

AXE 6: Puissance & Refroidissement

Charly Saugey

Paul Haardt

Corentin Colmel

Majeure Image / Epita
GPU distribué 2026

Mission

- Expliquer les enjeux de puissance et de refroidissement des infrastructures IA moderne
- Comparer les technologies de refroidissement et analyser les astuces de design des datacenters

Partie A : Evolution de la puissance des systèmes IA

Evolution de la puissance des GPU

Questions clés

- Comment la consommation électrique des GPU a-t-elle évolué ?
- Qu'est-ce qui provoque l'augmentation de la puissance ?
- Existe-t-il une limite maximale à la puissance des GPU ?
- Comment la puissance évolue-t-elle avec les performances ?

Partie A : Evolution de la puissance des système IA

GPU	Year	TDP (W)	FP16 TFLOPs	W/TFLOP	Process Node
V100	2017	300	125	2.40	TSMC 12nm FFN
A100	2020	400	624	0.64	TSMC 7nm
H100 SXM	2022	700	1979	0.35	TSMC 4N
H200 SXM	2024	700	1979	0.35	TSMC 4N
B200 SXM	2024	1000	4500	0.22	TSMC 4NP
B300	2025	1100	5000	0.22	TSMC 4NP
R100 (Rubin)	2026	~2300	~4000	~0.58	TSMC N3

Partie A : Evolution de la puissance des système IA

Puissance au niveau système (nœud complet)

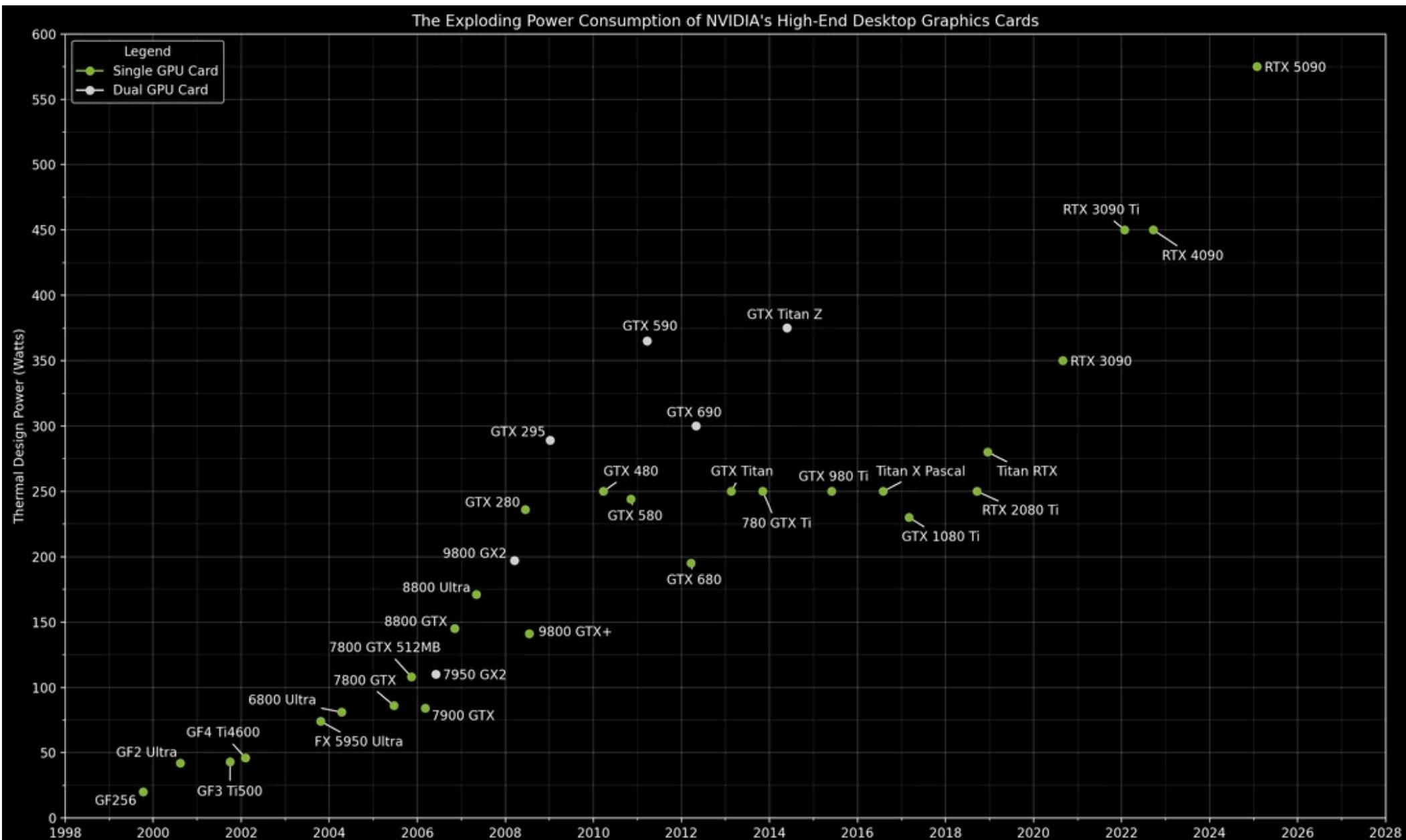
System	GPUs	GPU Power	Total System Power	Year
DGX-1	8 × V100	2.4 kW	~3.5 kW	2017
DGX A100	8 × A100	3.2 kW	~6.5 kW	2020
DGX H100	8 × H100 SXM	5.6 kW	~10.2 kW	2022
DGX B200	8 × B200 SXM	8.0 kW	~14.3 kW	2024
GB200 NVL72	72 × B200	72.0 kW	~120 kW	2024

