

MPLS JAPAN 2024

本検討書は（国研）産業技術総合研究所の承認を得た書類ではありません。

「AI時代のデータセンター」セッション TDP1,000W 最新GPUサーバの冷却方法



2024年10月24日

元（国研）産業技術総合研究所 情報・人間工学領域

デジタルアーキテクチャ研究センター

超分散アーキテクチャ研究チーム テクニカルスタッフ

株式会社LXスタイル

代表取締役 杉田 正

sugi@lxs.jp

目次

1. はじめに	• • • 3
2. 自己紹介	• • • 4
3. CPUクーラー	• • • 5
4. キャンパスタイプDC	• • 16
5. データセンター省エネ技術	• • 18
6. 外気導入とMetaの間接外気導入	• • 28
7. 産総研GreenITプロジェクト（2007-2012）	• • 38
8. 空冷と水冷のハイブリッド排熱	• • 46

1.はじめに

このセッションでは、AI時代に使われる大型GPUで必要な排熱技術をご紹介します。
データセンターでの大規模言語モデルLLMを構築するとき
必須な省エネデータセンター構築技術をご紹介します。

LLMとは「Large Language Model（大規模言語モデル）」の略で、大量のテキストデータを学習して自然言語処理や生成のタスクを実行できるAIモデルを指します。

主に次のような特徴があります：

- 1.大量のデータを学習：**数十億～数千億単語を含むテキストデータを学習しており、さまざまな分野の情報や知識を蓄えています。
- 2.多様なタスクに対応：**質問応答、要約生成、翻訳、コード生成、会話など、幅広い言語タスクに対応可能です。
- 3.高い言語理解力：**コンテキストを把握し、適切な返答や推論を行えるため、人間との自然なやりとりができます。

2.自己紹介 杉田 正

- ・元（国研）産総研 デジアキ テクニカルスタッフ
- ・JDCC東大次世代データセンター勉強会世話役メンバ
- ・新規事業開発経歴35年

ストレージRAIDNAS開発、SCSIネットワーク、Webサーバー開発、低PUE_DC開発、日本初DC_ISMS取得、HA_DC開発、レンサーバ専用サーバ開発、コンビニ向けSaaS、無料Windowsホスティング、VPSサービス、外気導入DC、井戸水利用DC、雪冷熱利用DC、世界一空調省エネクラウドECサービスDC、世界一省エネAIスパコンDCなど、どんなサイズDCでもPUE=1.1を実現、世界初を含み新規事業開発多数

世界一省エネなデータセンターを空冷・液冷ともに日本で。もう7年。

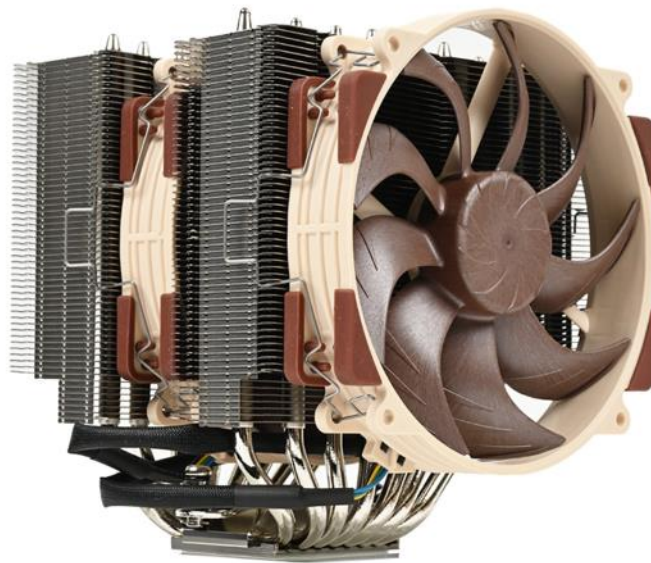
※産総研は情報を売る組織。NDAから外れる公知な事実を本日ご案内。

3.CPUクーラー

一般PC用クーラーはデカイ、うるさい

DeepCool

noctuaノクチュア
空冷最高とされる

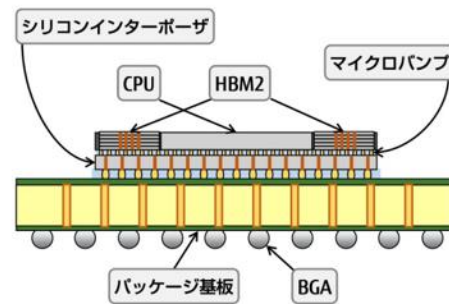


ヒートパイプが16本

TDP100W~250W? とされる。
前後に吸排気空間も必要

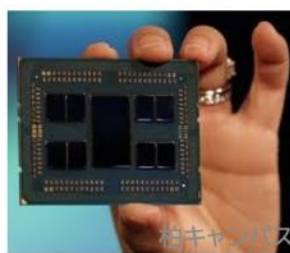
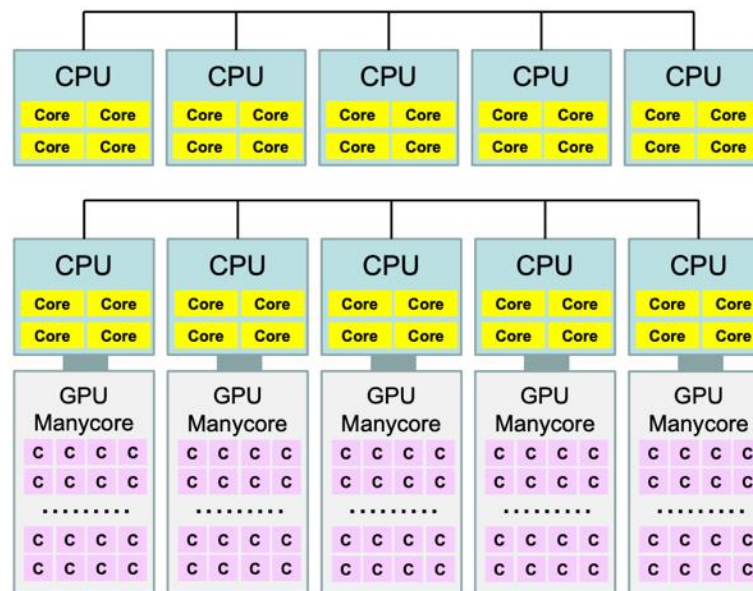
スパコンの速さの源泉： プロセッサ・メモリ・ネットワーク

- CPU (Central Processing Unit, 中央演算装置)
 - 並列計算が基本: PC, スマホも同様
 - 最新スパコンでは 40~100コア程度、それを1~2個
 - マルチコア、メニーコア
- GPU (Graphic Processing Unit)
 - 描画用途⇒高いデータ処理能力を計算に利用
 - 数百のコアを有する「メニーコア」
- メモリ: 三次元積層、高速なデータアクセス
- **マルチコア・メニーコアCPU(+GPU)を
超高速ネットワークで接続した並列計算機
(Parallel Computer) = 「スパコン」**
 - 家庭用LANの数百倍, 100Gbps~800Gbps



「whitepaper
富士通
PRIMEHPC
FX1000」
<https://www.fujitsu.com/downloads/JP/jsuper/primehpc-fx1000-hard-ja.pdf>

図 12 2.5次元パッケージの構成 (断面模式図)



柏キャンパス一般公開

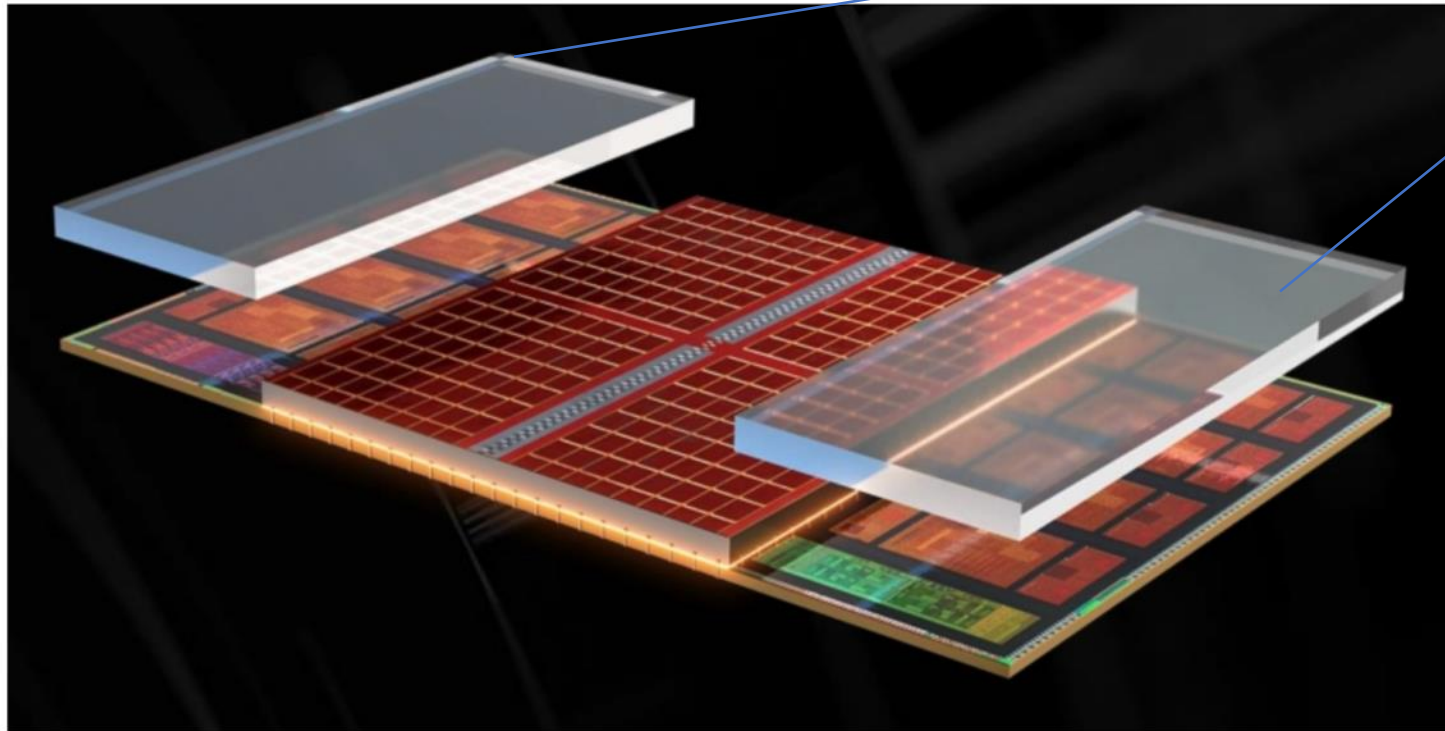
引用先： AIとスパコンが生み出す次世代の 科学技術シミュレーション
東京大学 情報基盤センター スーパーコンピューティング研究部門 埴 敏博

CPU/GPUに大容量メモリを搭載する時代。 メモリ耐熱は50℃前後

AMD's V-Cache technology does two things, it increases L3 cache capacity, allowing more data to be stored on-chip, and it significantly increases AMD's L3 cache bandwidth. Both of these changes will impact Milan X's performance. By keeping more data on-chip, Milan X processors will not need to access DRAM as frequently, freeing up that bandwidth for other tasks while granting users faster access to on-chip data.

You can join the discussion on **AMD's Milan-X CPU specifications** on the **OC3D Forums**.

製造工程で
L3キャッシュは別物



引用先 : https://www.overclock3d.net/news/cpu_mainboard/amd_s_milan-x_processors_specs_have_leaked_revealing_skus_with_768mb_of_l3_cache/1?fbclid=IwAR3TJi7JP2mvKcTZpqdOddpwK7JBQ8W58NjElhoC3IQsYHvKXjb_8wGfF3I

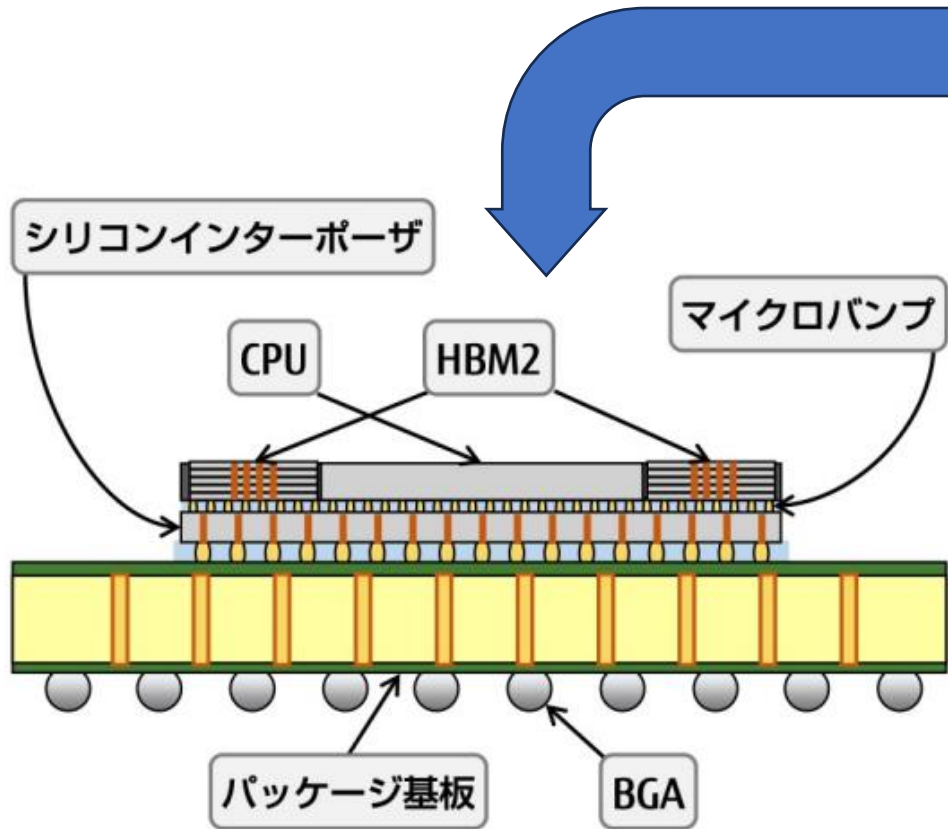
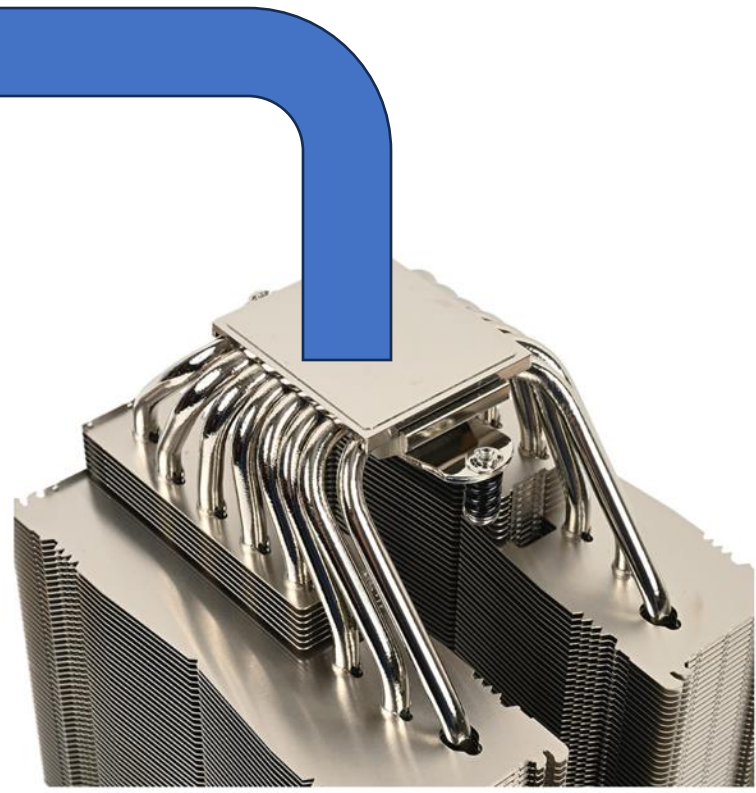
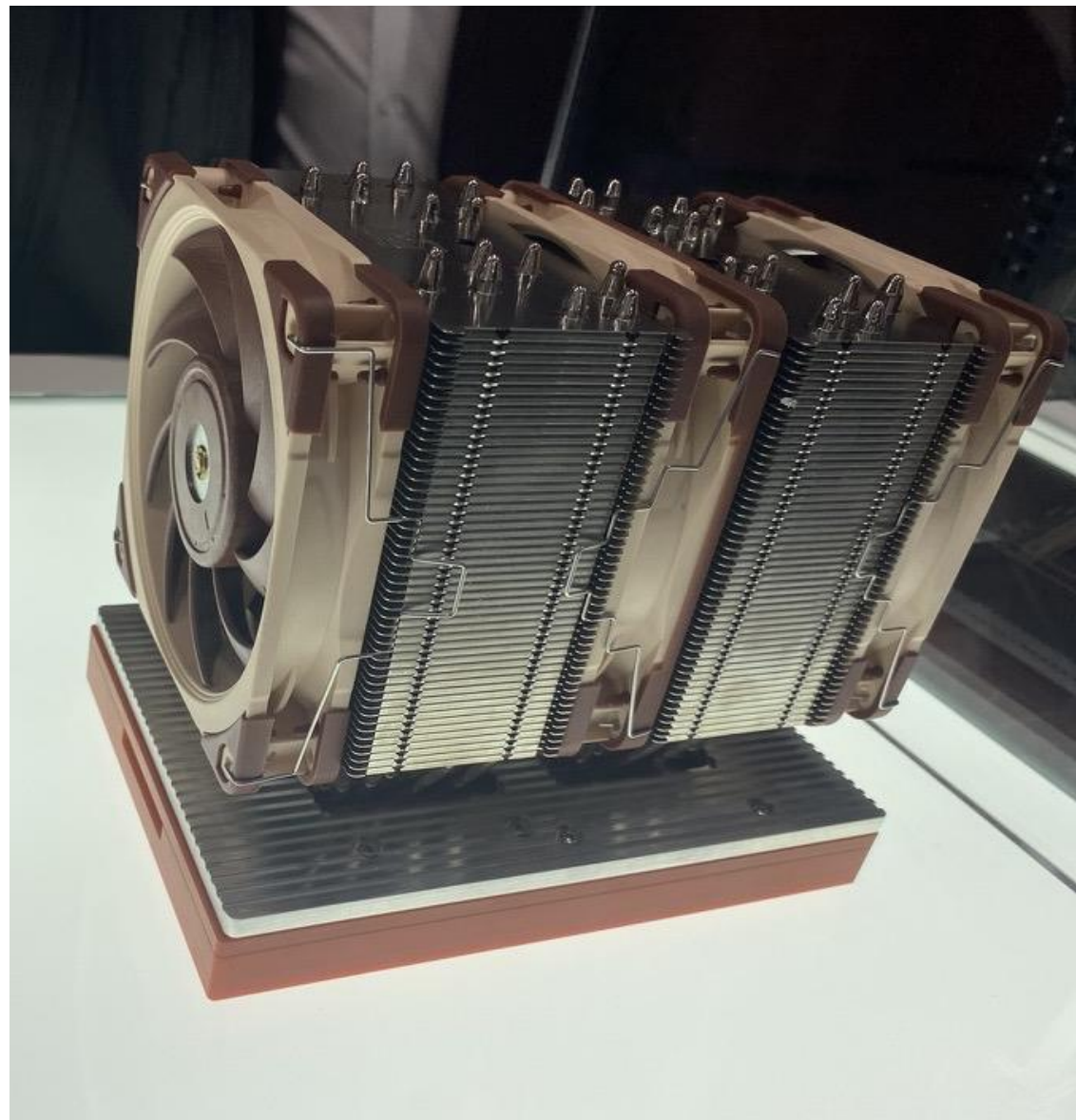


