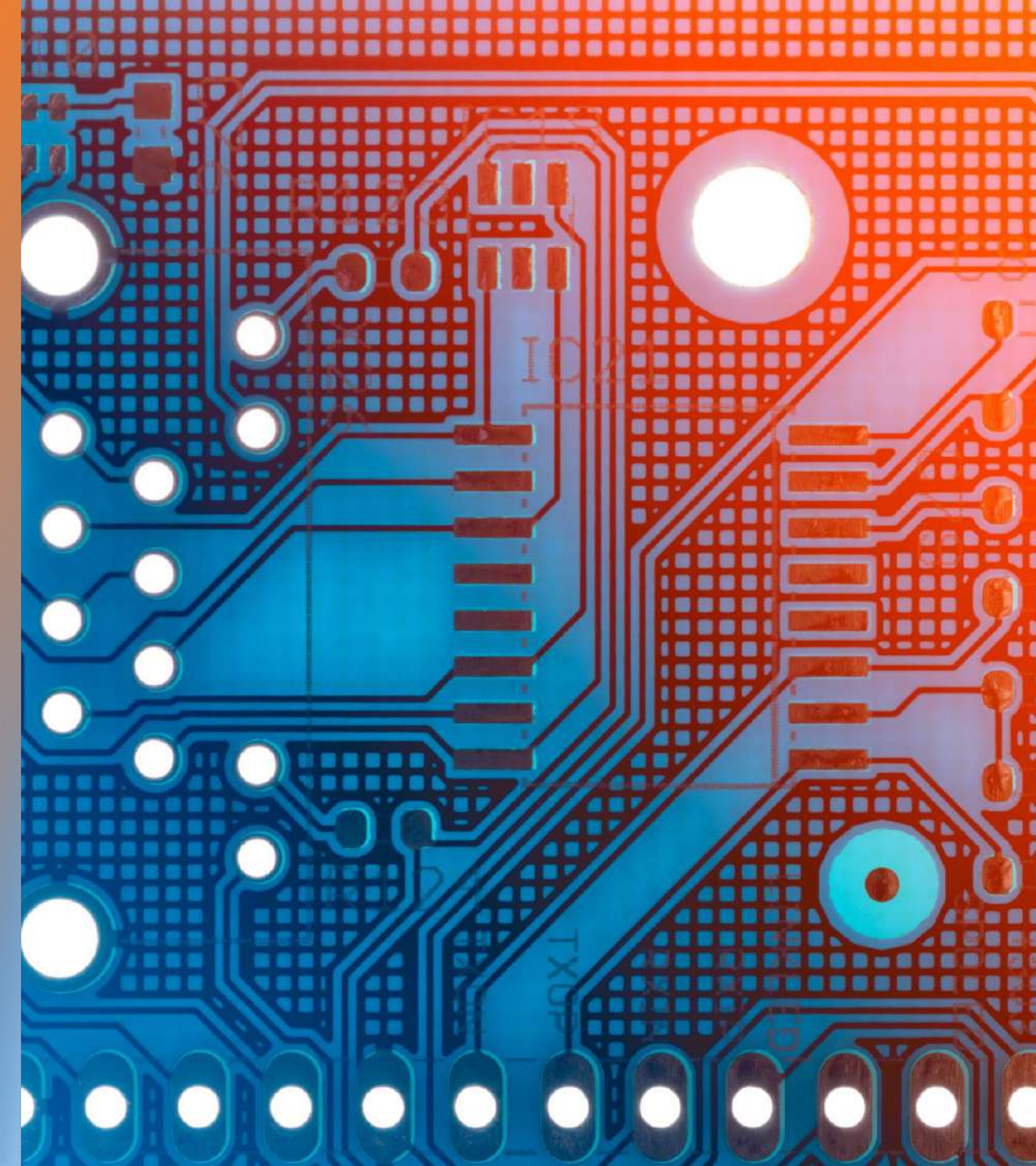




FUTURE PROOFING THE DATA CENTER

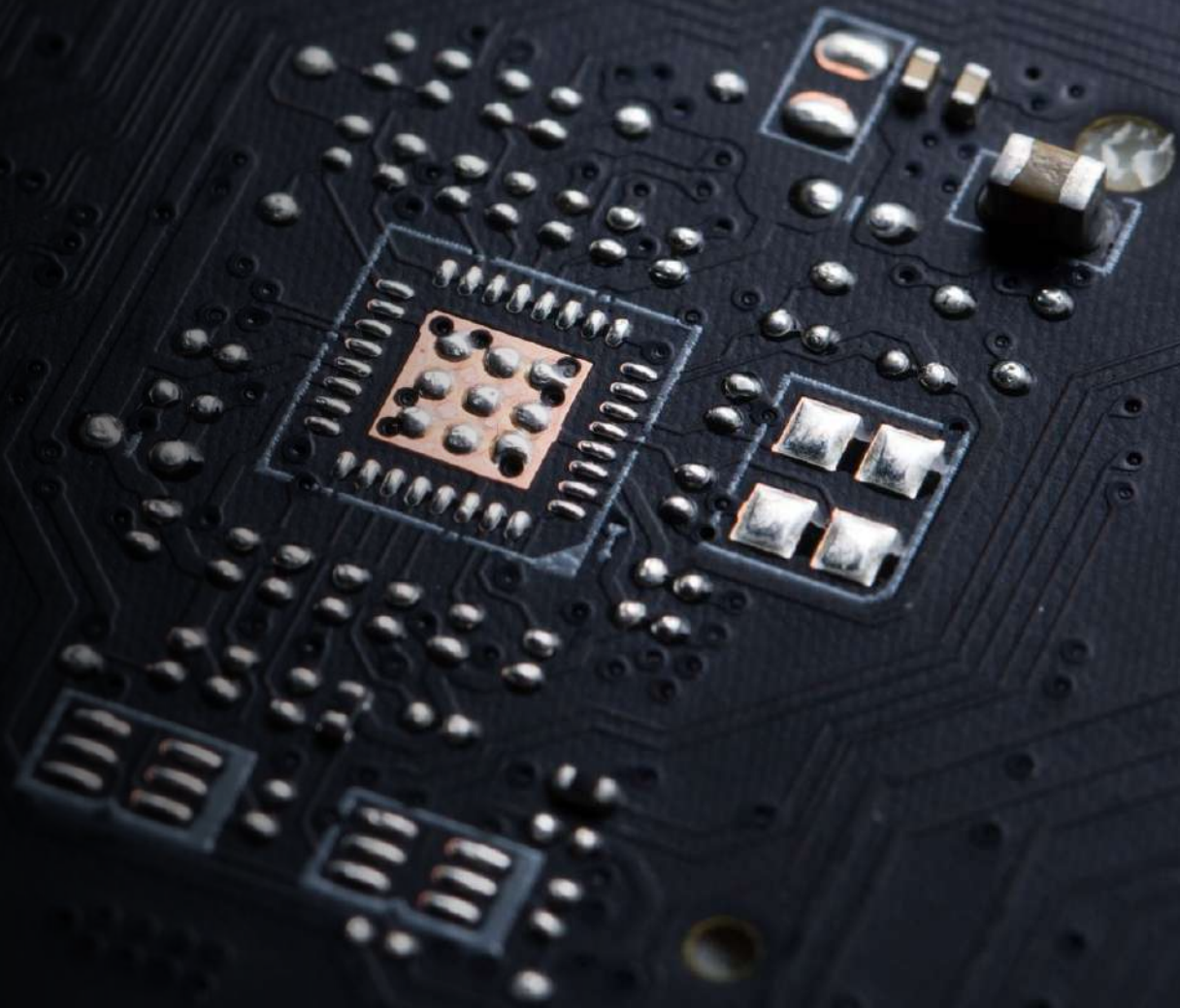
Objective statement - To instigate discussions around future proofing the data center with better designs



AGENDA

- The Chip Design
- Impact to the Data Centre Facility
- The Demand Forecast
- Not a cooling problem it's a Power Problem
- Data Centre Cooling: RDHx
- Summary

