



PCCCワークショップ

2024/6/28

AI/機械学習部会イベント

パネルセッション

生成AI時代の高性能計算インフラの抱える課題 —これからの日本DC設計に求められるもの—

(発言まとめの文責は遠藤

ご指摘・ご修正歓迎。

未来パートのQ&Aは項目のみ)

モデレータ: 遠藤敏夫 (AI/ML部会長／東京工業大学)

パネリスト:

愛甲浩史 (エヌビディア合同会社)

高野了成 (産業技術総合研究所)

矢澤克巳 (インテル株式会社)

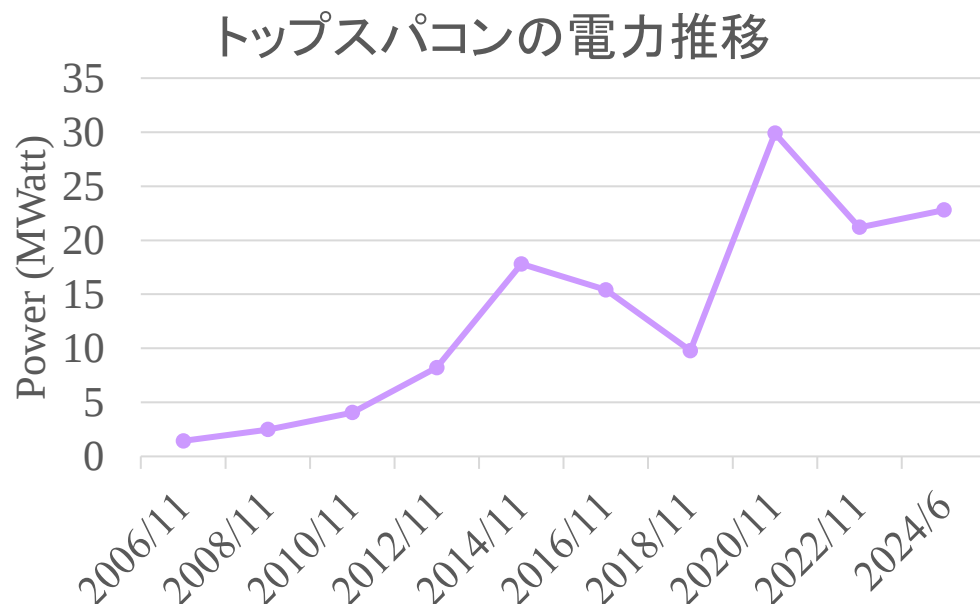
大原久樹 (日本AMD株式会社)

石橋史康 (レノボ・ジャパン合同会社)

司会: 岩谷正樹 (エヌビディア合同会社)

- 2006/11のトップスパコン BlueGene/L
 - 0.28PFlops, 1.43MWatt, 0.20GFlops/Watt
- 2024/6のトップスパコン Frontier
 - 1194PFlops, 22.7MWatt, 52.9GFlops/Watt

17.5年で**速度 4300倍**、**電力 16倍**、**電力性能比 270倍**



1サイトに作れるシステムの電力は明らかに上昇している
GPUシステム増加も影響

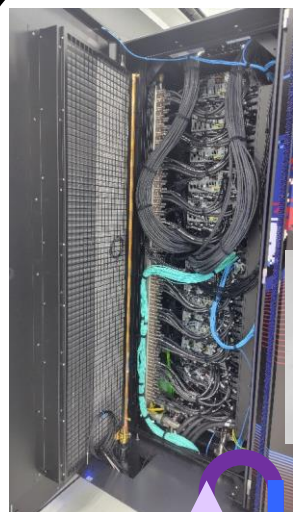
- ➔ 電力上昇させなければ計算需要に対応できない！
- ➔ ましてや、今後AI/LLMの計算需要爆発により、DCとスパコンは近づいていく

全体消費電力 計算機ラック数 23

- 設計値 1.8MW, 通常運用時 0.7MW程度

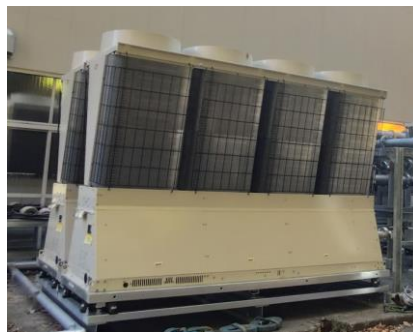
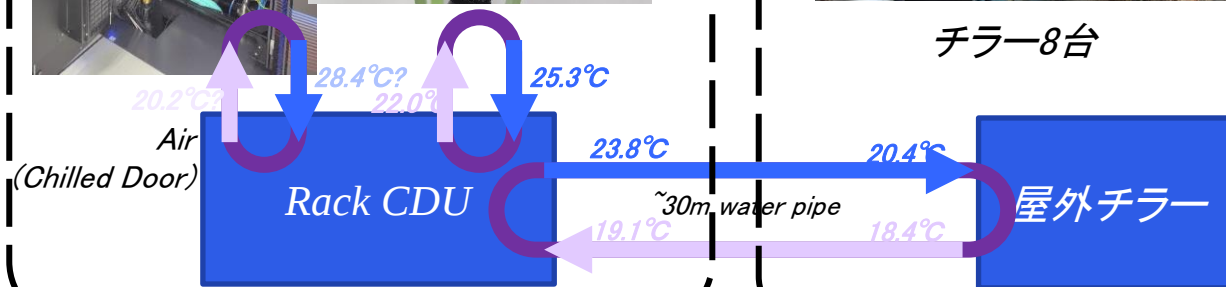
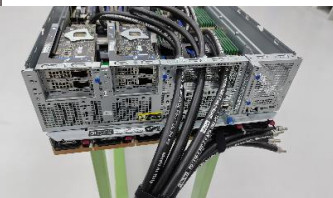
計算機ラックあたり消費電力