図 12 2.5次元パッケージの構成 (断面模式図)

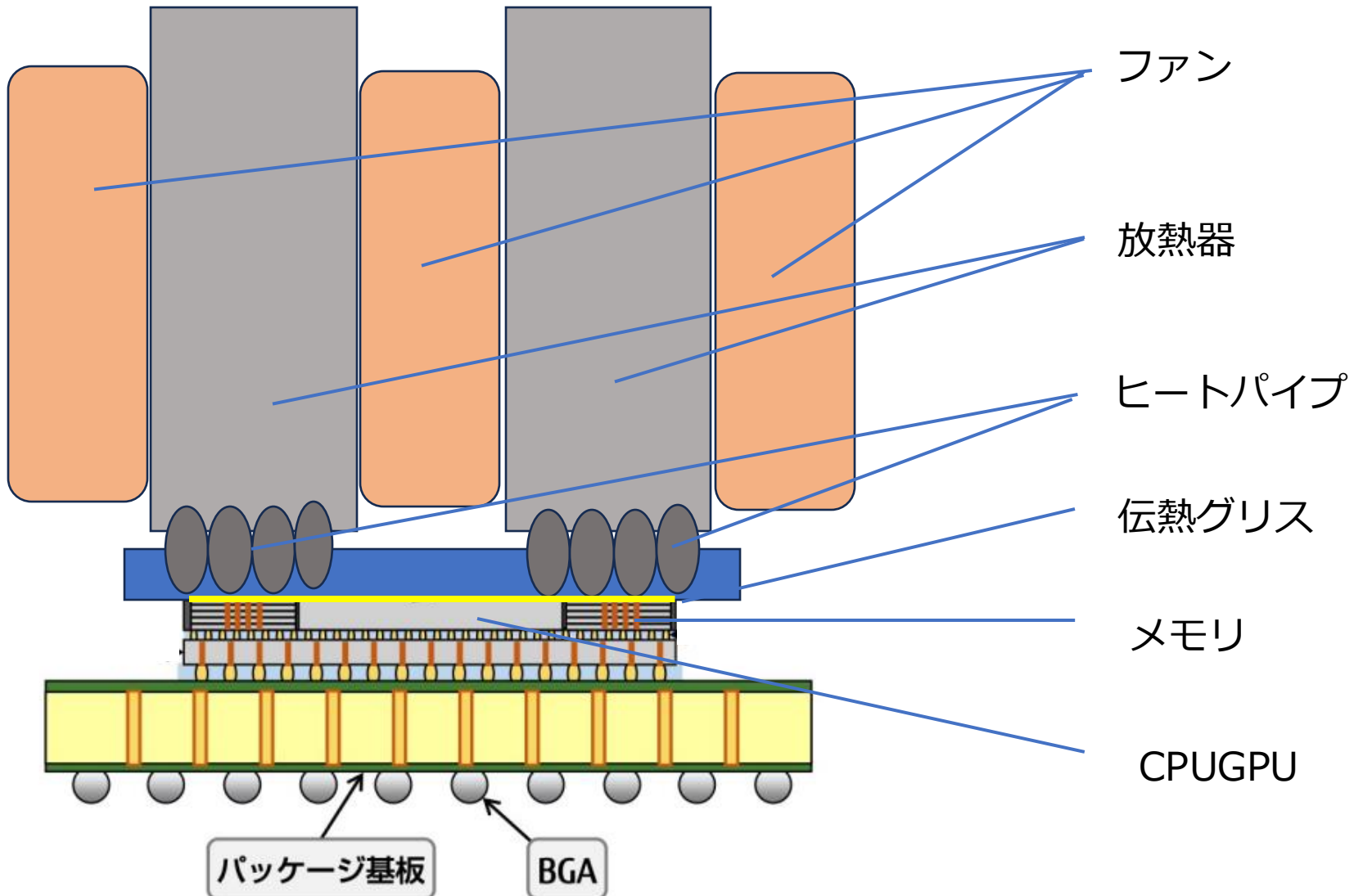


ヒートパイプが16本

引用先：AIとスパコンが生み出す次世代の 科学技術シミュレーション
 東京大学 情報基盤センター スーパーコンピューティング研究部門 埴 敏博



ComputexTaipei2024 筆者撮影



- 熱伝導率:11.8W/m・k
- 粘度:140~190pas

図 12 2.5次元パッケージの構成 (断面模式図)

ColdPlate式は、 ヒートパイプが無い ため熱伝導が悪い

この厚みが問題になる

流れ出さないグラフェンシートなどがある。

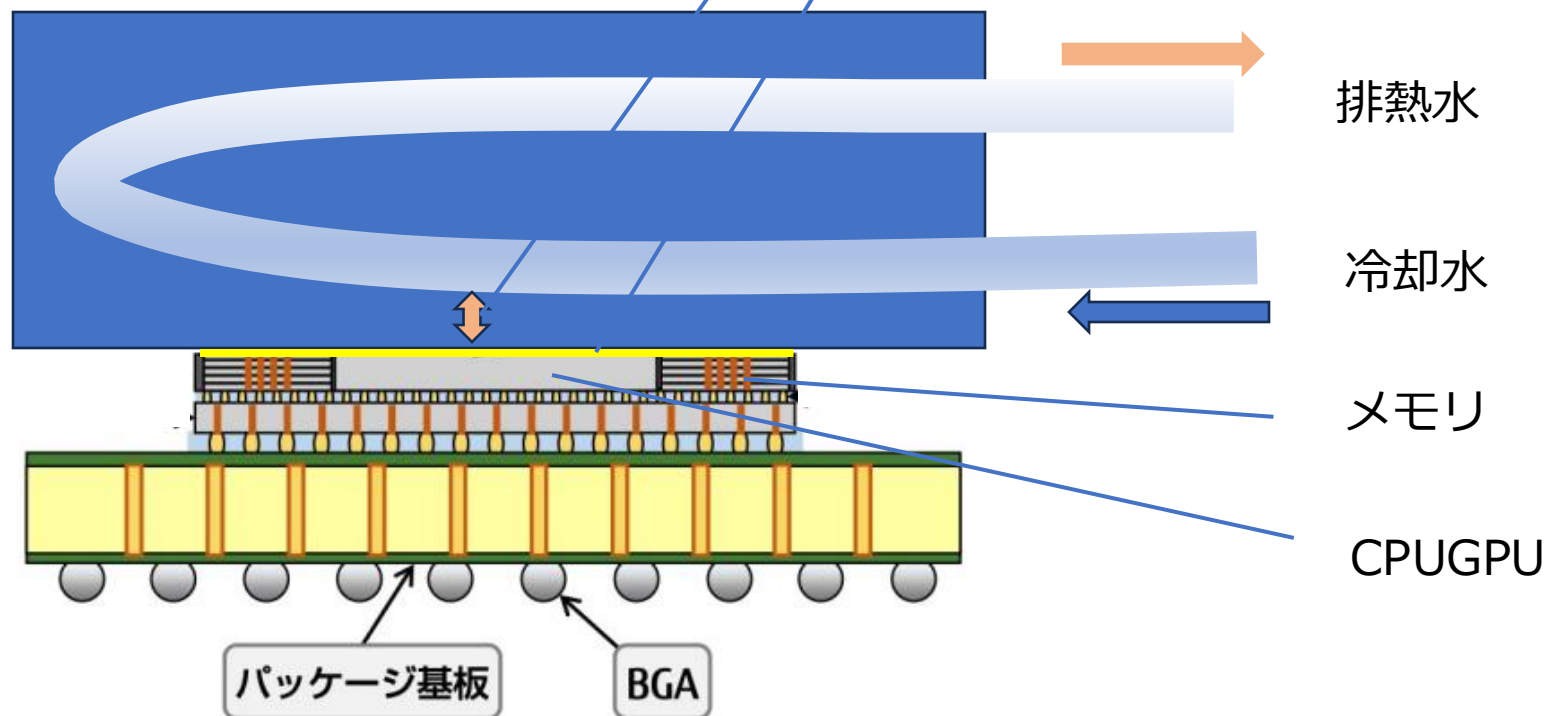
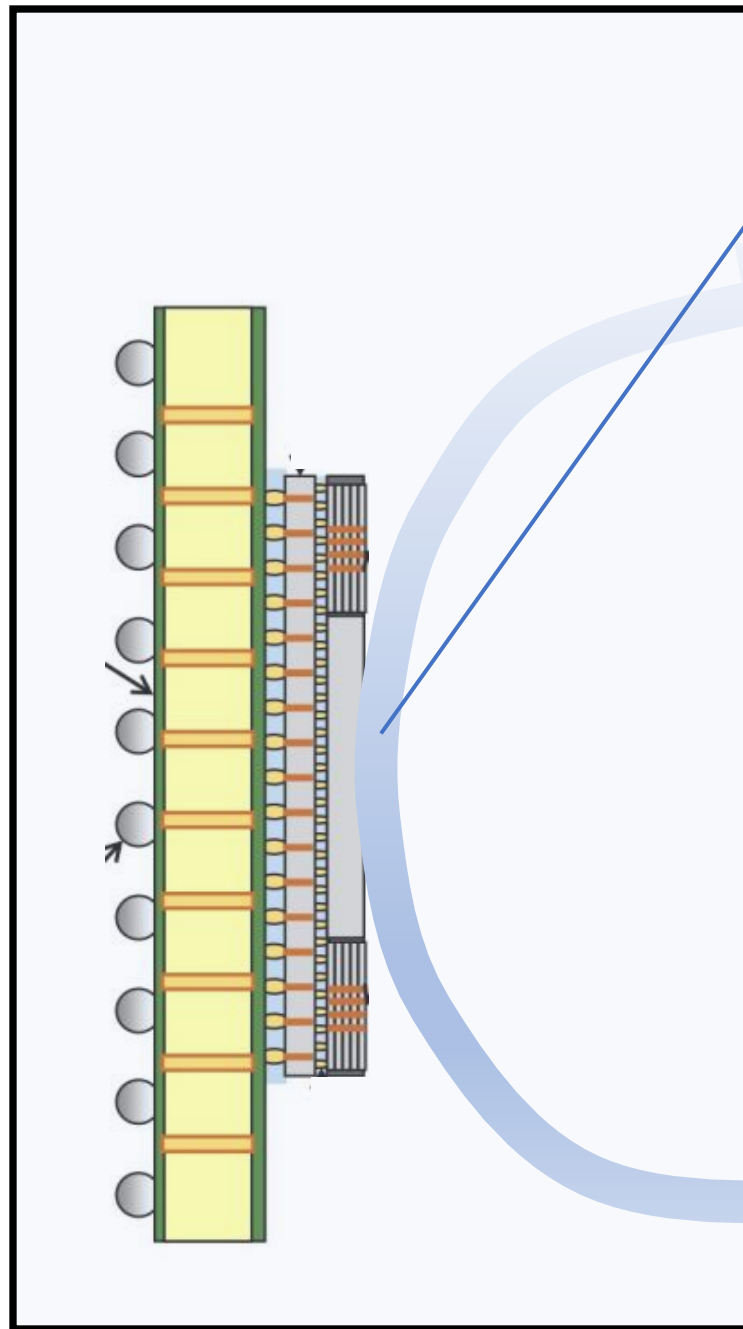