THE CHIP DESIGN



THE CHIP DESIGN

M1 delivers up to 3.5x faster CPU performance, up to 6x faster GPU performance, and up to 15x faster machine learning, all while enabling battery life up to 2x longer than previous-generation Macs. With its profound increase in performance and efficiency, M1 delivers the biggest leap ever for the Mac.1

 <p>MTIA develops its own silicon, the "Meta training and inference accelerator".</p>	 <p>Hopper Places 7nm And 5nm Orders For Next-Gen Ampere GPUs.</p>	 <p>Research Alliance Produces Industry's First 7nm Node Chips , July 9, 2015.</p>	 <p>M1 5nm node and is expected to begin production using 3nm by 2023.</p>	 <p>IDM 2.0 7nm PC Chip Will Arrive in 2023 Using TSMC's Tech.</p>
---	---	---	---	---

- 8 X SXMS GPU's at 700W power draw circa 5.6kW
- 4+2 3000W (240v) 80 PLUS Titanium power supplies,
- Single server in a 5U Chassis
- Can be accompted by a Bluefield-3 additional 16 Cores
- Network cards upwards of 2kW

G593-SD0 (rev. AAX1)

HPC Server - 4th Gen Intel® Xeon® - 5U DP SXMS H100 8-GPU Server System | Application: AI, AI Training, AI Inference & HPC



CPU vs. GPU vs. DPU

CPU	GPU	DPU
Several cores	Many cores	Dozens of cores
Low latency	High throughput	Higher degree of versatility
Ideal for serial processing	Ideal for parallel processing	Ideal for big data processing
Handles a handful of operations at once	Handles thousands of operations at once	Handles thousands of operations at scale