Partie A : Evolution de la puissance des système IA

Évolution de la densité de puissance des racks

Era	Typical Rack Power kW/rack	kW/rack	Cooling Method
Traditional IT (2010)	5–10 kW/rack	~8	Air
GPU compute (2018)	15–25 kW/rack	~20	Air
AI training (2022)	30–50 kW/rack	~40	Air / DLC
AI training (2024)	60–120 kW/rack	~80	DLC required
AI training (2026)	150–300 kW/rack	~200	DLC / Immersion

Partie A : Evolution de la puissance des système IA



Partie B: Distribution de puissance aux GPU

Comment fonctionne l'alimentation électrique d'un GPU ?

Questions clés

- Du réseau électrique à la puce : quelles sont les étapes de conversion ?
- Où se situent les pertes d'efficacité ?
- Comment la puissance est-elle distribuée à l'intérieur d'un GPU ?



Partie B: Distribution de puissance aux GPU

Stage	Input	Output	Typical Efficiency	Loss
Utility transformer	HV AC (10–30 kV)	LV AC (400–480 V)	98–99%	1–2%
UPS	AC	AC (conditioned)	94–97%	3–6%
PDU	AC	AC	98–99%	1–2%
PSU (AC-DC)	AC (200–240 V)	DC (12 V / 48 V)	92–96%	4–8%
VRM (DC-DC)	DC (12 V / 48 V)	DC (~0.7–1.1 V)	85–95%	5–15%

Modules régulateurs de tension (VRM)

Questions

- Qu'est-ce que VRM et pourquoi est-ce un élément critique ?
- Qu'est-ce que les phases et comment fonctionnent-elles ?
- Pourquoi le rendement d'un VRM est important ?
- Quels sont les défis thermiques ?

Partie B: Distribution de puissance aux GPU

VRM Aspect	Description	Typical Value
Input voltage	Voltage supplied to VRM	12 V or 48 V
Output voltage	Core supply voltage (Vcore)	0.7–1.1 V
Phase count	Parallel VRM phases	10–20 phases
Efficiency	DC-DC conversion efficiency	85–95%
Power loss (1kW GPU)	Heat dissipated by VRM	50–150 W

Distribution de puissance 12V vs 48V

- En quoi le 48V est meilleur pour les systèmes de haute puissance?
- Qui promeut le 48V?
- Quels sont les défis?