図 12 2.5次元パッケージの構成（断面模式図）

液浸1相式
CPU表面を
流れる流量が
不足する。



発生する泡も伝熱効率を下げる



排熱水

気化温度の低い液体を
使えば、CPU/GPU表面で
気化して大きく排熱出来る
→2相式

※フロンが気化した
ガスは人体に有害

冷却水





eSports用
液浸冷却システム
by 篠原電機

PUE=1.02

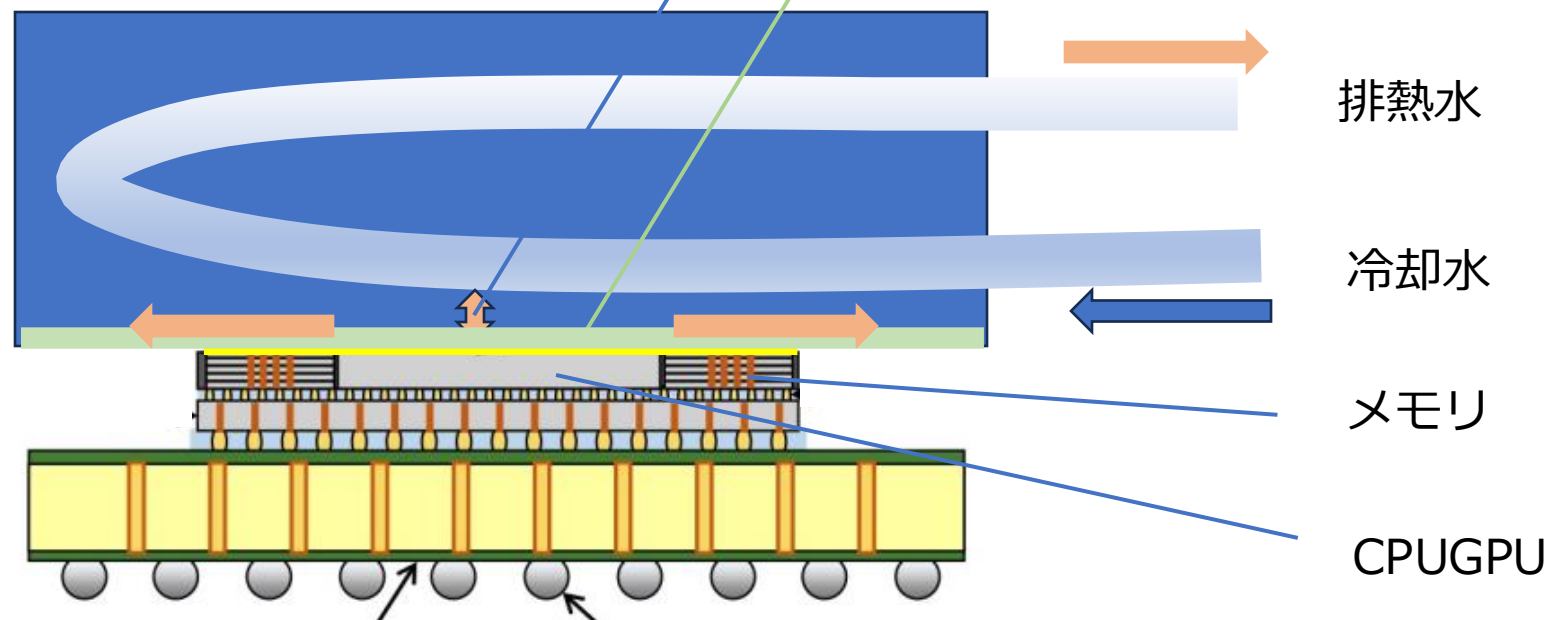
新型液体を開発中

液浸冷却システムでは、
相変化や泡の使い方が
開発ターゲット
中国ベンチャーが
26億円調達とか、

普通の素材は、
熱を一方向にしか
送れない。

この厚みが問題になる

横方向に高速大容量で熱搬送出来る素材
例：炭素セラミック、人工ダイヤモンドなど



電子負荷 と 熱伝導素材

4.AI時代のデータセンター キャンパス型DC-AI仕様と呼んでいます。

Internet
DC



@2KVA



Campus
type
DC



@10KVA

Hyper
scaler
DC



@15KVA



@10KVA









AI Specifications
Campus type DC



@70KVA

@200KVA

違いは、1ラック当たりの消費電力が大きい

Internet DC	Campus type DC	Hyperscaler DC	AI Specifications Campus type DC
 <p>@2KVA</p> 	 <p>@10KVA</p> 	 <p>@15KVA</p>  <p>@10KVA</p>	 <p>@70KVA</p>  <p>@200KVA</p>

Server Type

AALC: Air-Assisted Liquid Cooling, also called Direct Liquid Cooling

1U, 2U, blade	Rack Scale	OCP, dedicated machine	GPU supercomputer AALC*.
1Rack			
0.5KVA to 6KVA	6KVA to 12KVA	6KVA to 15KVA	6KVA to 15KVA
0.2KW to 4KW	6KW to 8KW	8KW to 12KW	8KW to 12KW
			+ liquid waste heat 40KVA to 250KVA

5. データセンター省エネ技術

冷却せず、動力を使わず排熱すれば省エネになります。

pPUE=1.1以下は、外部エネルギー（水の気化熱や雪冷熱）を利用してpPUE=1.015まで実現しています。

データセンター建造時にコストを掛けるとpPUEは下げられますが、ほどよいpPUEを設計時に決めて、設計出来る時代になりました。

寒暖分離 空調の基本キャッピング その後は冷水温度でPUEが決まる

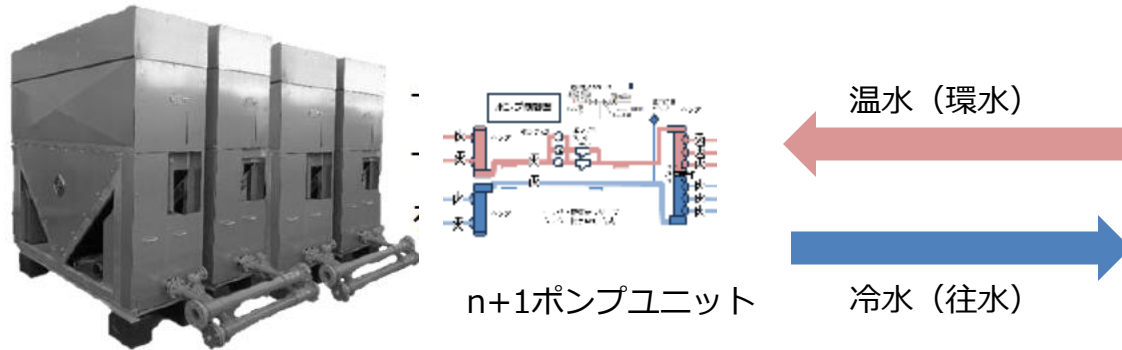
PUE=1.15以下 建造コストと運用コストで
PUEは決まる（設計出来る）

- ・ サーバ排気を高い温度で「水」に熱交換する。

高性能なAHU（Inrow）を使うほど戻り温度が上がる
チルドタワーなどフリークーリング機の排熱効率が上がる



DC in DC by 中空パネル
必要：煙感知器

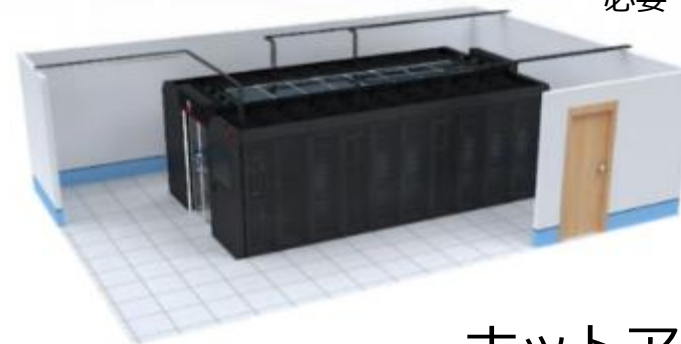


ドライクーラー
チルドタワーAシリーズ

必要な省電力に対して
スケールアップが容易な
付帯設備 80KW～1.8MW

17℃ PUE=1.65？
20℃ PUE=1.2以下
23℃ PUE=1.15以下
26℃ PUE=1.1以下

※冷水温度を上げるとpPUE値は
室内AHU性能に影響される。



ホットアイルキャッピング
ラックモジュール20ラック
@80KW～1.8MW
コールドアイル **27℃**

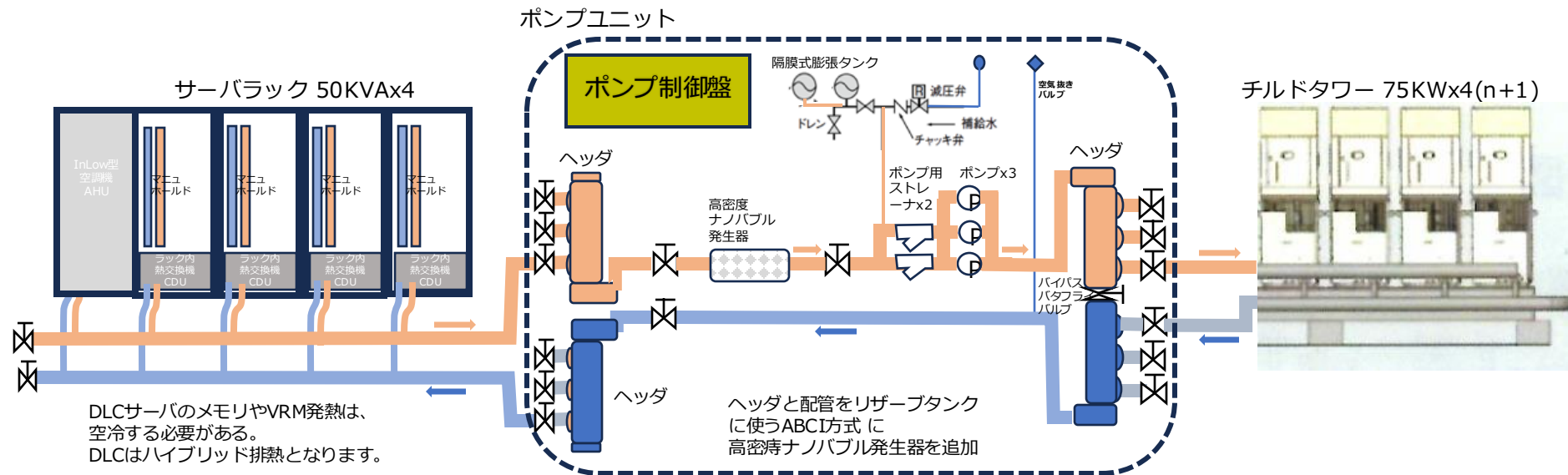
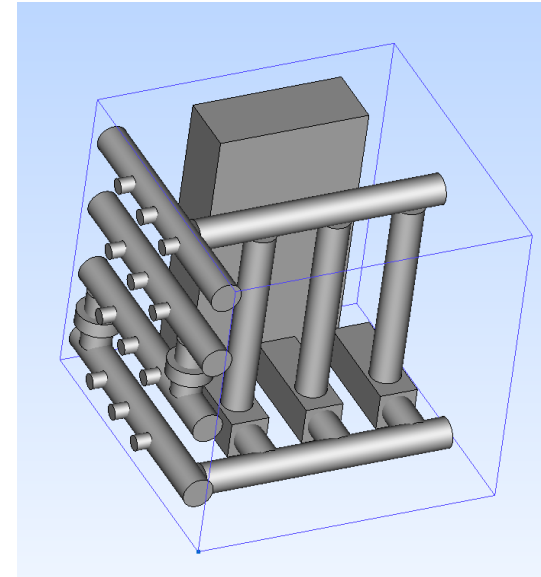
フリークーリング省エネ設計例

省エネ排熱ループ

瞬停対策

ポンプが止まるとサーバも止まる

重要なポンプユニット



フリークーリング30年 データセンター向け新型機 元日立金属だが、海外販売を行っていなかった。

桑名金属工業 チルドタワーAシリーズ (ドライクーラー)

NEW 低GWP冷媒R32を採用
NEW モジュールタイプを採用

供給温度範囲: 10~30°C

チルドタワー

チラー 冷却塔



■冷却塔(ドライクーラー)と空冷チラーを一つの機器に集約

自然エネルギーを利用したシステムを、チラーや冷却塔を組み合わせるよりも単純化でき、設計・施工の省力化、省スペース化を図れます。

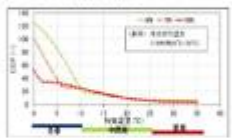
■圧縮機定格出力7.5kW未満のため、有資格者による定期点検が不要※

番号	点検	点検頻度	点検実施者
全ての第1種特定製造	3ヶ月に1回以上	お客様ご自身で点検可能	
50kW以上の冷凍機器	1年に1回以上	有資格者による点検が必要	
7.5kW以上の冷凍冷蔵機器	3年に1回以上	有資格者による点検が必要	

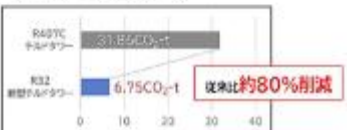
※「冷凍機器の点検」に基づく機器管理に係る「計画」の基準で定められる機器の点検において、定期点検が必要とされる対象機器ではありません。(3ヶ月に1回以上の稼働点検は必要となります。)

特長

▶ インバータ圧縮機の採用
インバータ圧縮機を採用し、低負荷率においても高いCOPを実現し、省ランニングコストを実現。

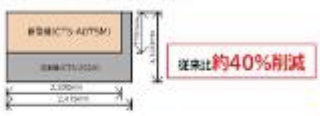


▶ R32冷媒の採用で環境負荷低減
GWP(地球温暖化係数)がR407Cの約1/3。冷媒充填量も従行の約1/2。



冷媒による地球温暖化への影響の比較(GWP換算 = GWP×冷媒量)

▶ 省スペースでの設置が可能
従来機種と比較してフットプリントを縮小。高さ寸法も縮小し、搬送性を向上。



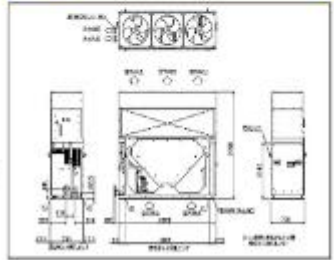
▶ モジュール化によりフレキシブルな設置、運転が可能
最大25台連結によるグループ制御可能。*



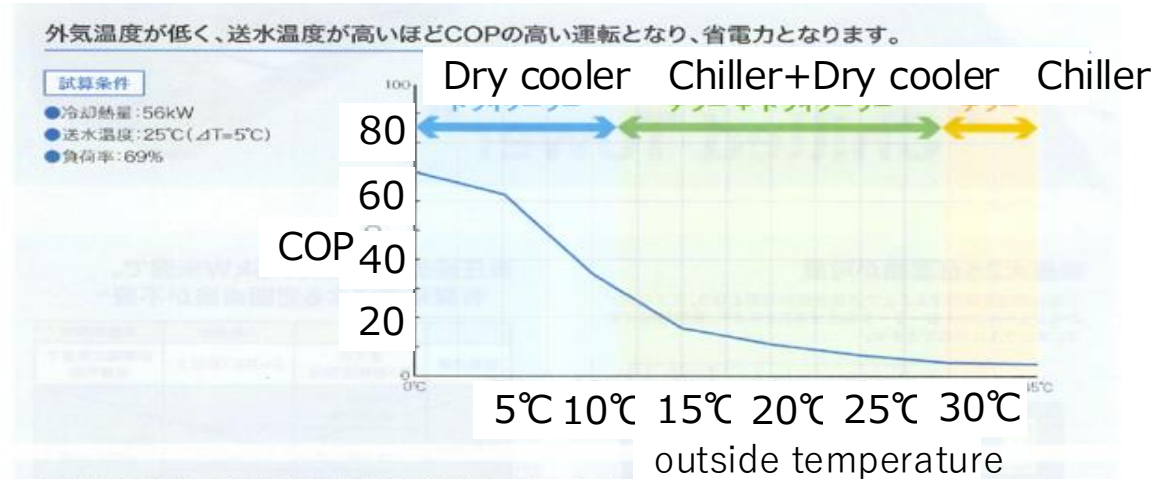
仕様表

仕様	項目	値
性能	入口温度25°C 冷卻能力	kW 75.3
	出口温度20°C 冷卻水量	m³/h 12.96
	入口温度35°C 冷卻能力	kW 83.3
	出口温度25°C 冷卻水量	m³/h 7.16
	冷水循環量(最小~最大)	m³/h 6 ~ 18
圧縮機	圧縮機出力	MPa/kW 0.053 (1.2)
	高圧ガス保安法適用区分 ¹⁾	種別不要
圧縮機	冷媒の種類 ²⁾	R-32
	圧縮機容量 ³⁾	kg 10
外形寸法	幅	mm 770
	奥行寸	mm 2,200
	高さ	mm 2,200
圧縮機	型式	スクロ-6.25中容量型電動圧縮機
	圧縮機出力	kW/kVA 5.5/4
運転圧力	最高運転圧力	MPa 6.7
	運転圧力	kW 20.9
電圧	運転電圧	A 64
	電圧	AC3φ 200V 50/60Hz
配管寸法	冷水出入口	JIS 10Kフランジ 50A
	制御電線(運転電線)	kg 980 (1,500)
質量	質量	kg 66
	質量	kg -10 ~ +43
冷却能力	冷却能力	PS 30wt%
	冷却能力	L 60
冷却	冷却能力	L 20
	冷却能力	L 20

外形図



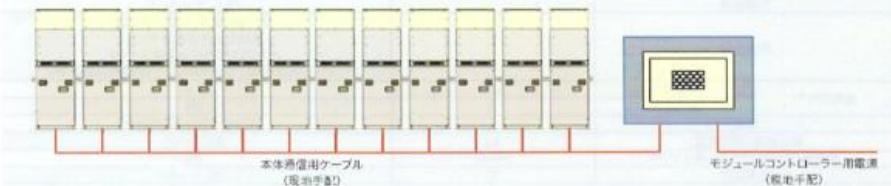
フリークーリング30年 @75KWを最大25台連結なスケーラビリティ 低コストでn+1を実現、軽量化屋上設置にも対応



*弊社での実験値から試算したものであり、実際の値は使用条件によって異なります。気象データは気象庁ホームページ2022年の観測値を用いています。

モジュール連結により逐次増設にも対応可能

オプションのモジュールコントローラーを使用すれば最大25台まで連結して運転可能です。



■ 設置イメージ(6台連結する場合の例)



