr

**VOUS AVEZ UNE QUESTION SUR CE LIVRE BLANC,
N'HÉSITEZ PAS À NOUS CONTACTER !**