Partie B: Distribution de puissance aux GPU

Aspect	12V Distribution	48V Distribution
Intensité pour 1kW	~83 A	~21 A
Epaisseur du câble	Thick (high current)	Thinner
Pertes par effet Joule	High ($\times 16$ vs 48V)	Much lower
Taille du connecteur	Large / multiple pins	Smaller / fewer pins
Adoption industrielle	Legacy servers	AI & hyperscale

Partie B: Distribution de puissance aux GPU

Considérations sur les connecteurs

Connector	Max Power	Pins	Issues
8-pin PCIe	150 W	8	Limited power, multiple cables needed
12VHPWR	600 W	16 (12+4 sense)	Overheating, insertion sensitivity
12V-2x6	600 W	16 (12+4 sense)	Improved safety, still high current

800V DC pour les usines IA

Questions

- Pourquoi NVIDIA promeut-t-il 800V DC?
- Quels sont les avantages en termes d'efficacité?
- Comment cela change-t-il la conception des datacenter?
- Quels sont les considérations de sécurité?

Partie C: Technologies de refroidissement

Refroidissement par air

Questions

- Comment fonctionne le refroidissement par air?
- Quels sont les principes de conception d'un heat sink?
- Quels sont les limites?
- Quand le refroidissement par air échoue-t-il?
- A quel TDP l'air devient inutile?
- Quels sont les limites acoustiques?
- Comment l'altitude affecte-t-elle le refroidissement par air?

Partie C: Technologies de refroidissement

Air Cooling Aspect	Typical Value	Limit
Max heat dissipation per GPU	400–600 W	~700–800 W
Max rack density	20–40 kW	~50 kW
Airflow per rack	$85\text{--}170 \text{ } m^3m^{-1}$	$\sim 225 \text{ } m^3m^{-1}$
Fan power overhead	5–10%	~15%

Direct Liquid Cooling (DLC)

Comment fonctionne le DLC ?

- Conception et fixation des cold plates
- Manifolds et systèmes de distribution
- Types de fluides et débits
- Rejet de chaleur (CDU, dry coolers, tours de refroidissement)

Partie C: Technologies de refroidissement

DLC Component	Function	Key Specifications
Cold plate	Transfert de chaleur du GPU vers le liquide	Surface alu/cuivre, low ΔT
Manifold	Distribution du fluide aux cold plates	Multi-port, équilibrage de débit
Quick disconnects	Connexions sans fuite pour maintenance	$\leq 0.5\%$ fuite, haute pression
CDU (Coolant Distribution Unit)	Pompe + échangeur + filtration	Débit 10–100 L/min, filtres
Facility water loop	Rejet de chaleur vers le bâtiment	Dry coolers ou cooling towers

Partie C: Technologies de refroidissement

Aspect	Air Cooling	Direct Liquid Cooling
Max GPU TDP supported	~700–800 W	~1500+ W
Max rack density	~40–50 kW	~100–200+ kW
PUE impact	Higher overhead	Lower overhead
Maintenance complexity	Low	Medium/High
Capital cost	Lower	Higher (install + CDU)
Operating cost	Higher fan energy	Lower pumping energy

Partie C: Technologies de refroidissement

Coolant	Thermal Properties	Cost	Safety	Use Case
Water/glycol	Very high heat capacity, high thermal conductivity	Low	Conductive, leak risk	Cold plates, DLC loops
Propylene glycol	Slightly lower than water, freeze protection	Low-Medium	Less toxic than ethylene glycol	Cold climates, DLC
Dielectric fluids	Lower than water, electrically insulating	High	Non-conductive, safer for electronics	Immersion cooling

Immersion Cooling

Sujets

- Monophasé vs diphasé
- Conception des bacs et gestion des fluides
- Méthodes de rejet thermique
- Considérations de maintenance

Partie C: Technologies de refroidissement

Aspect	Single-Phase Immersion	Two-Phase Immersion
Fluid type	Dielectric oil	Low-boiling dielectric fluid
Operating principle	Liquid absorbs heat, stays liquid	Liquid boils, vapor condenses
Max heat flux	High	Very high
Fluid cost	Medium	High
Complexity	Medium	High
Maturity	Commercially mature	Emerging / niche

Partie C: Technologies de refroidissement

Advantage	Challenge
Highest heat density	Specialized fluids
No fans required	Hardware compatibility
Reduced PUE	Higher CAPEX
Component longevity	Maintenance procedures

Rear-Door Heat Exchangers (RDHx)

Questions

- Qu'est-ce que RDHx ?
- Comment est-ce que cela aide le refroidissement à air ?
- Quelles puissances peut-il supporter ?
- Quand est-ce le bon choix ?