■ モジュールコントローラー

- 台数制御
- 故障時の自動スキップ機能
- 機器の運転時間平準化
- 運転/停止、温度設定
- 運転状態・内部計測温度出力
- 遠隔運転操作入力
- 運転信号・異常信号出力

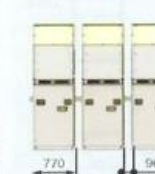
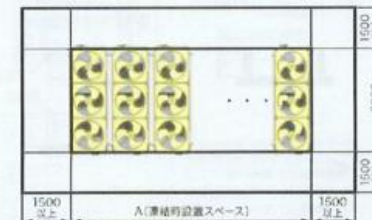


■ 通信プロトコル

- ▶ Modbus/RTU, Modbus/TCPでの外部とのデータのやり取り、遠隔制御が可能です。
- ▶ BACnetアダプタ(オプション)による通信も可能です。

*ModbusはSchneider Electric USA Incの登録商標です。
*BACnetはASHRAEの登録商標です。

■ 設置スペース



モジュール接続台数	A(最小)mm
1台	770
2台	1,630
3台	2,490
4台	3,350
...	...
8台	6,790
...	...
25台	21,410

どうして省エネデータセンターを作れるか？
サーバ設計から同様にデータセンターを作る。

データセンターはサーバを置く場所
冷却チラーが付いているサーバは世界に1台も無い
※量子コンピュータには1チップに巨大な冷凍機が付いている。

コンデンサ耐熱温度が20年で30℃以上あがっている。
部屋の温度を上げられる。カーナビ・スマホは何℃で動く？
つまり「冷却は不要」になってきた。

→ 「動力を使わず排熱するのが大事」

どうして省エネデータセンターを作れるか？
サーバ設計から同様にデータセンターを作る。
高密度実装して取れる排熱温度を上げる。
無料Windowsホスティング用に開発していたアルミ筐体サーバ



マザーボードは最終的に別物

なぜアルミ？

1ラック積載限界750kgで
高密度実装するため。

VCCI対策が難しい。

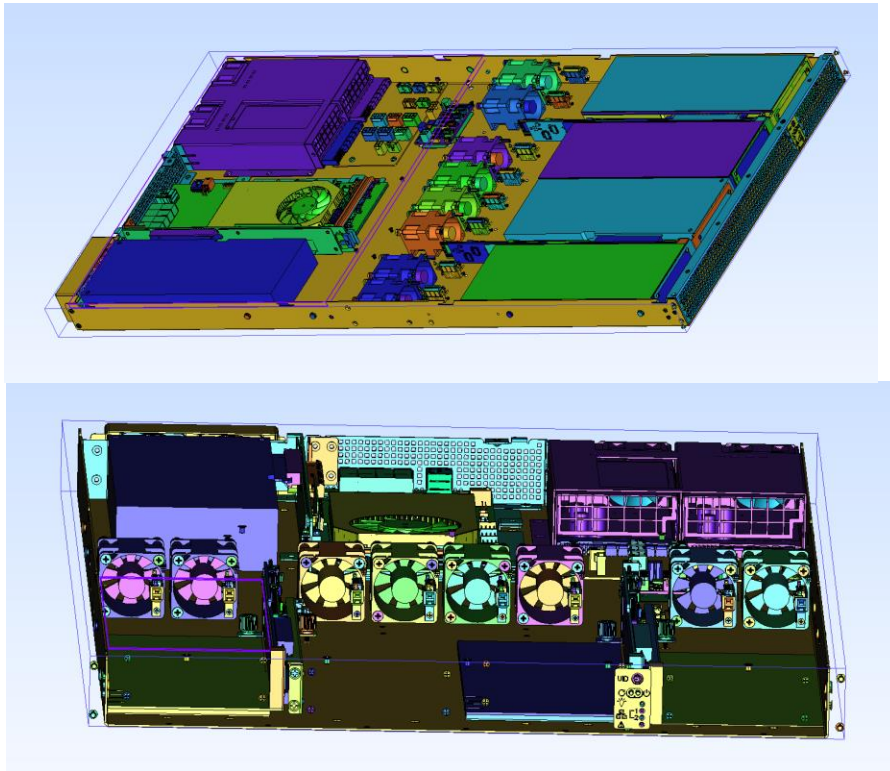
データセンターはサーバを置く場所。
ラックのアースを工夫してクリア。

2005年

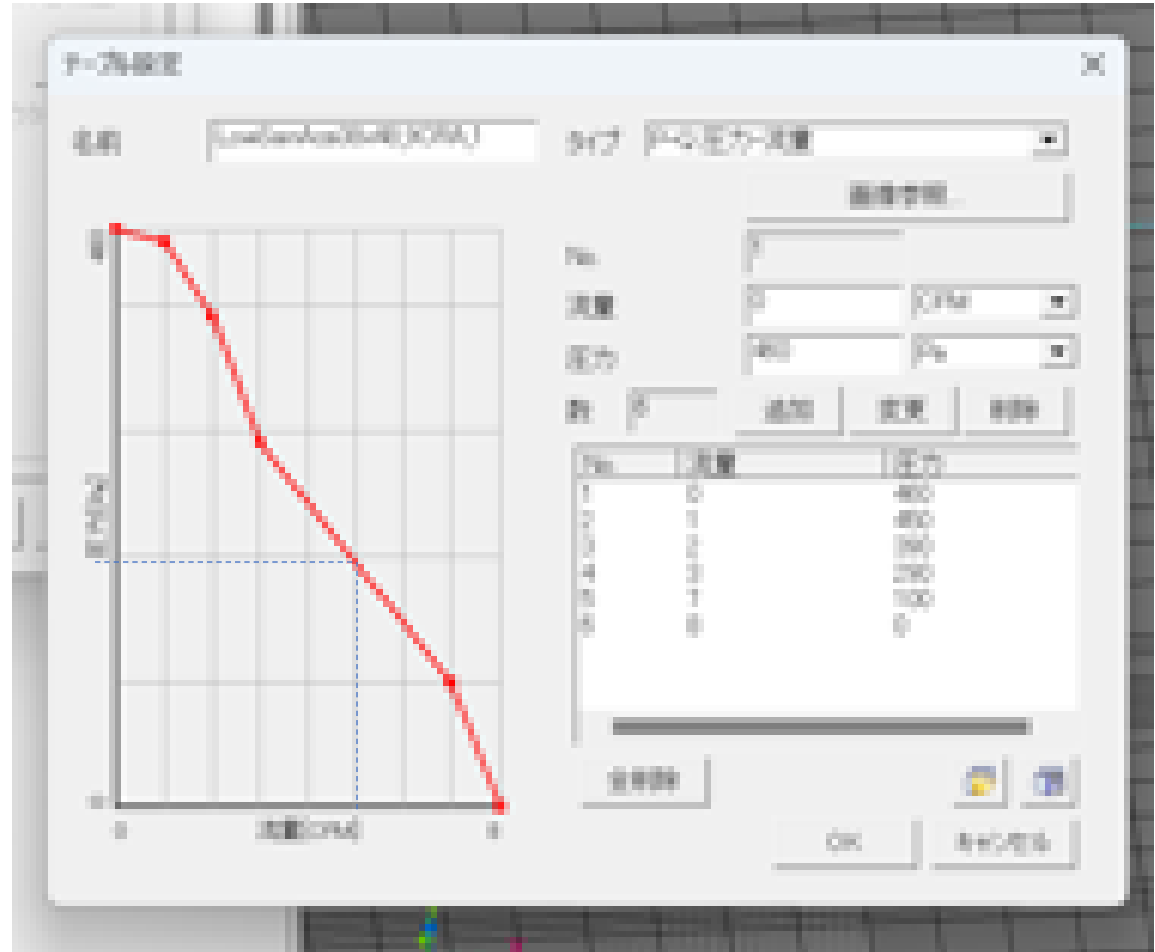
その後

NECマイクロサーバを使った
低価格VPSサービスの構築

1Uサーバ 内蔵ファン 大容量2重反転ファン8台でも100CFM
静圧力は1万台でも 200Pa以下
サーバ内蔵ファンを無視したデータセンターを設計するから「暑くなる」



資料提供 株式会社コアマイクロシステムズ



サーバとデータセンター サイズは違うが・・・

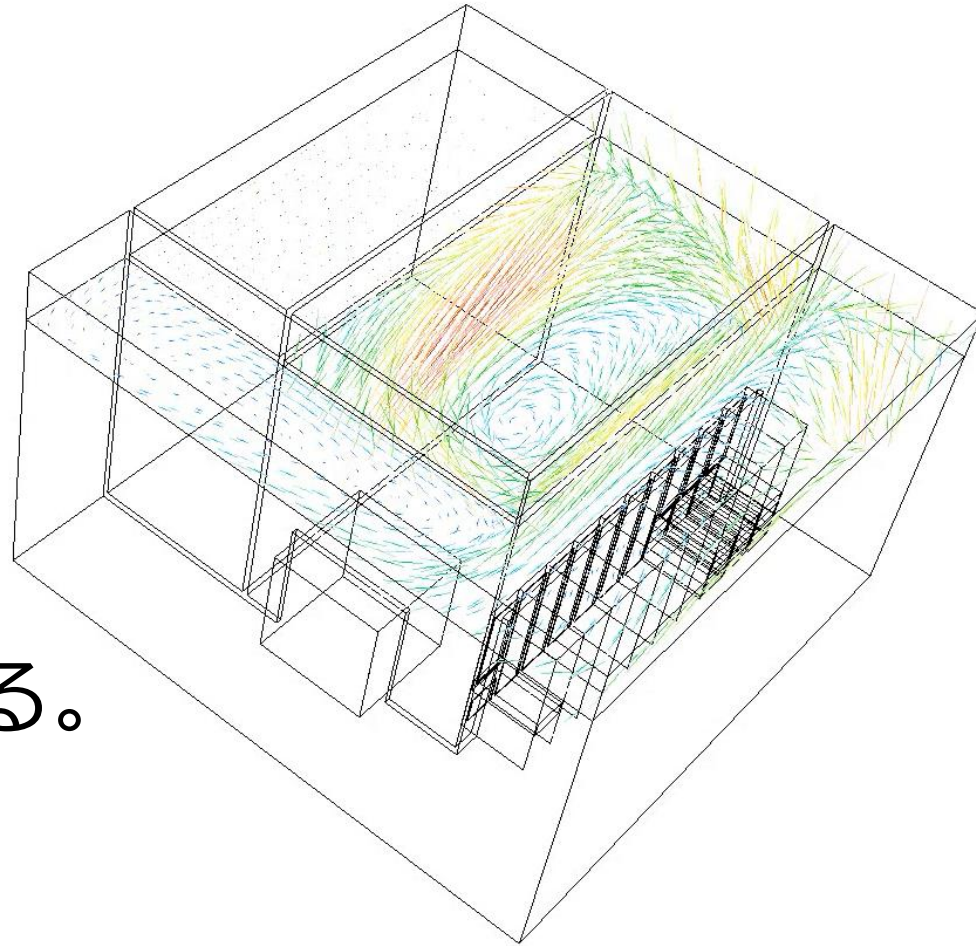
1ラック10KVA
CFD (流体解析)

HyperscalerDCも
空冷で10KVA

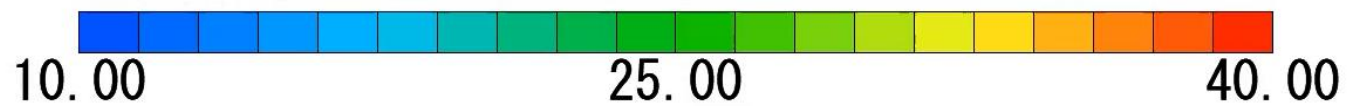
超えると液体で排熱となる。

温度傾斜5°C~7°C

1 KVA150CFM

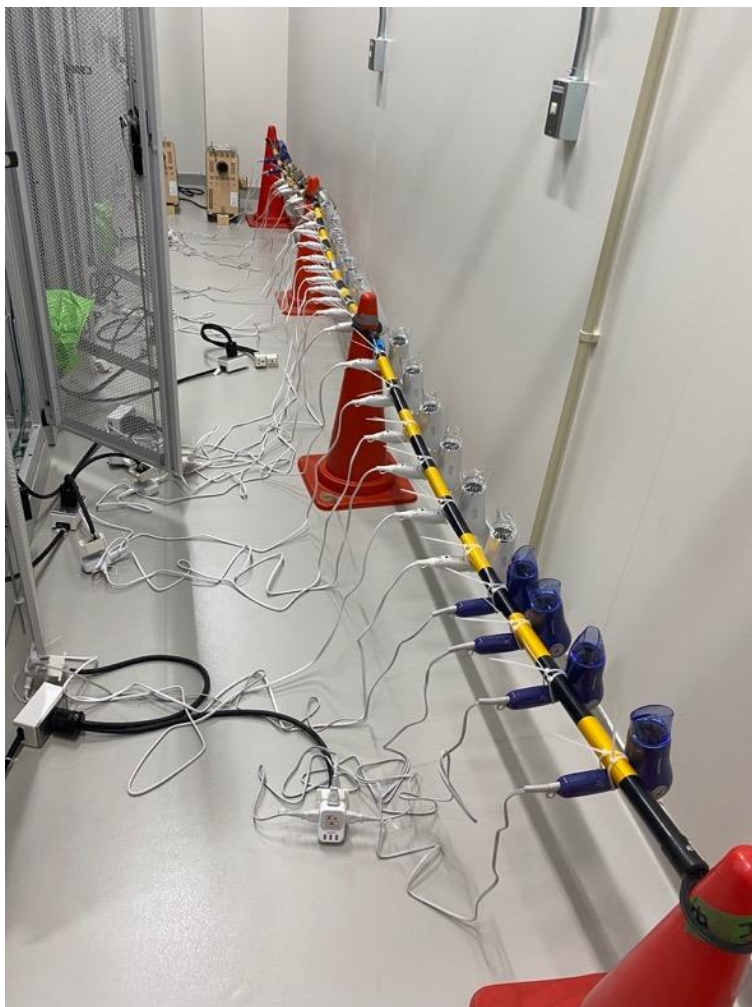


温度 [°C]

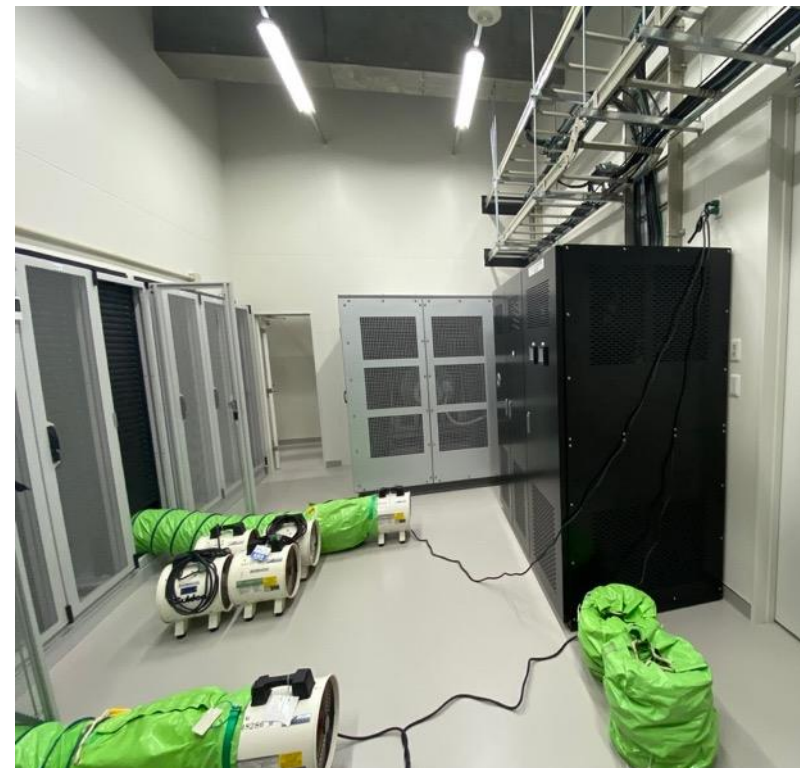


ビデオ

熱源としてヘアドライヤ 60KW



+1KW 150CFMの
風量を押し込む



フリークーリング対応 ドライクーラーで
排熱すると pPUE=1.1を実現。

6.外気導入 実験室 6年後サーバが錆びた



限界を調査する。
「空気質」の問題
結果 日本では湿度が問題

外気導入

外気を部分的に導入すれば、空調電力を10%~20%削減でき、PUE=1.08を目指せる。
 JEITA ETR-3001 Class-Rに適合する場合は、空冷ヘッドに外気を導入する。
 空調機の運転時間が短縮され、省エネ、機器の長寿命化が図れる。