Network Card : Bluefield-3 additional 16 Cores

5 nanometer process

Machine learning accelerators

16-core Neural Engine

11 billion operations per second

Up to 8-core GPU

8-core CPU

16 billion transistors

Industry-leading performance per watt

Transistor / USB 4 controller

Media encoder and decoder engines

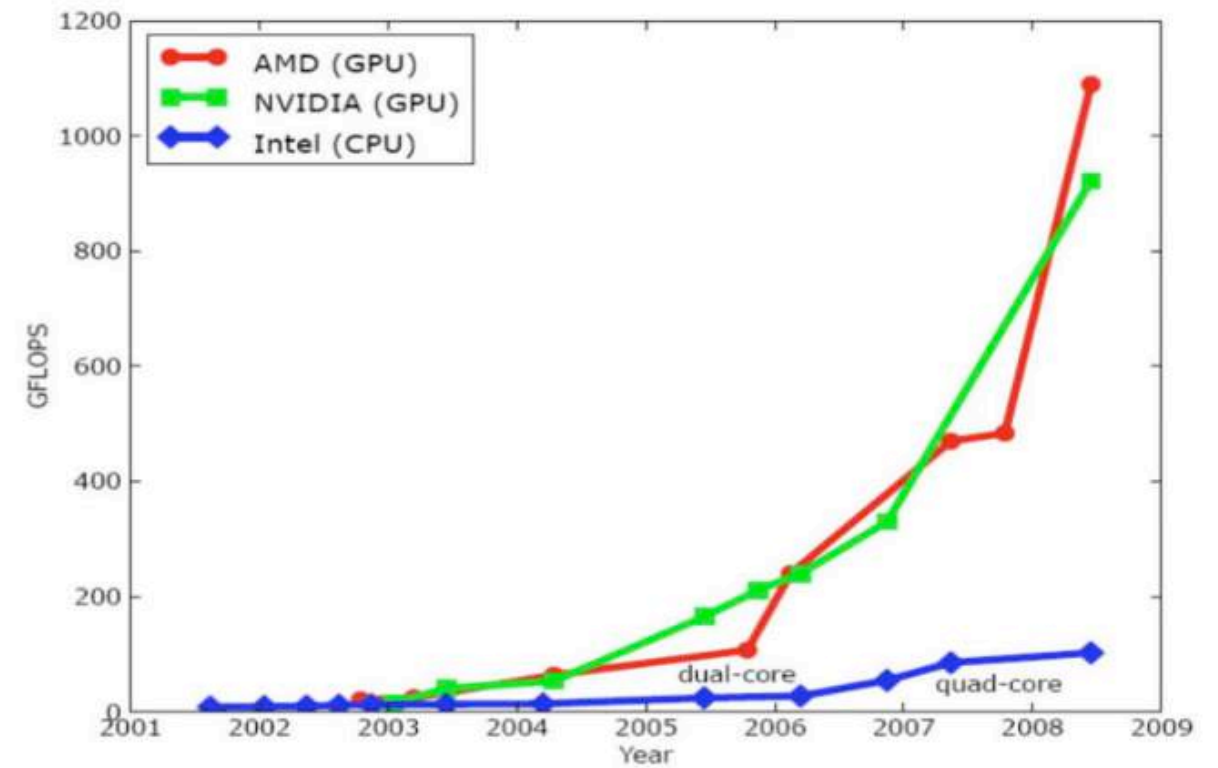
Apple M1

Advanced image signal processor

Secure Enclave

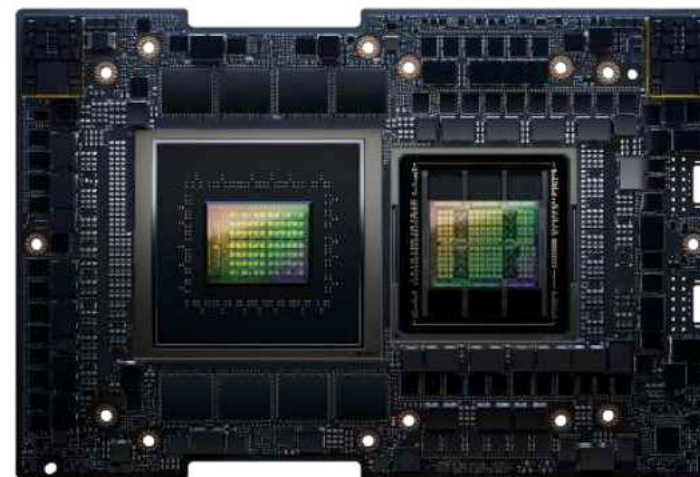
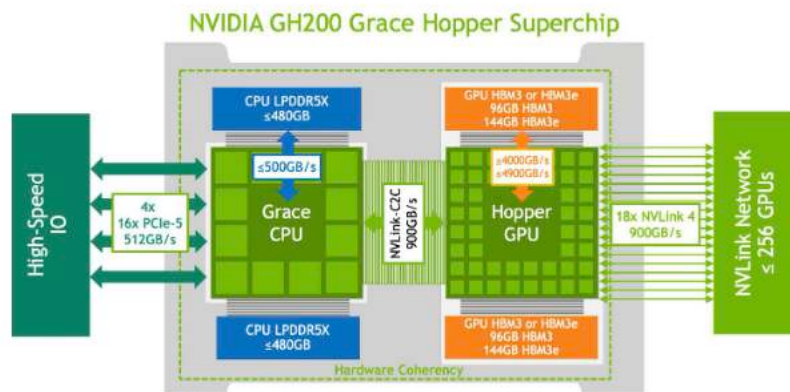
Unified memory architecture

IT Hardware Development is on a steep curve of capability



CPU (from Intel) versus GPU (from Nvidia and AMD) performance growth [Nvi10a]

nVidia Grace Hopper Super chip
1kW in Power max.
 (GPU+CPU+Memory)





IMPACT TO THE DATA CENTER

WHEN DID A DC STOP DELIVERING WHAT IT SHOULD DO?

A Data Center is supposed to provide Space, Power and Cooling for IT Hardware.