Partie C: Technologies de refroidissement

RDHx Type	Cooling Capacity	Best For
Passive RDHx	~20 000–35 000 W par rack	Racks densité moyenne, retrofit
Active RDHx	~30 000–80 000 W par rack	Racks haute densité (GPU)

Partie D: Rendement énergétique (PUE)

Comprendre le PUE

Définition

- PUE = Power Usage Effectiveness
- PUE = Total Facility Power / IT Equipment Power
- $1 \text{ PUE} = (\text{IT Load} + \text{Cooling} + \text{Power Distribution} + \text{Lighting} + \text{Other}) / \text{IT Load}$

Partie D: Rendement énergétique (PUE)

Questions

- Que mesure le PUE et que ne mesure-t-il pas ?
- Comment le choix du refroidissement affecte-t-il le PUE ?
- Quels sont les composants du surcoût énergétique ?

Partie D: Rendement énergétique (PUE)

PUE Component	Typical % of Overhead	Reduction Strategies
Cooling	40–60 %	DLC, free cooling, confinement
Power distribution losses	10–15 %	Haut rendement UPS/PSU
Lighting and other	5 %	LED, automatisation

Partie D: Rendement énergétique (PUE)

PUE benchmarks

Datacenter Type	Typical PUE	Best-in-Class PUE
Legacy enterprise	1.8–2.5	1.6
Modern enterprise	1.4–1.6	1.3
Hyperscale (air)	1.2–1.4	1.1
Hyperscale (DLC)	1.1–1.2	1.05
AI-optimized	1.1–1.3	1.05

Partie D: Rendement énergétique (PUE)

Cooling Method	Typical PUE	Why
Traditional air (CRAC)	~1.6	Composants énergivores
Hot/cold aisle containment	~1.4	Airflow optimisé
Free air cooling	~1.2	Peu de refroidissement actif
Direct liquid cooling	~1.1	Transfert thermique direct
Immersion cooling	~1.05	Quasi suppression HVAC

Partie D: Rendement énergétique (PUE)

Exemples de référence (best-in-class)

Company	Facility	PUE	How Achieved
Google	Hyperscale DC	1.10	Free cooling + AI control
Meta	Hyperscale DC	1.09	Free cooling
Microsoft	Azure DC	1.12	Free air + DLC
NVIDIA DGX Cloud	AI DC	~1.1	Liquid cooling

Partie D: Rendement énergétique (PUE)

Au-delà du PUE : autres métriques d'efficacité

Metric	Definition	Typical Values	Best-in-Class
PUE	Ratio total/IT	1.1–1.6	~1.05
WUE	L d'eau / kWh IT	0.2–1.8	<0.2
CUE	kgCO ₂ / kWh IT	Variable	Near zero

Partie E: Spécifications de l'infrastructure

Planification de la densité de puissance

Questions

- Comment planifier l'augmentation de densité ?
- Quelles mises à niveau d'infrastructure sont nécessaires ?
- Comment gérer des densités mixtes ?

Partie E: Spécifications de l'infrastructure

Density Tier	kW/Rack	Infrastructure Requirements
Low density	5 000–10 000 W	Air cooling, standard power
Medium density	10 000–30 000 W	Hot/cold aisle, high airflow
High density	30 000–80 000 W	RDHx or DLC
Ultra-high density	>80 000 W	DLC or immersion

Partie E: Spécifications de l'infrastructure

Architecture de distribution électrique

Component	Function	Sizing Consideration
Utility feed	Power from grid	Peak MW + redundancy
Main switchgear	Distribution & protection	Fault current rating
UPS systems	Backup & conditioning	MW load, minutes runtime
PDUs	Rack-level distribution	kW per rack
RPPs	Branch circuits	Density zones

Partie E: Spécifications de l'infrastructure

Planification de capacité de refroidissement

Cooling capacity Unit	Conversion	Context
1 ton of cooling	12 000 BTU/h	3.52 kW
1 kW IT load (air)	0.0003 ton	Includes overhead
1 kW IT load (DLC)	0.00028 ton	Direct rejection