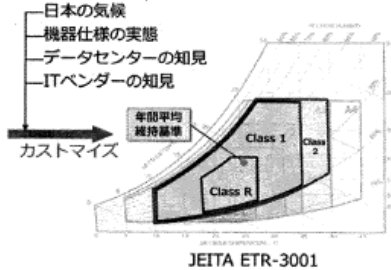
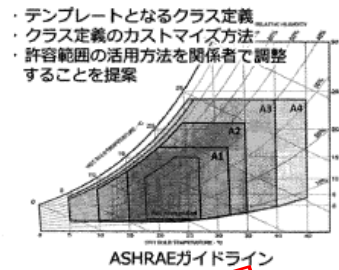
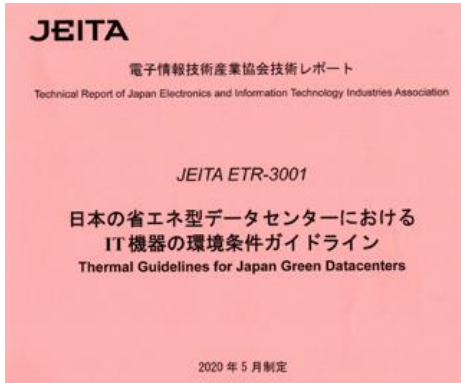


図4-1 JEITA ETR-3001とASHRAEガイドラインの関係

サーバ会社10社が確認



外気導入口 シールド時に湿気の侵入なし



8年の実績 倉敷ChiroroNetデータセンター



部分外気導入モードは、外気を少量導入して、サーバーの排熱と混合して運用します。またサーバーの排熱の一部は、天井より排出されます。

Metaが使う間接外気冷却機 MuntersOASYS 国内実績 12MW 東京都内 PUE=1.07以下 Meta社（Facebook社）より省エネ高性能



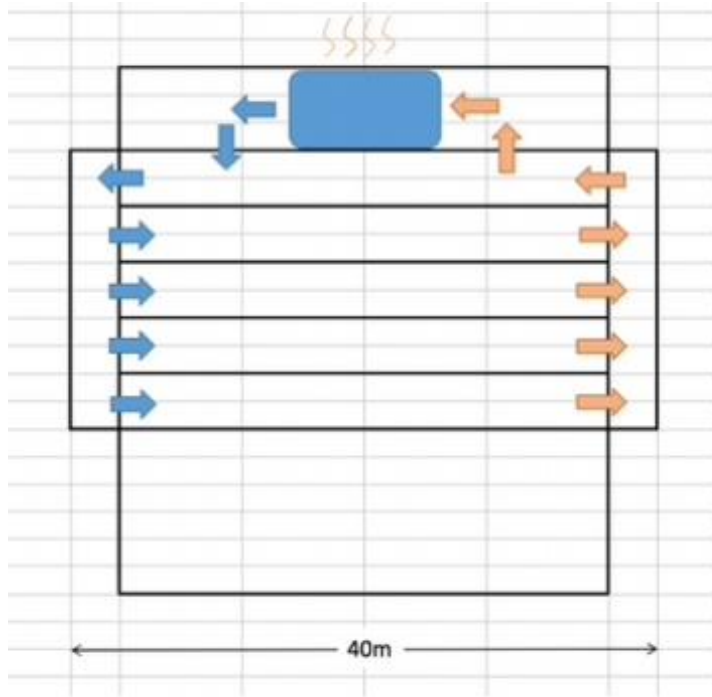
外部に大型ダクトを配置
PUEが低く消費電力は低いが
建設コストが高い

屋上に設置出来る台数に制限
平屋構造だと省エネにならない

湿度侵入対策済み
搬送動力削減にも成功

Munters OASYSはNTTファシリ
ティーズが取り扱っている。

間接外気排熱機 国内実績 12MW 東京都内 PUE=1.07以下 Meta社（Facebook社）より省エネ高性能



5層上昇気流を使う
直角ラック配置で静圧を制御
コールドアイル・ホットアイル間の
微差圧も大きい

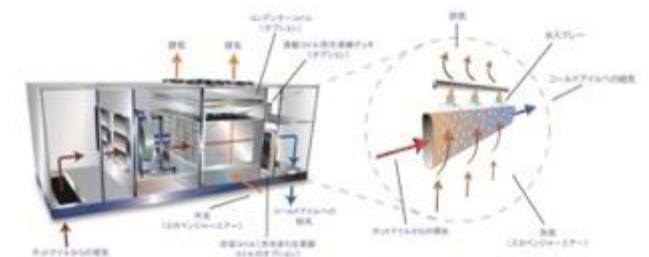
おそらく間接外気冷却機を使う
データセンターとして、省エネ 世界一

チルドタワーより大型機なため
40m²屋上に排熱量10MWしか設置出来ない。



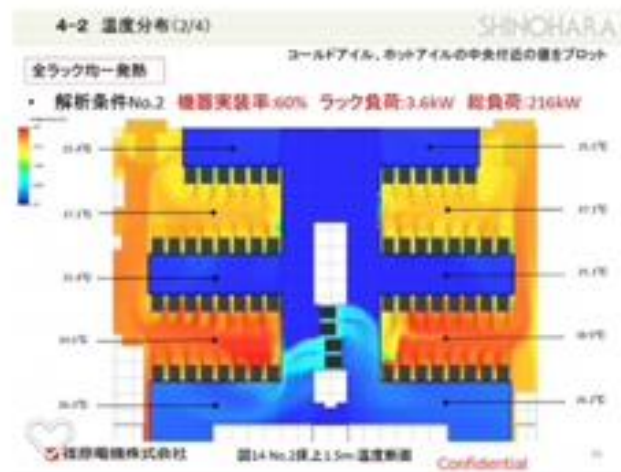
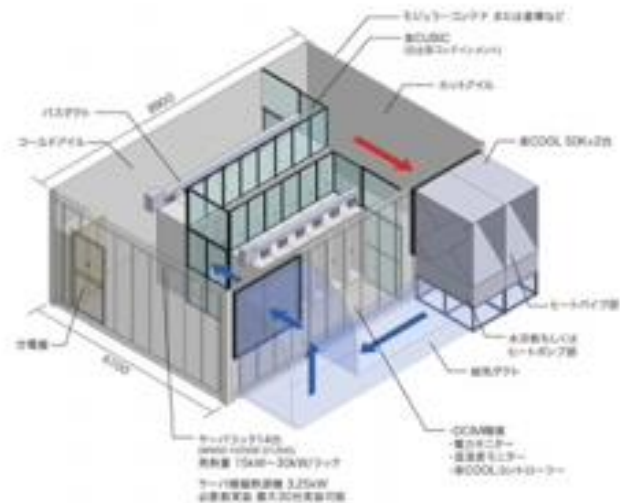
Indirect Air Side Economizer

Polymer Tube Hybrid Heat Exchanger (Direct Water Spray on HX)
(wet bulb during hot dry conditions and dry bulb during cold conditions)

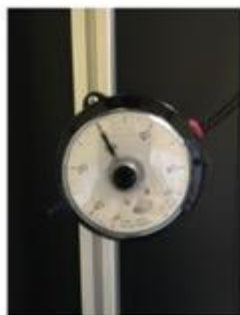


参考：東京大学 次世代データセンター勉強会 間接蒸発式空調 第28回

サーバー排気温度を上げて、 排熱効率を上げる「くし形キャッピング」 篠原電機と共同開発



空冷ファン2台で
400KW排熱



実証試験

Meta (FaceBook) が使う間接外気導入機 全世界300MW10カ所で使われている。
最近チューニングが進みPUE=1.08達成とアナウンスがあった。
このMuntersOASYS部門はMetaに買収され、入手困難。



湿度が上がると能力が落ちる。270KW→190KW
特殊なプラスチックを使っているので蒸発で熱を取れる。
流す水は再生する仕組みを持っているが、大量に消費する。
排気ファンの騒音が高い。