Additional Infrastructure:-

- Aisle Containment
- High Density Vent Tiles
- 3 Tile Aisles
- Hot Air Plenum Ceilings
- In Row Coolers
- Product separation



Images taken from regmedia.co.uk

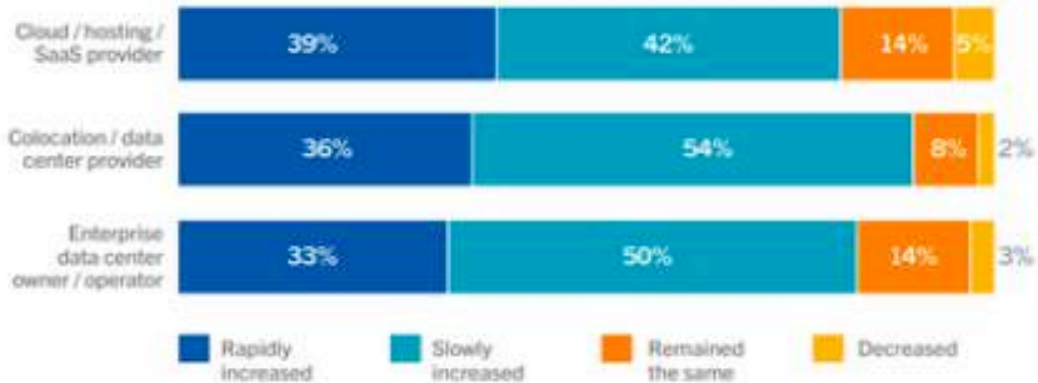
“The demand of AI has seen that these additional infrastructure components are no longer applicable to support the Data Center”

DC Design 5kW - stretched to 25kW per Rack
Maybe 30kW with Indirect cooling solutions

IMPACT TO THE DATA CENTER

Rack power density rising fast across major data center segments

Over the past three years, how has the most common (modal average) rack power density deployed in your organization / colocation data centers changed? (n=611)

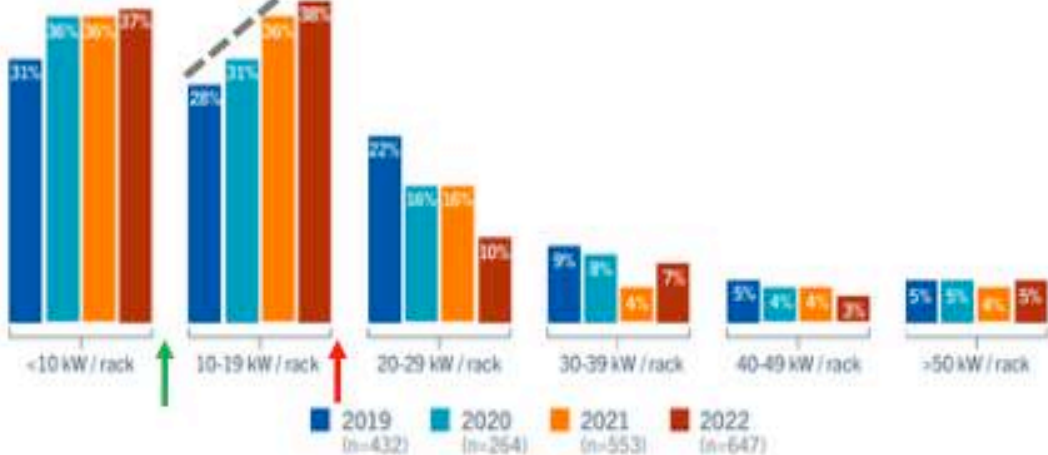


UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2022

UptimeInstitute | INTELLIGENCE

2022 - { 2026 - 2030/ '35}

What is the highest server rack density deployed in your site?