Partie E: Spécifications de l'infrastructure

Besoins en eau pour le refroidissement liquide

Cooling Method	Water Usage	m ³ /h per MW
Evaporative (cooling tower)	High	1.5-2
Dry cooler	Low	0.1-0.3
DLC (closed loop)	Very low	~0.05

Partie E: Spécifications de l'infrastructure

Systèmes d'alimentation de secours

UPS Type	Efficiency	Runtime	Best For
Double conversion	94–97%	5–15 min	Critical loads
Line interactive	96–98%	5–10 min	Edge DC
Rotary UPS	96–98%	Seconds	Large DC
Battery + flywheel	95–98%	Seconds–minutes	High power

Partie E: Spécifications de l'infrastructure

Exigences des générateurs

Facility Size	Generator Capacity	Fuel Storage	Startup Time
10 MW	10-12 MW	Hours	<10 s
100 MW	120 MW	Hours	<10 s
500 MW	600 MW	Hours	<15 s
1 GW	1.2 GW	Hours	<15 s

Partie E: Spécifications de l'infrastructure

Connexion réseau pour sites à l'échelle du GigaWatt

Scale	Grid Requirements	Typical Lead Time
50 MW	Substation tie-in	1–2 years
200 MW	Dedicated substation	2–3 years
500 MW	Transmission upgrade	3–5 years
1 GW+	New transmission lines	5+ years

Stratégies d'approvisionnement énergétique

Strategy	Description	Pros	Cons
Grid connection	Utility power	Simple	Carbon intensity
On-site generation	Gas/diesel	Fast backup	Emissions
PPA	Long-term renewable contract	Green energy	Price lock
Behind-the-meter solar/wind	Local generation	Low carbon	Intermittent
Nuclear (SMR)	Small modular reactors	Stable low-carbon	Regulatory

Partie F: Sujets approfondis

Gestion de l'énergie au niveau puce

Mise à l'échelle dynamique tension/fréquence (DVFS)

- Comment fonctionne le DVFS ?
- Quelle est la relation puissance/fréquence ?

Partie F: Sujets approfondis

Power State	Voltage	Frequency	Power	Use Case
Max boost	Élevée	Élevée	Très élevée	Peak compute
Base clock	Nominale	Nominale	Élevée	Sustained
Idle	Basse	Basse	Faible	Low utilization
Sleep	Très basse	~0	Très faible	Inactive

Comportement du throttling thermique

Threshold	Temperature	Action
Target	~83 °C	Fréquence nominale
Throttle start	~85 °C	Réduction fréquence
Max operating	~90 °C	Limitation puissance
Shutdown	~95 °C	Arrêt matériel

Conception des dissipateurs et cold plates

Parameter	Impact	Tradeoff
Fin density	Surface d'échange	Restriction airflow
Base thickness	Diffusion chaleur	Masse
Heat pipe count	Transport chaleur	Coût
Material (Cu vs Al)	Conductivité	Poids / prix

Partie F: Sujets approfondis

Design Aspect	Consideration	Best Practice
Contact area	Résistance thermique	Couverture maximale
Channel design	Turbulence	Micro-canaux
Flow rate	ΔT liquide	Suffisant sans excès
Pressure drop	Charge pompe	Minimiser

Puissance échouée (« Stranded Power ») dans les datacenters

Questions

- Qu'est-ce que la puissance échouée ?
- Pourquoi apparaît-elle ?
- Quelle part de puissance est typiquement échouée ?
- Comment minimiser cette capacité perdue ?

Partie G: Entreprises & paysage industriel

Constructeurs de systèmes

Company	Products	Cooling Approach	Market Position
NVIDIA	DGX, MGX, HGX	Air + DLC ready	Leader AI platforms
Dell	PowerEdge XE	Air + DLC	Enterprise AI servers
HPE	Cray EX	DLC	HPC & AI leader

Partie G: Entreprises & paysage industriel

Company	Products	Cooling Approach	Market Position
Supermicro	GPU servers	Air + DLC	Broad OEM supplier
Lenovo	ThinkSystem	Air + DLC	Enterprise/HPC