1ラック10KWが限度なので、
NVIDIA GDXが1ラック2台
しか搭載されていない。

6.サーバ室CFD

流体シミュレーション
サーバ向かい合わせ

1ラック

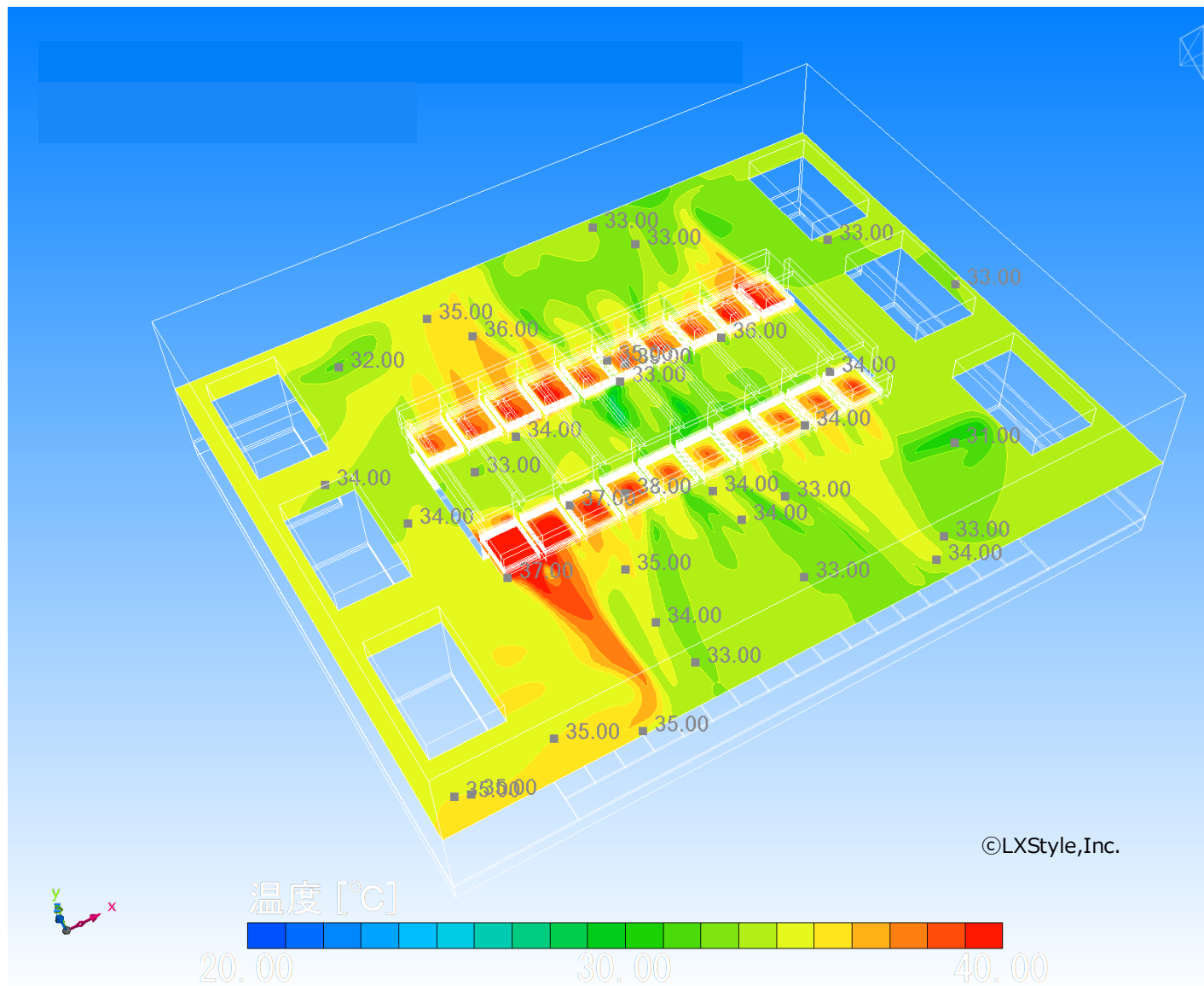
500W x 12 = 6KW

6KW x 20ラック = 120KW

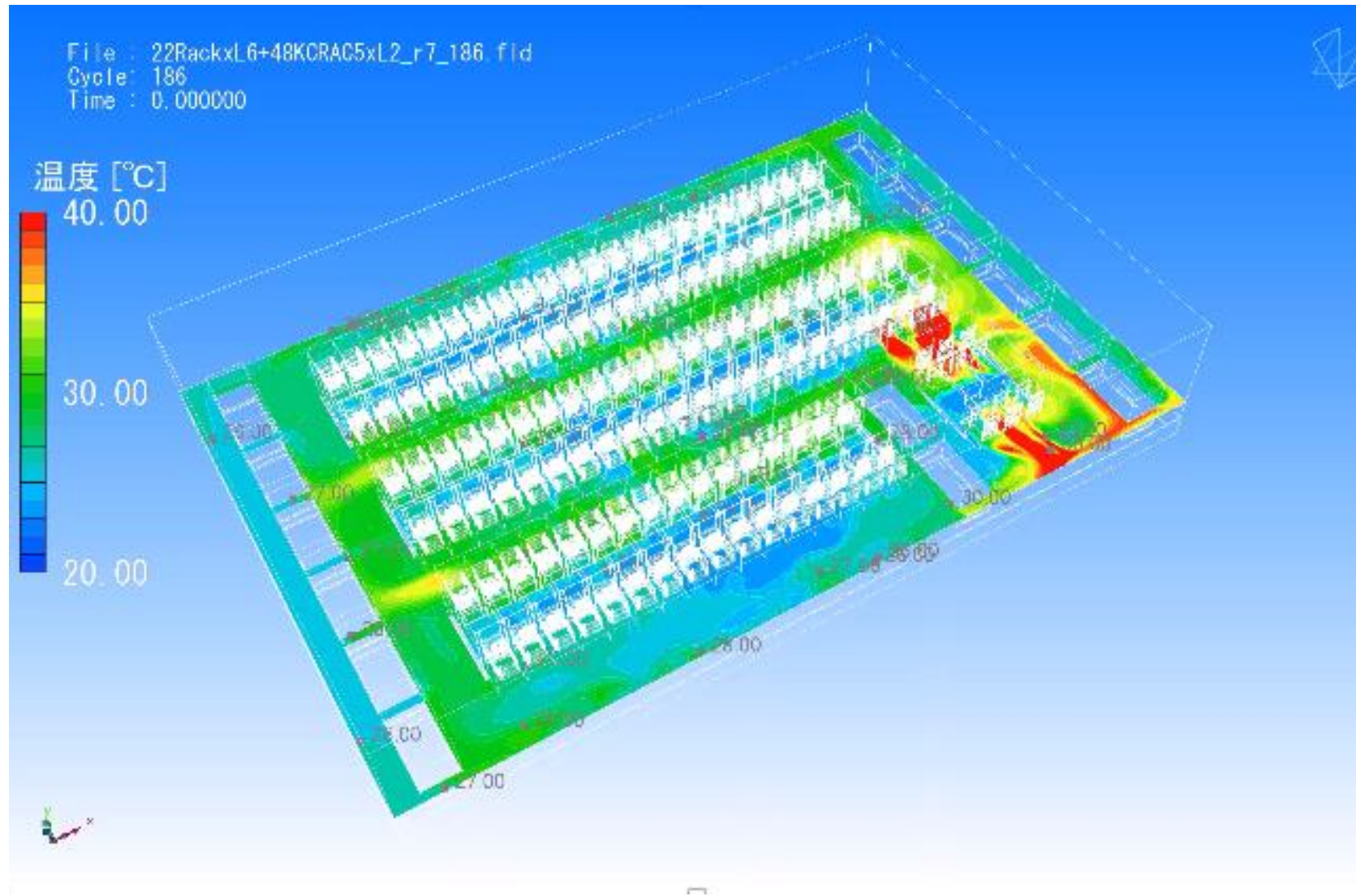
空調機56KW
(稼働率50%) x 6台 = 168KW

コールドアイル
キャッピング
が効いている。

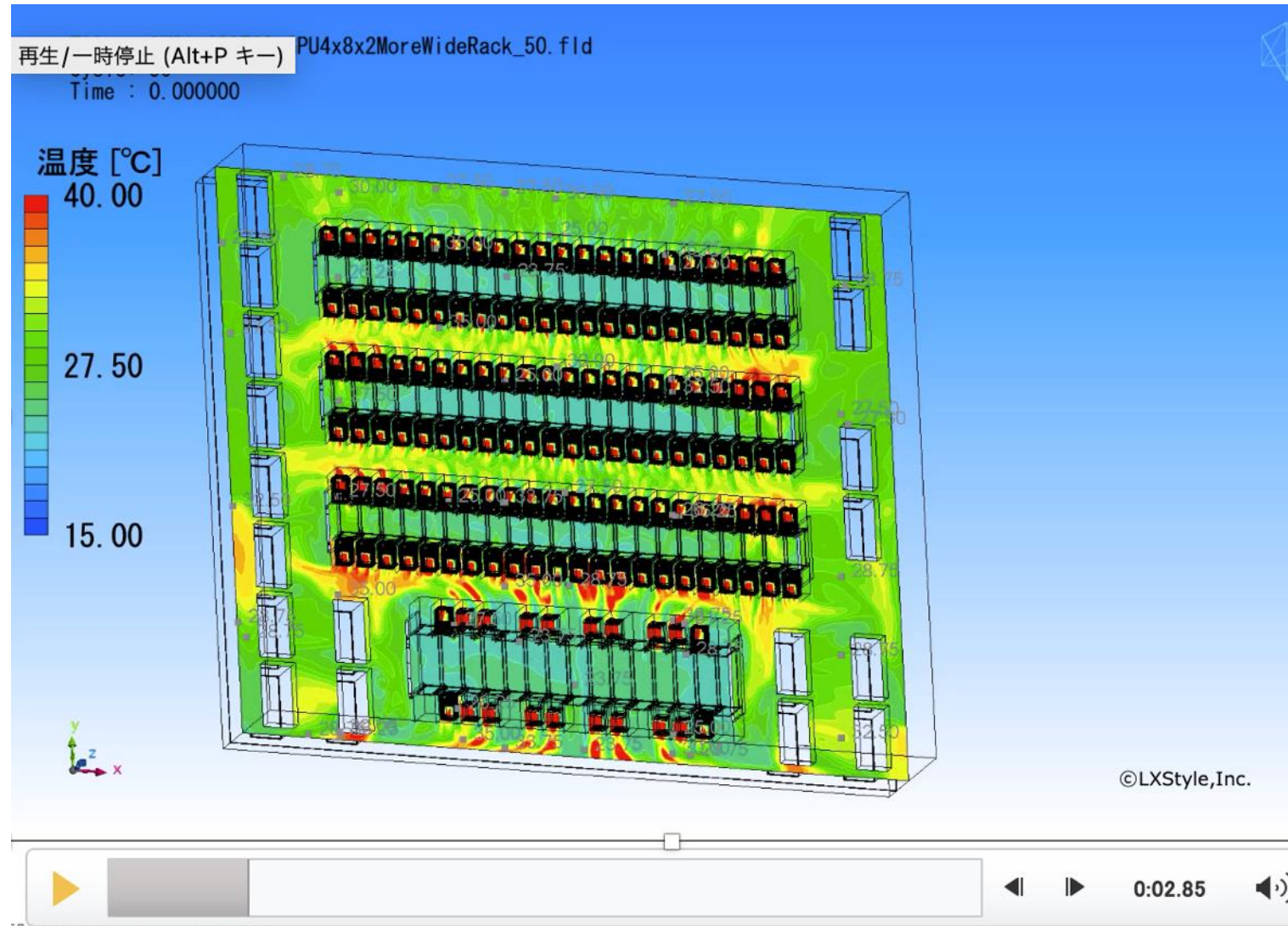
10ラック x 2列 = 20ラック



5KWGPUを3台 6 Rack CRACエアコンを2台追加・・・OKか？



GPUゾーンを長くした コールドアイルに熱風が無い OK!





Analyst Finds AI Boom Lifts Hyperscale Cloud Spend, Propels Data Center Infrastructure, As NVIDIA Presides

We look into a trio of ongoing cloud and AI data center sector reports recently released by industry analyst Dell'Oro Group.

引用先 : [https://www.datacenterfrontier.com/hyperscale/article/55017061/analyst-finds-ai-boom-lifts-hyperscale-cloud-spend-propels-data-center-infrastructure-as-nvidia-presides?o_eid=9847F3472278D4D&oly_enc_id=9847F3472278D4D&rdx.ident\[pull\]=omeda|9847F3472278D4D&utm_campaign=CPS240411163&utm_medium=email&utm_source=DCF+Newsletter](https://www.datacenterfrontier.com/hyperscale/article/55017061/analyst-finds-ai-boom-lifts-hyperscale-cloud-spend-propels-data-center-infrastructure-as-nvidia-presides?o_eid=9847F3472278D4D&oly_enc_id=9847F3472278D4D&rdx.ident[pull]=omeda|9847F3472278D4D&utm_campaign=CPS240411163&utm_medium=email&utm_source=DCF+Newsletter)

この時代を15年前に予想した本がある。

7. 産総研GreenITプロジェクト(2007~2012年)

PUE ISO化委員長 伊藤智先生の著書 2010年

省エネ対策バイブル



た。熱通過のみで誘熱する空調システムセンターの設置環境としては、外気温が低いことに加え、過度な空気の流れがあることも有利に働くだろう。

5) 高温環境下でのサーバーの動作
放熱だけで内部の温度管理をしようとすると、どうしてもマシナリー内部の温度はある程度の高温になってしまう。一般的なサーバーの動作保証温度の上限は35°Cとされている。フロア内部の温度、特にサーバーの吸気側の温度を35°Cまでコントロールできればよいが、放熱だけでこの温度を実現することは難しい。サーバーが高温でも動作するのであれば、放熱のみでの誘熱の可能性も高まるだろう。ただし、サーバーには電解コンデンサなどの高温に弱い部品が使用されており、チップも高温で使用すると寿命が短くなるなどの課題もあり、長期間の安定運用を想定した設備、安定運用の実証テストも必要になる。

では、サーバーほどの程度の温度まで動作するだろうか。吸気温度が高い場合のサーバーの動作を調べた例を図4-37、4-38に示す。サーバーの前面(吸気側)、背面(排気側)などに温度センサーを設置し、吸気温度を上げながら消費電力を測定した。図4-37では排気温度は吸気温度より5~10°C程度高い。吸気側の温度が35°Cを超えても突発故障がなかった。特にシステムの動作の異常は見られなかった。吸気温度が30°Cを超えるあたりで内部のファンの回転数が増し、消費電力が増加している。また、図4-38は別のサーバーでの吸気温度を変動させて機器前後の温度を測定したものである。このサーバーでも突発故障が起きたこともあり、吸気温度が40°C前後でもシステムに異常は見られなかった。

このように短時間の測定ではサーバーの動作保証温度の上限である35°C程度ではサーバーに異常は見られなかったが、情報シス

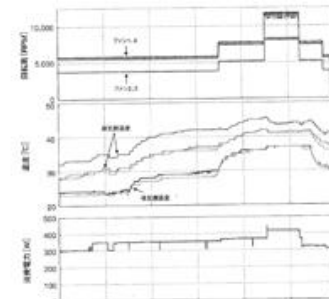


図4-37 高温環境下でのサーバーの動作

テムの基盤としてサーバー、データセンターを考えると、高い温度で継続して運用した場合にハードウェアの故障やシステムの動作に不具合が発生しないか、パフォーマンスが低下しないかなどを調べていかなければならない。

また、温度と電力の測定結果からわかるように、データセンター全体の省エネを考える場合には、吸気温度とサーバーの消費電力の関係にも注意してはならない。空調を使わず、もしくは空調の設定温度を従来よりも高く設定して空調に要するエネルギーの削減を目指しても、サーバー筐体内の冷却ファンの動作にともなってサーバーの電力消費が大きくなり、期待通りの省エネが実

15年前に高温環境下でのサーバ動作を検討している

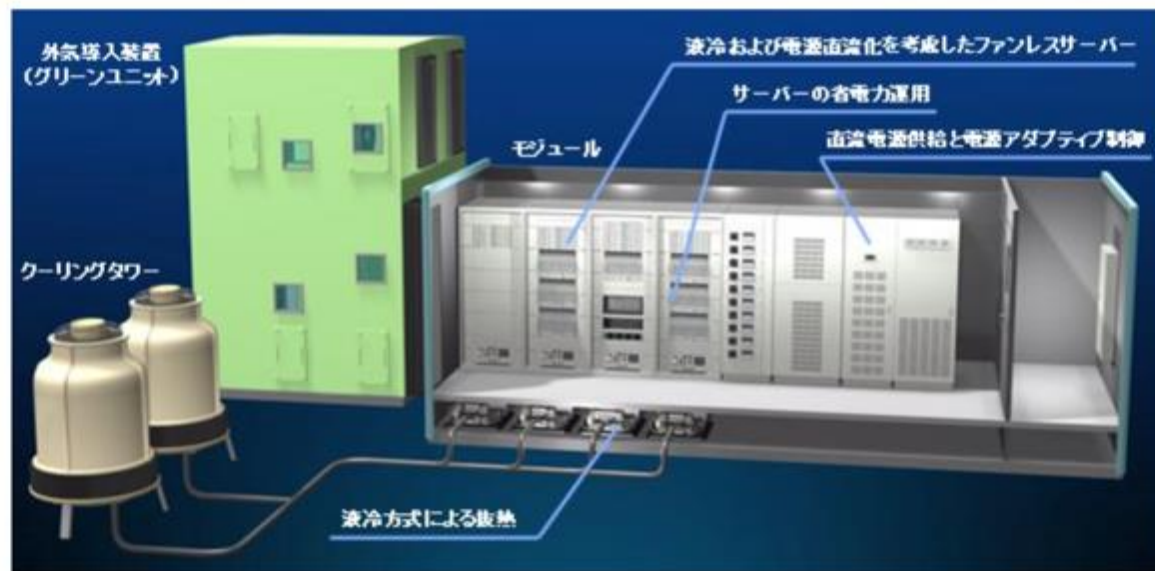
伊藤先生がムントースを英国展示会で見つけてきた。データセンター高温化で多大な省エネ貢献を果たしました。

本の内容を実証した試験設備

間接外気導入機、外気導入システム、直流電源化、DLC（直接液体冷却）サーバなどの研究開発

NEDOは、2008年度に開始したグリーンITプロジェクトにおける成果の一つとして、省エネ技術を結集した次世代モジュール型データセンターを構築しました。これまで開発してきた省エネ基盤技術である、高電圧直流電源技術、サーバー液冷技術、グリーンクラウド運用技術、データセンターモデリング・評価技術に、今回新たに開発した外気導入技術（特許出願中）を組み合わせ、エネルギー利用効率を最適化し、本事業の目標とした総消費電力を従来に比べ30%削減（比較のために構築した従来モジュール型データセンターの総消費電力28kWが次世代型で19.6kW以下に）できることを検証します。また、商用電力の供給量が制限された際に、制限内で効率良くデータセンターの運用が行える運用技術も開発します。

クラウドコンピューティングの普及とともに震災以降需要の高まるデータセンターにおいて、省電力は必須の技術となっています。技術開発を進めていくことで引き続き持続可能な社会の実現に貢献することが期待されます。



次世代モジュール型グリーンデータセンターの構成図

NTTグループと研究結果、データセンター室温を上げると大幅に省エネになることがわかった。

コールドアイル30℃を目指して省エネ化

産総研はPUEのISO制定に協力。 このころからDLC（直接液体冷却）サーバを研究開発

(1) 液冷を用いたファンレスサーバ

データセンターにおけるサーバの集積密度は年々増加傾向にあり、数年前は2 kVA～6 kVAだったラックあたりの消費電力は、近年では8 kVA～20 kVAにもなろうとしています。サーバ内で発生した熱は、通常サーバ内のファンによってサーバ外に排熱されます。ラックあたりのサーバ集積度が上がるとデータセンターのサーバ一室内に排出される熱は非常に膨大となり、空気中に分散した熱をデータセンター外に排出するには強力な空冷装置が必要となります。そこで、空気よりも熱伝導率が高い液体を用いて、熱をサーバ室内に排出せずに効率的に除去する手法を採用しました。

今回、株式会社SOHkiがNEDOグリーンITプロジェクトの一環として開発した液冷ジャケットを既製品サーバに装着し、大部分の熱を冷却液によりサーバから除去することとしました（図1）。

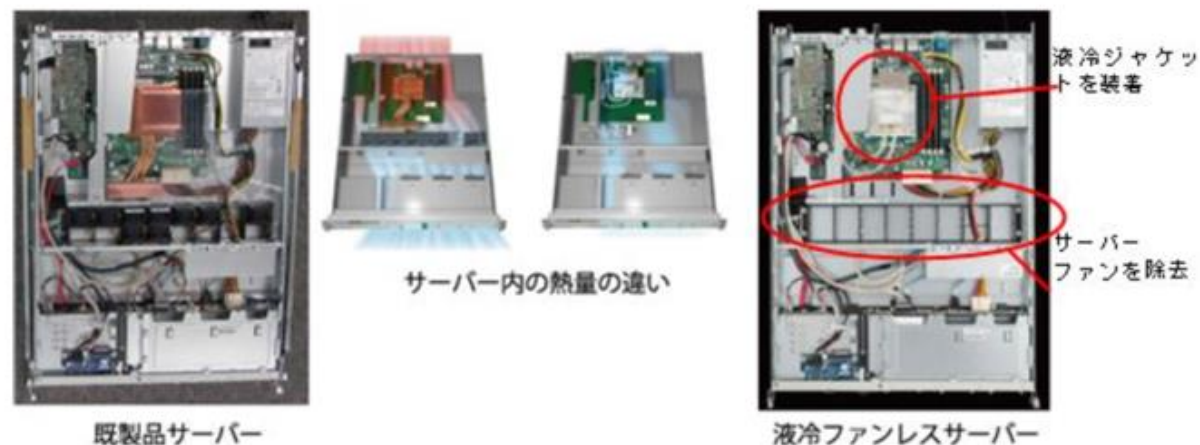
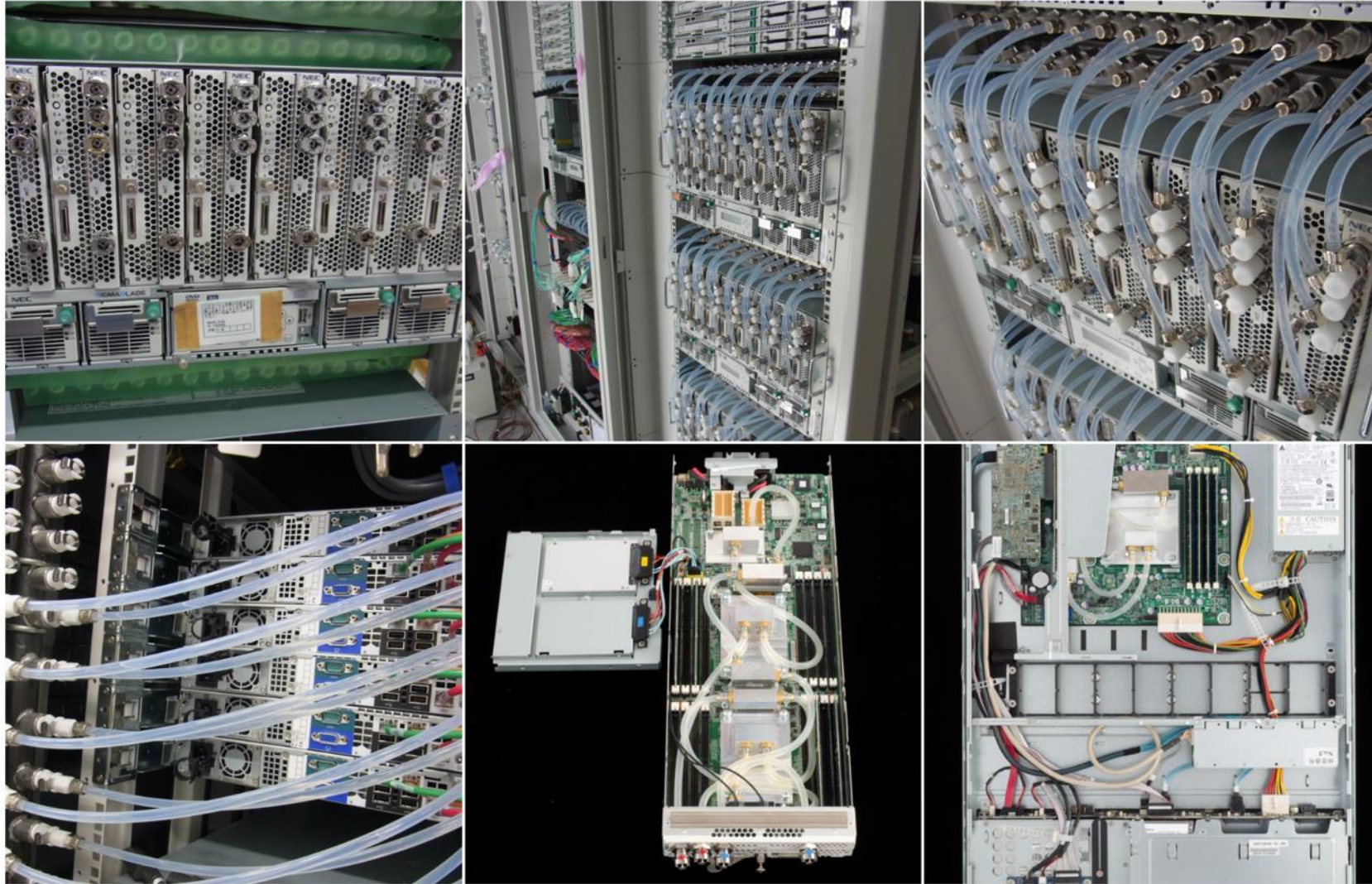