- 設計値 55kW, 通常運用時 25kW程度



各Motivairラックに
冷却機能

- 背面冷却ドア
- 最下部にCDU・熱交換



チラー8台

ノードの液冷部分 → 水パイプがラックCDUに接続

ノードの空冷部分 → 暖かい空気は冷却ドアで冷却・CDUに接続

システム全体の冷却は最終的に屋外チラーに依存

※TSUBAME3と異なりフリークーリングではない

※左図の温度はある時点のセンサー値;設計の温度ではない

DCにおける電力について

- 生成AIの急速な普及に伴い、DC・ラック・サーバ・パーツに求められる性能電力比は高くなります。上記の変化に、どう対応してきました？
していますか？
- 今後のDCにおいて消費電力・性能電力比はどう変化すると予想しますか？そしてその対応は？

□ 愛甲

- AIの演算要件: 2018Transformer登場後に激増
- GPUの性能・性能電力比の向上よりも演算需要増が多い
- CPU+GPUパッケージは省エネに

□ 高野

- 2024以降予定の日本のDC例。8-50MW、1000-4000ラック
- 産総研は2.4MW → 5MW
- 今後GPUは高火力。ピーク使いきれず、電力変動が多くなるおそれ

□ 矢澤

- チップあたり電力は上がっている。さまざまな最適化しつつも次世代CPU 500W
- 生成AIアクセラレータも

□ 大原

- CPU 400W, GPU 750W。単純外挿では、2035年ZFlopsで2140GFlops/Watt, 500MW ??
- CPU+GPUパッケージ、3D積層(ロジック・メモリ含む) → プロセスルール微細化はゆるやかになるが、電力はあがる
- Processing-in-memoryも含め、半導体は面白い

□ 石橋

- プロセッサ以外も電力があがっている。DIMM(600W), HCA(150W)
- 空冷の限界; ラックあたり数年後には80-150kW? Fan電力13%、空冷で+30%
- 水冷で冷却電力1/4程度へ

- Q: 海外では発電所つきDCを見る。日本では難しい？
- 高野: 日本ではあまりなさそう。送電網がDCに近くなる例はある。発電所が難しい理由あるか？

DCにおける冷却について

- DC・ラック・サーバ・パーツの冷却手法について、
どういう技術を使ってきましたか？
- 直接液冷や液浸冷却は必要と考えますか？
- 今後のDCにおいて冷却手法・冷却性能はどう変化すると予想しますか？

□ 愛甲:

- GPUの電力密度により空冷だけでは厳しく水冷が避けられない
- 液浸についてはNW含めて経験あるが、機器ベンダーとしては保証的な心配。ユーザと綿密な検証必要

□ 高野:

- ABCIは水冷+空冷、フリークーリング。日本は高温多湿だが、上記動いてきた。真夏のinlet 37°C
- 年間PUE1.1以下を続けてきた。PUEという指標への疑問
- 液浸は、もっと困ってくるとありうるかも

□ 矢澤:

- Intel DCについて。40万台、半導体設計EDA用に利用。
- 昔は天井排気, aisle isolation, free air, 密結合蒸発冷却... ラックあたり5 → 57kW
- DLCなしの空冷で上記を実現してきた。3D積層が進むと液浸も??

□ 大原:

- OEM次第だが、GPU 750W, CPU 400Wが空冷で動いている。数年後以降は水冷が避けられない
- 液浸については、冷媒の環境負荷問題で後ろ向き

□ 石橋:

- DLC、液浸はPUE同じくらい。液浸は今は欠点(warranty, 環境負荷) → DLCメイン。グリコールより純水
- ヒートシンクの面積を上げる工夫で、空冷ぎりぎり動いている

今後日本が作るべきDC像は？

- 今後重要になるのはSAAS・PAAS・IAAS・ホスティング・ハウジング・それ以外？
- DCインフラ(建物・受電設備ふくむ)は変化していくべき？コンテナ型DCは普及する？
- 今後生成AI等の需要の拡大に対応するため、日本が作るべきDC像は？

□ 愛甲:

- オンプレミスvsクラウド問題; その時その時で何を重要とするかによる
- 基本的に大きく作ると効率はよくなる。クラウドの出発の一つは、CPU/Diskは基本ひまなので詰め込もう
- チップは専門化・専用プロセッサが進むだろう。AIだけでなく暗号

□ 高野:

- Altman「計算は未来の通貨」。今後の計算は？
- 量子コンピュータ ready DCも必要では。一般的な施設では×。液体窒素・キャリブレーションも必要(気候にも依存)。Green quantum advantage?

□ 矢澤:

- 国内DCの典型は10kW/rack → 継続的な新陳代謝が必要

□ 大原:

- LLM隆盛にむけたDCは絶対に必要。Tier-1だけでなくTier-2が必要では。アイデアを実現する場。DCインフラの変化の必要性(冷却ふくめ)。
- Transformerの専用チップでてきたが、さらなる新規アイデアが出た場合は？

□ 石橋:

- コンテナDCについて、水冷型を実証実験予定。顧客は建物をいじることを嫌う。ドイツで予定されている
- 高温の水ができる。熱再利用により「カーボンネガティブ」へ

- GPUは将来1kW, 10kW? 廃熱の利用(焼肉含む)
- 米国の大企業はやりたいことに対してすぐDCを作る。日本の体質について、応用側のプレイヤーとシーズの距離が遠いのではないか。