Fournisseurs d'infrastructure de refroidissement

Company	Products	Technology Focus
Vertiv	Liebert, CDUs	Full stack cooling
Schneider Electric	APC, cooling	Power + thermal
Asetek	Cold plates, CDUs	DLC pioneer
CoolIT	DLC systems	Rack-level DLC
GRC	ICERaQ	Single-phase immersion
LiquidCool	Immersion tanks	Two-phase immersion
Submer	SmartPod	Immersion systems

Fournisseurs d'infrastructure électrique

Company	Products	Specialty
Schneider Electric	UPS, PDUs, switchgear	End-to-end power
Vertiv	UPS, PDUs	Critical power
Eaton	UPS, PDUs	Power distribution
ABB	Transformers, switchgear	Utility-scale
Caterpillar	Generators	Backup power

Partie H: Comparaison finale

Comparaison des technologies de refroidissement

Aspect	Air Cooling	Rear-Door HX	Direct Liquid	Single-Phase Immersion	Two-Phase Immersion
Max kW/rack	~20–50 kW	~20–80 kW	~50–200+ kW	~80–250+ kW	~100–300+ kW
PUE achievable	~1.2–1.6	~1.15–1.4	~1.05–1.2	~1.03	~1.02
Capital cost	Faible	Moyen	Élevé	Élevé	Très élevé
Operating cost	Moyen/ Élevé	Moyen	Faible/Moyen	Faible	Faible
Maintenance	Simple	Moyenne	Moyenne/Complexe	Complexe	Très complexe
Maturity	Très mature	Mature	Mature	Mature (commercial)	Plus niche
GPU compatibility	Universelle	Universelle	Cold-plate requis	Matériel compatible immersion	Matériel compatible immersion

Résumé de puissance des systèmes IA

System	GPUs	Total Power	Cooling Method	Rack Density
DGX A100	8× A100	~6.5 kW	Air	Variable (dépend du nombre de serveurs/rack)
DGX H100	8× H100	~10.2 kW	Air/DLC	Variable
DGX B200	8× B200	~14.3 kW	DLC	Variable
GB200 NVL72	72× B200	~120 kW	DLC	~120 kW/rack
AMD MI300X (8-way)	8× MI300X	~12.5 kW	Air	Variable
Google TPU v5p pod	8 960 TPU v5p chips	Non communiqué	DLC	Non communiqué

Comparaison d'efficacité des datacenters

Operator	Facility Type	PUE	WUE	Cooling Method
Google	Hyperscale	1.09 (2024)	Non communiqué (global public)	Mix (free cooling + liquid selon sites)
Meta	Hyperscale	1.08 (2023 avg)	WUE communiqué (voir rapport)	Mix (optimisations air + water systems)
Microsoft	Azure	Non communiqué (global public)	Amélioration WUE (ordre de grandeur communiqué)	Mix (air + innovations “zéro eau” sur nouveaux sites)
AWS	Cloud	1.15 (global)	Non communiqué (global public)	Mix (air + optimisations)
CoreWeave	AI-focused	1.15 (site Barcelone annoncé)	“Zéro eau” annoncé (site Barcelone)	Air (free cooling) + design optimisé
Lambda Labs	AI-focused	Non communiqué	Non communiqué	Non communiqué

Impact du choix de refroidissement sur le TCO

Cost Component	Air Cooling	DLC	Immersion
Capital (\$/kW IT)	Faible	Élevé	Élevé/Très élevé
Power cost (\$/kW-yr)	Élevé (ventilos/HVAC)	Plus faible	Plus faible
Maintenance (\$/kW-yr)	Faible	Moyen	Élevé
Floor space (\$/kW-yr)	Élevé (densité limitée)	Plus faible	Plus faible
5-year TCO (\$/kW)	Variable (souvent ↑ à haute densité)	Souvent ↓ à haute densité	Souvent ↓ si densité extrême

Sources

Part A

- <https://www.nvidia.com>
- <https://www.tsmc.com/english/dedicatedFoundry/technology>
- <https://www.anandtech.com/>
- <https://www.servethehome.com/>
- <https://www.nvidia.com/en-us/data-center/>
- <https://www.servethehome.com/nvidia-gb200-nvl72-power-cooling-and-rack-scale-design/>
- <https://www.anandtech.com/>