図1 既製品サーバと液冷ファンレスサーバ

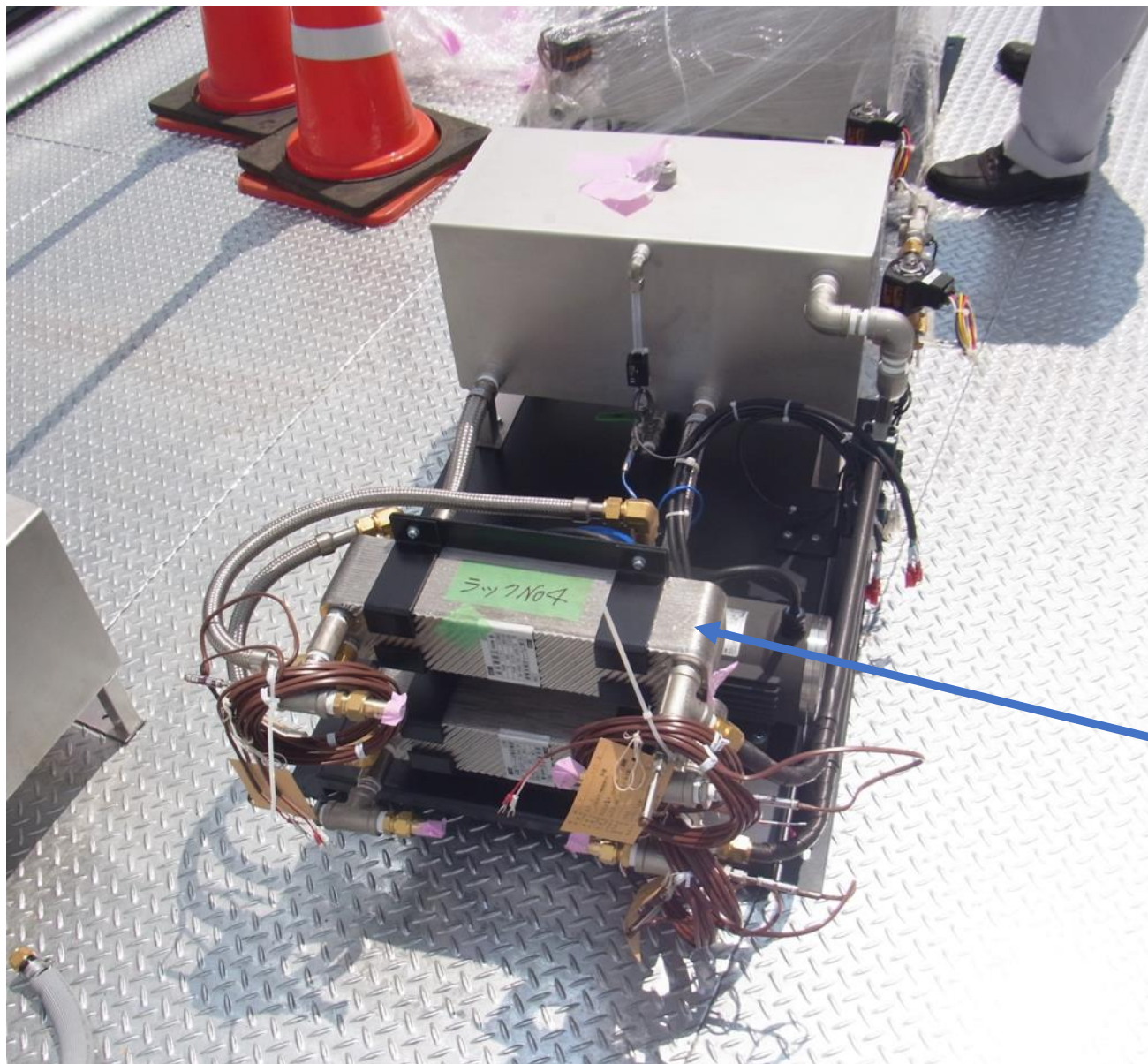
NOVECを使用
フッ素系液体は
GWPが高いので
産総研では使用禁止

絶縁性があるので
フッ素系液体だったが
比熱が低いので流量を
上げるためポンプ電力
が上がる

15年前からNEDO予算で、DLCサーバを開発していた。



筆者撮影 場所 つくば 産総研



コンテナで 使ったCDU

小型プレート式
熱交換器

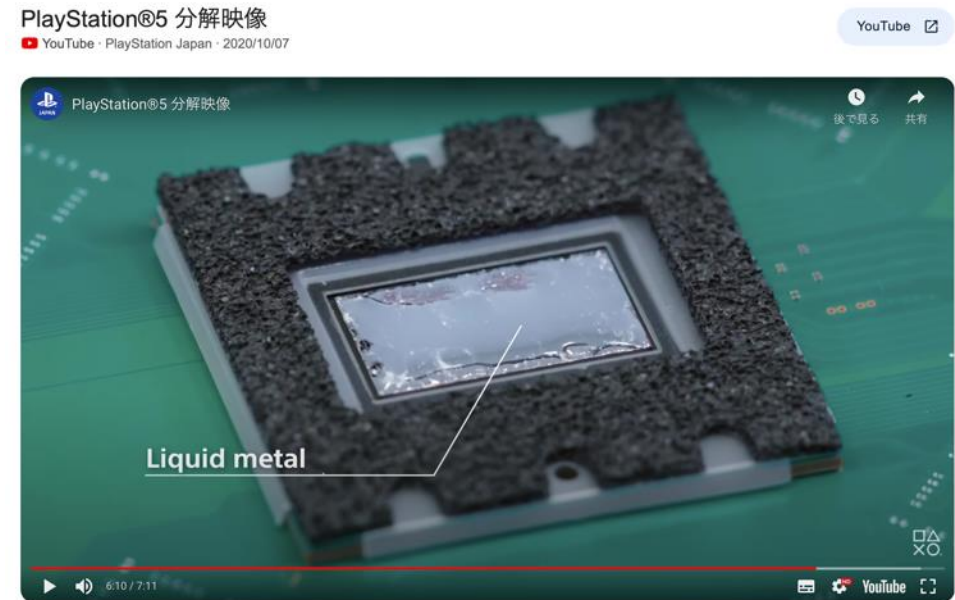
筆者撮影 場所 つくば 産総研

余談1：速さを求めた15,000回転HDD Cheetah
RAIDでサーバに装備されたが冷却風量不足で数年で故障
SUNが潰れた原因と思っています。
どうして？HDDキャニスタを熱が伝えない樹脂で作った。



筆者撮影 場所 つくば 産総研

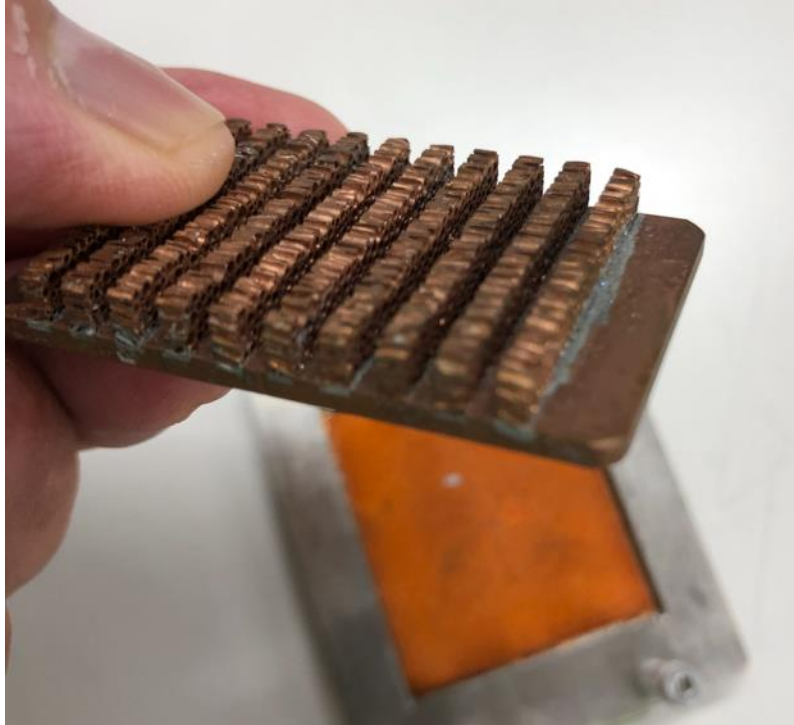
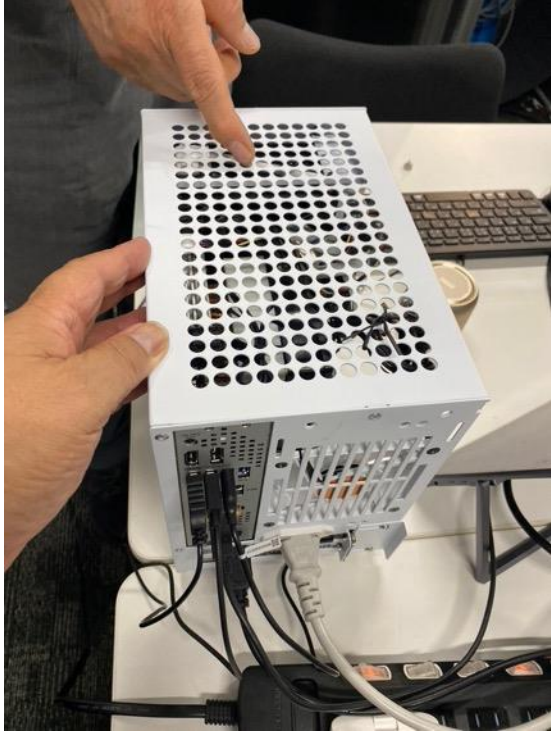
余談 2 : SONY PS5 冷却システム ホコリを吸わずに、風量を下げて排熱



https://youtu.be/iLKvWhcA_KU

集中した熱を 液体金属Liquid metal で拡げて、専用ファンで低速で排熱している。

余談 3 : CPU熱を高効率に排熱 するとPCを小型化出来る。



TDP200WクラスがGPUとともに小型ケースに収納

8.空冷と水冷のハイブリッド排熱

大量の熱を発生源の特徴に合わせて水冷DLC（直接液体冷却）する。
CPU/GPUは高温OKなので熱を取りやすい。

残ったメモリやVRM熱は空冷する。
空冷は10KW程度必要で、簡単に実現出来ない。

産総研スパコンABCIIは7年前に設計施工、PUE=1.1以下を実現してきた。



- 経産省「人工知能に関するグローバル研究拠点整備事業」(H28二次補正)の一環として整備
- 我が国における産学官によるAI研究開発を加速するオープンイノベーションプラットフォーム
- 高い計算能力を活用したAI技術の研究開発・実証、社会実装の推進、AI分野の最重要課題への挑戦が目的
- 新型コロナウイルス感染症対策に無償提供
- 経産省「人工知能に関する橋渡しインフラ拡張」(R1補正)により昨年度末、大幅アップグレードを実現



2018年8月1日運用開始

2MW運用
設備容量3.4MW
コールドアタッチ型
PUE=1.1

5,000L/min
省エネポンプや配管
構造で省エネ世界一？

GPUを使うAI演算ベンチ
マークは2回世界一

レポート

産総研のABCIスパコンが大幅アップグレード

🕒 2021/05/07 08:30

著者：Hisa Ando

引用先：<https://news.mynavi.jp/article/20210507-1884603/>

夏期の日中は、32℃の水で排熱する「ホットウォータークーリング」

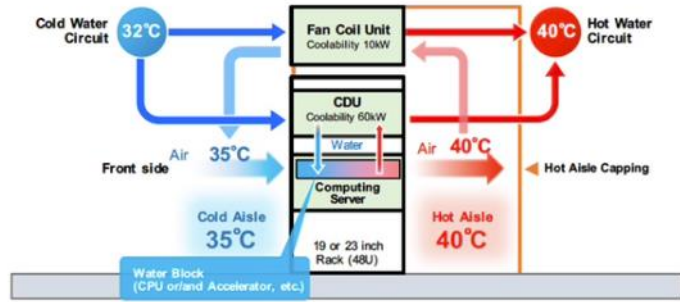
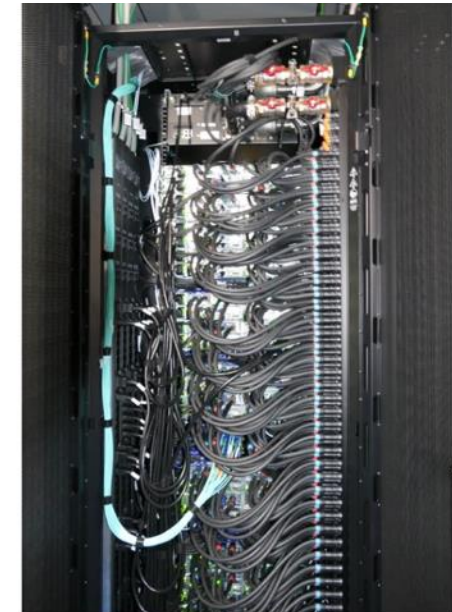


図 3 冷却装置の概要



空気に比べて「水」は、約1,000倍の熱搬送力があるので、
要求に応じて拡張（スケール）しやすい。



FCU故障時対策 レンタル600KW空調機

サーバCPUクーラー



背丈の高いフィンが使えない
→風量を増やす
→轟音を発生する



1Uだと
→放熱面積も足りなくなる



マイクロチャンネル熱交換器

溝付きの板を積層し、金属の拡散現象で接合した構造の熱交換器です。プレートフィン型よりも、肉厚が大きく取れるため、高圧・熱サイクルの厳しい環境での使用に適しています。



パワーデバイスクーラー

弊社のろう付け技術を応用した、高性能なパワー半導体用の冷却器です。通常の水冷・空冷だけでなく、相変化を利用した製品も扱っています。

一般PC用 おもに静音化が目的
ファンスピードを落とせる



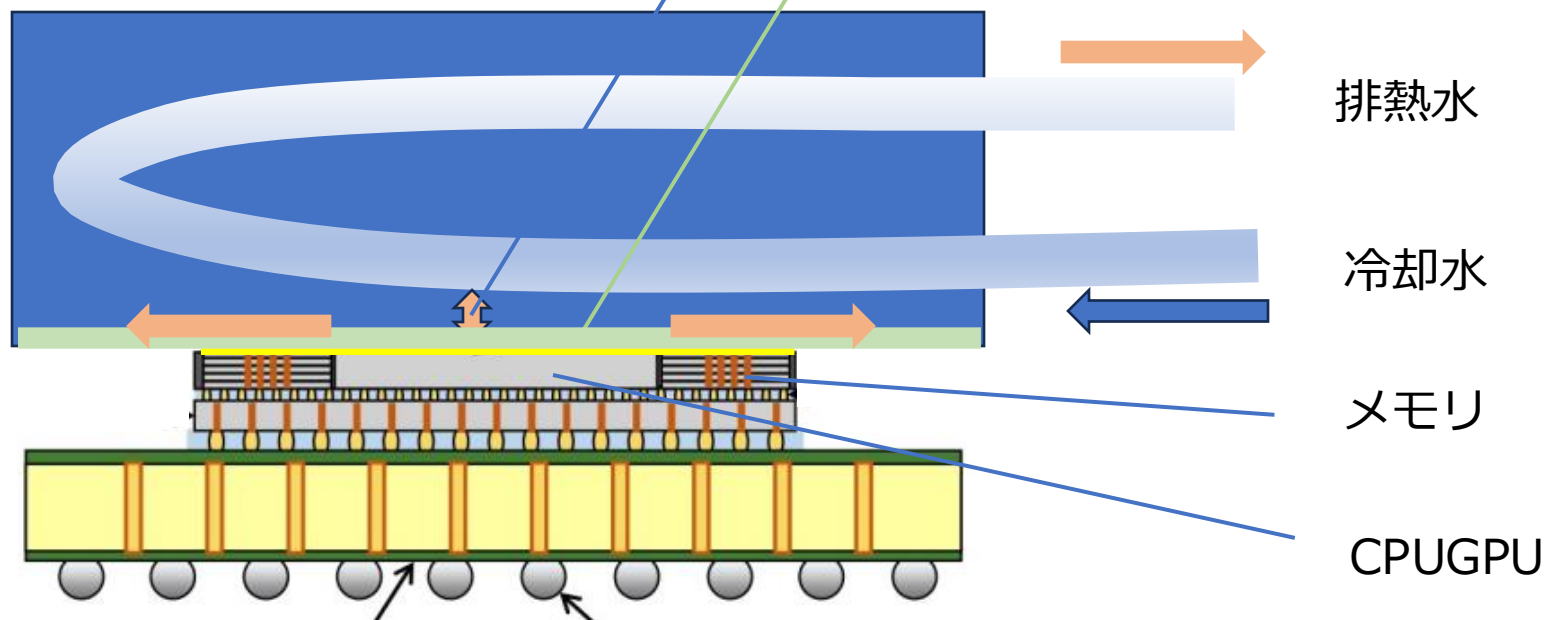
引用先：住友精密工業<https://www.spp.co.jp/netSU/products/>

引用先：https://www.dospara.co.jp/5info/cts_str_pc_wcooled.html

普通の素材は、
熱を一方向にしか
送れない。

この厚みが問題になる

横方向に高速大容量で熱搬送出来る素材を使い排熱する。
例：炭素セラミック、人工ダイヤモンドなど



Lenovo メモリも電源もDLC化 空冷がほぼ不要 2年前 ドイツ最速スパコンで採用 34°Cホットウォータークーリング



小型ヘッドで分岐配管
内径φ8~φ10

流速 不明
水圧 不明
CDUの設定に依る。

※聞くと すぐに
NDAと言う (> <)
どうやって?
付帯設備を設計する?

日経クロステックNEXT東京2024
筆者撮影

最近のサーバ 1ラック100KVA超え

HPEサーバを多く作っているInventec社

Inventec Artemis Rack Solution



MGX Design
Comply with 1U
2P4GPU GB200

Hi-power density design

- ~120kW for Oberon building blocks
- Power Bus bar – 1400A
- 4X 66kW power shelves (N+0 / full rating)
- Delivers 132kW per rack (N+N)

Liquid cooling solution

- Support different liquid cooling solution: in rack/sidecar/row based
- IEC self-developed CDU to comply with customer requirement
- Leakage detection solution (additional 1U tray)

ORv3 Rack integration

- 44OU ORv3 rack
- Oberon building blocks can also fit in EIA rack
- 1U 48*10G/1G TOR management switch
- Great serviceability with blind-mate design with NVLink Cartridge and Manifold



1ラック120KVA
UPSの方が大きく無い？
床対荷重が問題になる。

40MWデータセンターでも
400ラック？

※従来は1ラック10KVA
6KW として 12MWで
1,000ラックは置けた。

•
•

PUEが悪いと損害額も大きい



背面の写真

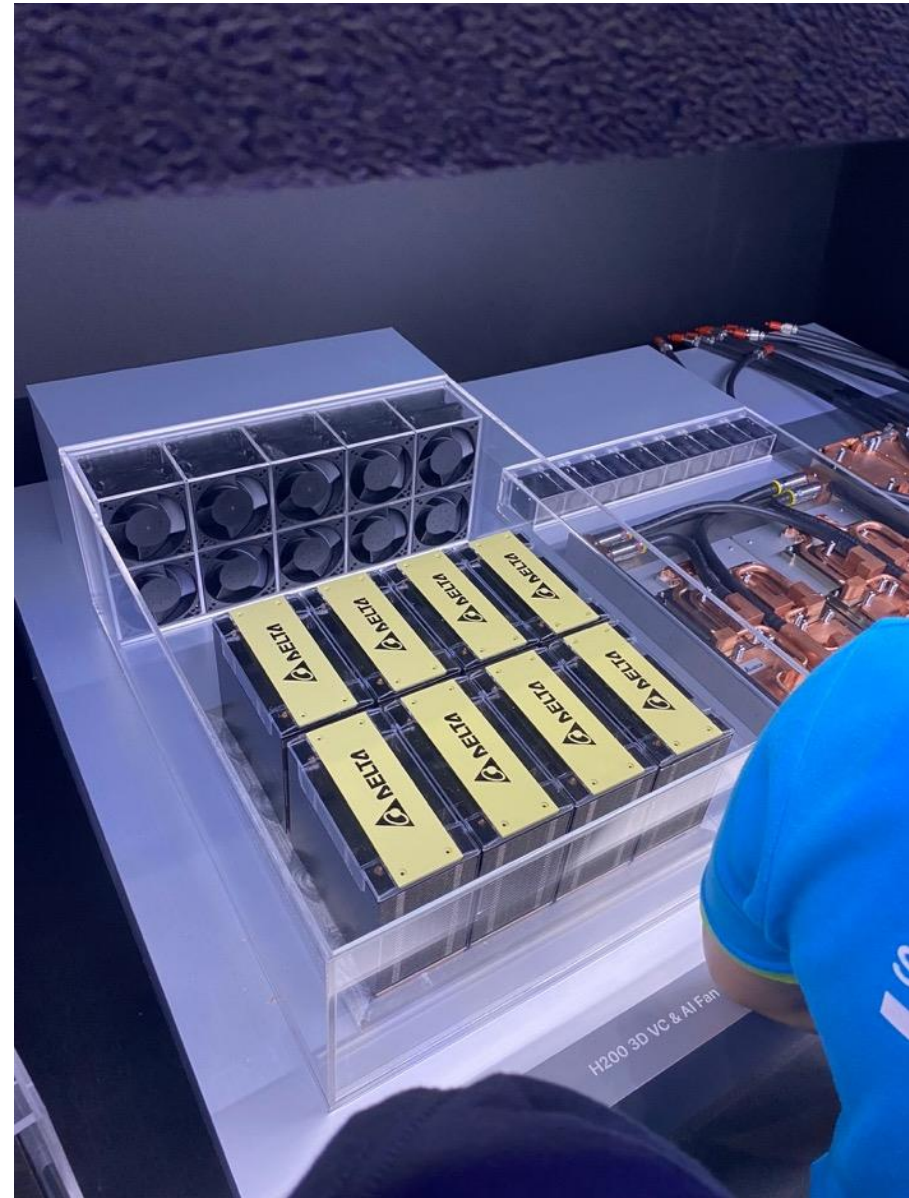
動かないですよね・・・4億円のモックサンプル

120KW/1ラックだが、

40Aシングル

50KW排熱が限度

- ・流速を上げたら？
CDUに内蔵されるプレート熱交換器
の性能で制限される。



ComputexTaipei2024 筆者撮影

GigaByte GPU8基 サーバ 空冷のみで評価中



データセンターの
風流設計を本サーバ
に合わせて行えば
DLCサーバは不要

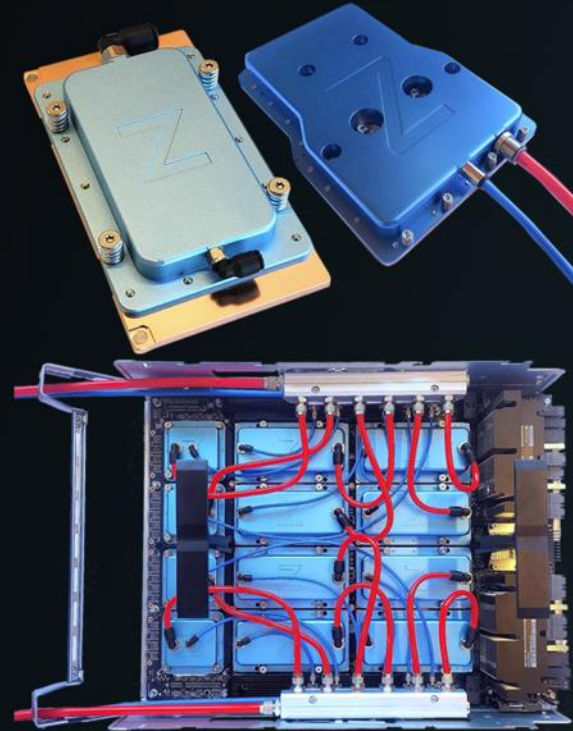
(株) コアマイクロシステムズ 戸田研究所 筆者撮影

WATERLESS コールドプレート

ZutaCore HyperCool最先端の水なしコールドプレート技術で先進の冷却を体験してください。

水なしコールドプレートは、CPUまたはGPUを冷却するためにマニホールドから供給される熱伝導流体を気化させ、その結果生じる蒸気を熱除去ユニット (HRU) に自然に流し、凝縮させて液体に戻します。この低圧オンデマンド冷却システムは、2800ワット以上を処理でき、最適な熱管理を保証します。

この環境に優しいシンプルでスケーラブルなソリューションは、電力使用量を82%削減し、計算密度を高め、データセンターに持続可能性をもたらします。Dell、Super Micro、Asusなどの主要なサーバーメーカーと互換性のあるHyperCoolは、スロットルなしで卓越したパフォーマンスを提供する持続可能な冷却ソリューションを求める企業に最適です。



NVIDIA
Bシリーズ
TDP500W対応

多くのサーバーメーカー製サーバに対応。サーバ内部にヘッダを持ち、専用マニフールド、専用CDUと組み合わせて使う。特殊な冷却液を使い気化熱で効率を上げている。

引用先：<https://zutacore.com/solutions/>



DLCサーバ + NidecCDU ColdPlate



Nidec高圧ポンプを使う
高圧ポンプは冗長化されていて
ホットスワップ出来る。

Coolant Distribution Unit (CDU)

Supermicro CDU contains 1+1 redundant hot-swappable pump systems that circulated the coolant to cold plates cooling down CPUs and GPUs heat. The CDU cooling capacity is up to 100kW enabling extremely high rack densities.



CPU/GPU Cold Plate

Cold plates are placed on top of CPUs and GPUs to cool down chips efficiently by letting coolant flowing through micro channels in the cold plate.



CPU Cold Plate



GPU Cold Plate

Coolant Distribution Manifold (CDM)

CDM are the distribution pipes that supply coolant to each server and collect the hotter coolant back to the CDU.



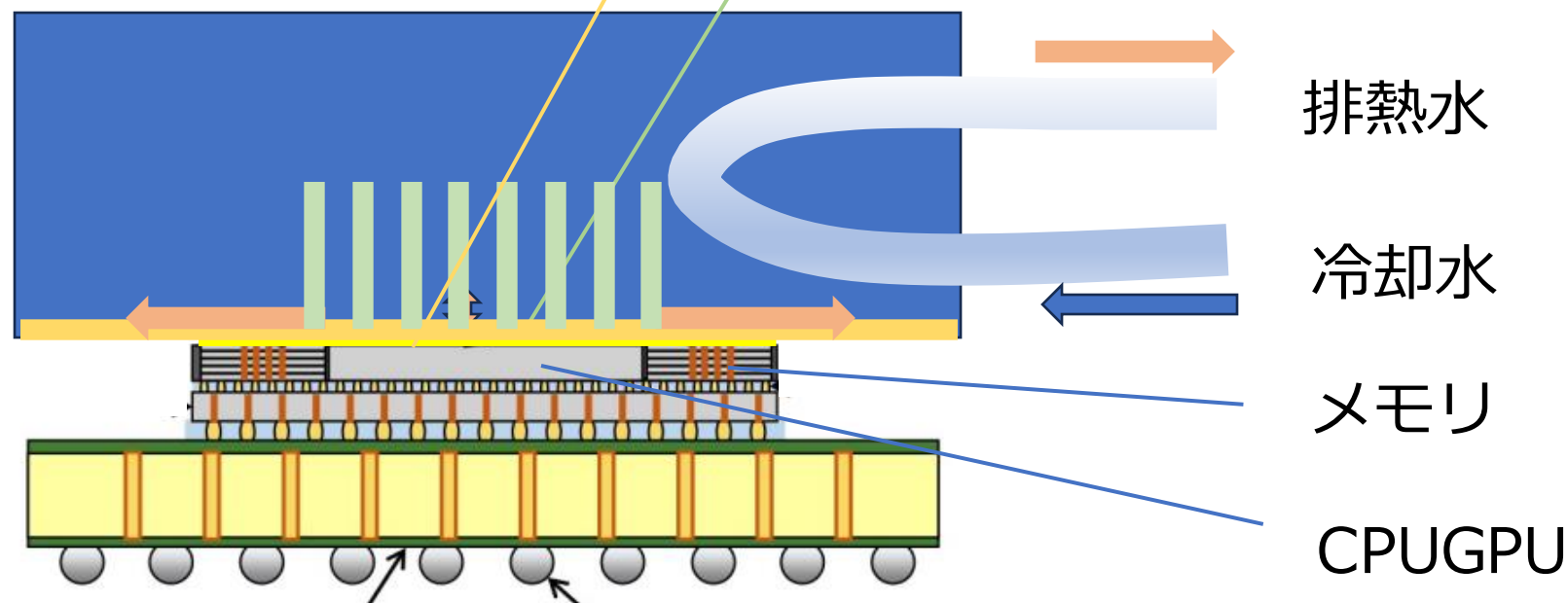
Nidec水冷システムは月産3,000システムと発表

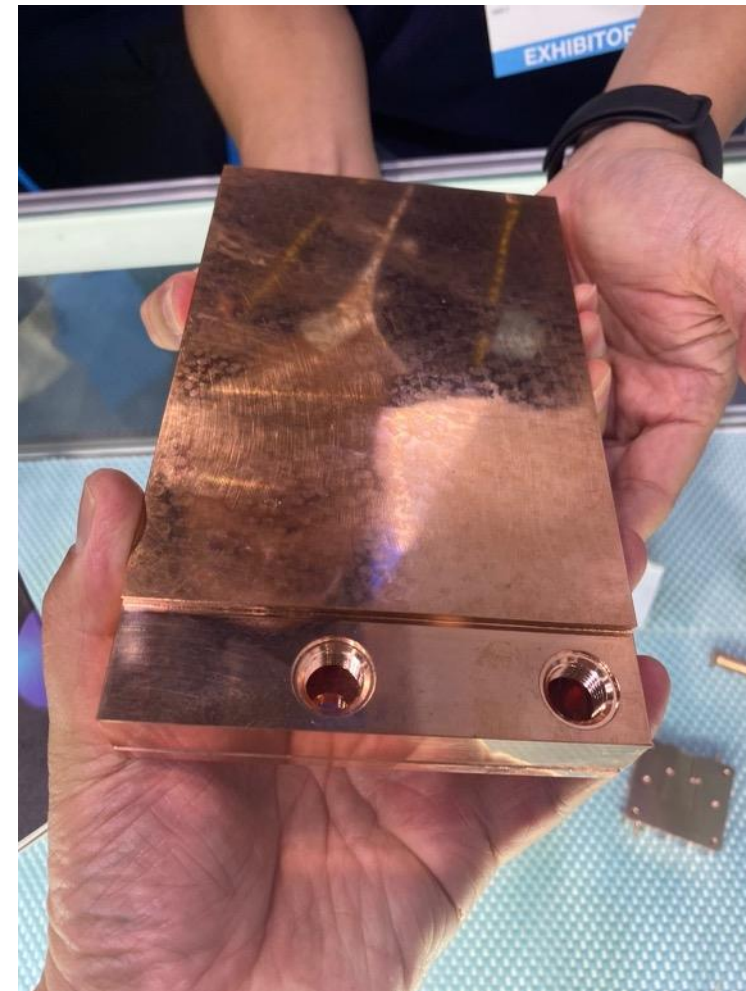
ColdPlateの課題

普通の素材は、
熱を一方向にしか
送れない。

この厚みが問題になる

横方向に高速大容量で熱搬送出来る素材
例：炭素セラミック、人工ダイヤモンドなど





ComputexTaipei2024で見かけた大型コールドプレート



ComputexTaipei2024 筆者撮影

最後に

コールドプレート高性能化で1ラック120KW排熱が可能になっても
空気中に熱を排熱するクーリングタワーやドライクーラーは
小型化出来ない。

ポンプが止まっても、排熱機が経年劣化で処理能力が落ちてても
対策しなければサーバは止まる。

日本の冬は厳しい。10年に1度ゴルフボール大の雹が降るし、、
一晩で1mも積もるドカ雪もある。

フリークーリングは自然を相手にするから“難しい”。

以上