Sources

- <https://www.se.com/ww/en/work/solutions/for-business/data-centers-and-networks/>
- <https://www.vertiv.com/en-us/about/news-and-insights/articles/white-papers/>
- <https://uptimeinstitute.com/resources/research-and-reports>
- <https://www.nvidia.com/en-us/data-center/>
- <https://www.servethehome.com/nvidia-gb200-nvl72-power-cooling-and-rack-scale-design/>
- <https://www.ashrae.org/technical-resources/bookstore/thermal-guidelines-for-data-processing-environments>
- <https://engineering.fb.com/>

Part B

- <https://www.pcisig.com/specifications/pciexpress>
- <https://www.nvidia.com/en-us/geforce/news/12vhpwr-connector/>
- <https://www.pcisig.com/news/pcisig-introduces-12v-2x6-connector>
- <https://www.anandtech.com/show/17530/nvidia-12vhpwr-issues-and-analysis>
- <https://www.opencompute.org/documents/ocp-48v-dc-power-distribution>
- <https://www.nvidia.com/en-us/data-center/energy-efficiency/>
- <https://www.ti.com/power-management/48v.html>

Sources

- <https://www.analog.com/en/analog-dialogue/articles/why-48v.html>
- <https://www.anandtech.com/show/17626/nvidia-hopper-h100-architecture-deep-dive>
- <https://www.schneider-electric.com/en/work/solutions/for-business/data-centers-and-networks/>
- <https://www.vertiv.com/en-us/about/news-and-insights/articles/white-papers/data-center-power-chain/>
- <https://www.eaton.com/us/en-us/products/backup-power-ups-surge-it-power-distribution/data-center-ups-efficiency.html>
- <https://www.plugloadsolutions.com/80pluspowersupplies.aspx>

Sources

- <https://www.ti.com/power-management/vrm.html>
- <https://www.intel.com/content/www/us/en/products/docs/processors/core/vrm-design-guide.html>
- <https://www.ti.com/power-management/vrm.html>
- <https://www.intel.com/content/www/us/en/products/docs/processors/core/vrm-design-guide.html>
- <https://www.analog.com/en/analog-dialogue/articles/multiphase-buck-converters.html>
- <https://www.anandtech.com/show/17626/nvidia-hopper-h100-architecture-deep-dive>

Part C

- <https://www.ashrae.org/technical-resources/bookstore/thermal-guidelines-for-data-processing-environments>
- <https://www.vertiv.com/en-us/about/news-and-insights/articles/white-papers/high-density-data-center-cooling/>
- <https://www.schneider-electric.com/en/work/solutions/for-business/data-centers-and-networks/>
- <https://www.servethehome.com/nvidia-gb200-nvl72-power-cooling-and-rack-scale-design/>
- <https://www.ashrae.org/technical-resources/bookstore/thermal-guidelines-for-data-processing-environments>

Sources

- <https://www.vertiv.com/en-us/about/news-and-insights/articles/white-papers/data-center-liquid-cooling/>
- <https://www.nvidia.com/en-us/data-center/energy-efficiency/>
- <https://www.servethehome.com/nvidia-gb200-nvl72-power-cooling-and-rack-scale-design/>
- https://www.coolingbestpractices.com/knowledge_center/whitepapers/direct_liquid_cooling
- <https://www.schneider-electric.com/en/work/solutions/business/data-centers-and-networks/>
- <https://www.ashrae.org/technical-resources/bookstore/thermal-guidelines-for-data-processing-environments>

Sources

- <https://www.submer.com/immersion-cooling/>
- <https://www.grcooling.com/immersion-cooling/>
- https://www.3m.com/3M/en_US/data-center-us/solutions/liquid-cooling/
- <https://www.vertiv.com/en-us/about/news-and-insights/articles/white-papers/data-center-liquid-cooling/>
- <https://www.vertiv.com/en-us/about/news-and-insights/articles/white-papers/rear-door-heat-exchangers/>
- <https://www.schneider-electric.com/en/work/solutions/business/data-centers-and-networks/rear-door-heat-exchanger/>

Sources

- <https://www.asetek.com/liquid-cooling/technologies/rear-door-heat-exchanger/>
- <https://www.ashrae.org/technical-resources/bookstore/thermal-guidelines-for-data-processing-environments>

Part D

- <https://uptimeinstitute.com/resources/what-is-pue>
- <https://www.google.com/about/datacenters/efficiency/>
- <https://engineering.fb.com/2020/03/12/data-center-engineering/data-center-efficiency/>
- <https://learn.microsoft.com/en-us/azure/sustainability/>
- <https://www.ashrae.org/technical-resources/bookstore/thermal-guidelines-for-data-processing-environments>
- <https://www.iea.org/reports/data-centres-and-data-transmission-networks>

Part E

- <https://uptimeinstitute.com/resources/what-is-pue>
- <https://www.ashrae.org/technical-resources/bookstore/thermal-guidelines-for-data-processing-environments>
- <https://www.vertiv.com/en-us/about/news-and-insights/articles/white-papers/data-center-power-and-cooling/>
- <https://www.schneider-electric.com/en/work/solutions/for-business/data-centers-and-networks/>
- <https://www.iea.org/reports/data-centres-and-data-transmission-networks>
- <https://www.energy.gov/ne/articles/small-modular-reactors>

Part F

- <https://developer.nvidia.com/blog/nvidia-hopper-architecture-in-depth/>
- <https://www.anandtech.com/show/17626/nvidia-hopper-h100-architecture-deep-dive>
- <https://www.intel.com/content/www/us/en/products/docs/processors/core/vrm-design-guide.html>
- <https://www.ashrae.org/technical-resources/bookstore/thermal-guidelines-for-data-processing-environments>

Sources

- <https://www.vertiv.com/en-us/about/news-and-insights/articles/white-papers/data-center-power-and-cooling/>
- <https://uptimeinstitute.com/resources>

Part G

- <https://www.nvidia.com/en-us/data-center/>
- <https://www.dell.com/en-us/dt/servers/poweredge-xe.htm>
- <https://www.hpe.com/us/en/compute/hpc/cray-ex.html>
- <https://www.supermicro.com/en/products/gpu>
- <https://www.lenovo.com/us/en/servers-storage/servers/thinksystem/>
- <https://www.vertiv.com/>
- <https://www.se.com/ww/en/work/solutions/for-business/data-centers-and-networks/>

Sources

- <https://www.asetek.com/data-center-liquid-cooling/>
- <https://www.coolitsystems.com/>
- <https://www.grcooling.com/>
- <https://www.submer.com/>
- <https://www.eaton.com/>
- <https://new.abb.com/>
- <https://www.caterpillar.com/>

Part H

- <https://images.nvidia.com/aem-dam/Solutions/Data-Center/nvidia-dgx-a100-datasheet.pdf>
- <https://docs.nvidia.com/dgx/dgxh100-user-guide/introduction-to-dgxh100.html>
- <https://www.nvidia.com/en-us/data-center/dgx-b200/>
- <https://www.nvidia.com/en-us/data-center/gb200-nvl72/>
- <https://www.sunbirddcim.com/blog/your-data-center-ready-nvidia-gb200-nvl72>
- https://www.dell.com/support/manuals/en-us/poweredge-xe9680/xe9680_ism_pub/safety-instructions?guid=guid-

Sources

- <https://docs.cloud.google.com/tpu/docs/v5p>
- <https://cloud.google.com/blog/products/ai-machine-learning/introducing-cloud-tpu-v5p-and-ai-hypercomputer>
- <https://datacenters.google/efficiency>
- <https://sustainability.atmeta.com/wp-content/uploads/2024/08/Meta-2024-Sustainability-Report.pdf>
- <https://aws.amazon.com/sustainability/data-centers/>
- <https://www.microsoft.com/en-us/microsoft-cloud/blog/2024/12/09/sustainable-by-design-next-generation-datacenters-consume-zero-water-for-cooling/>
- <https://www.edged.es/news/coreweave-partners-with-edged>

Sources

- <https://www.vertiv.com/en-us/about/news-and-insights/articles/white-papers/rear-door-heat-exchangers/>
- <https://submer.com/blog/single-phase-vs-two-phase-immersion-cooling/>
- <https://www.ashrae.org/technical-resources/bookstore/thermal-guidelines-for-data-processing-environments>