UPTIME INSTITUTE GLOBAL SURVEY OF IT AND DATA CENTER MANAGERS 2018-2022

UptimeInstitute | INTELLIGENCE

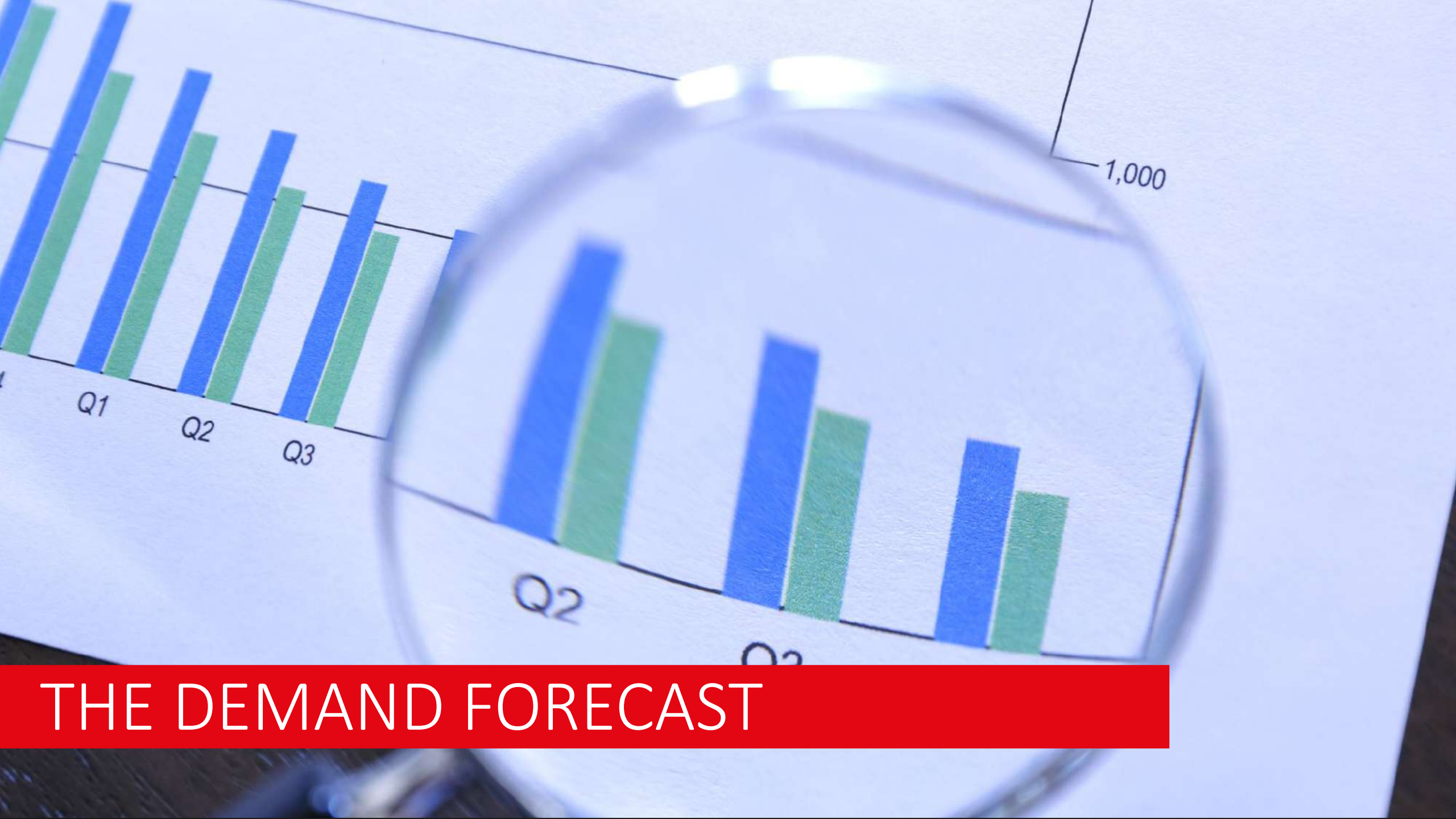
However, when it comes to the largest facilities at 10 MW and above, nearly half have **cabinets above 20 kW**, and almost one in five run some **racks over 40 kW power**. There is also a small but growing number of data centres that house **some cabinets above 70 kW power**, concentrated mostly in the largest facilities.

Source:

Uptime Institute : Global Data Center Survey 2022 _ report #78

RFP Requests in this month alone:-

- 5 Enquiries
- 564 RDHx
- Averaging 52kW per RDHx .
- Between 40-70kW capacity.



THE DEMAND FORECAST

**“IT’S ONLY GETTING WORSE - YET
DESIGNS ARE NOT CHANGING FAST
ENOUGH TO COPE WITH DEMAND!”**

ARE YOU?

The Demand Forecast

Abstract

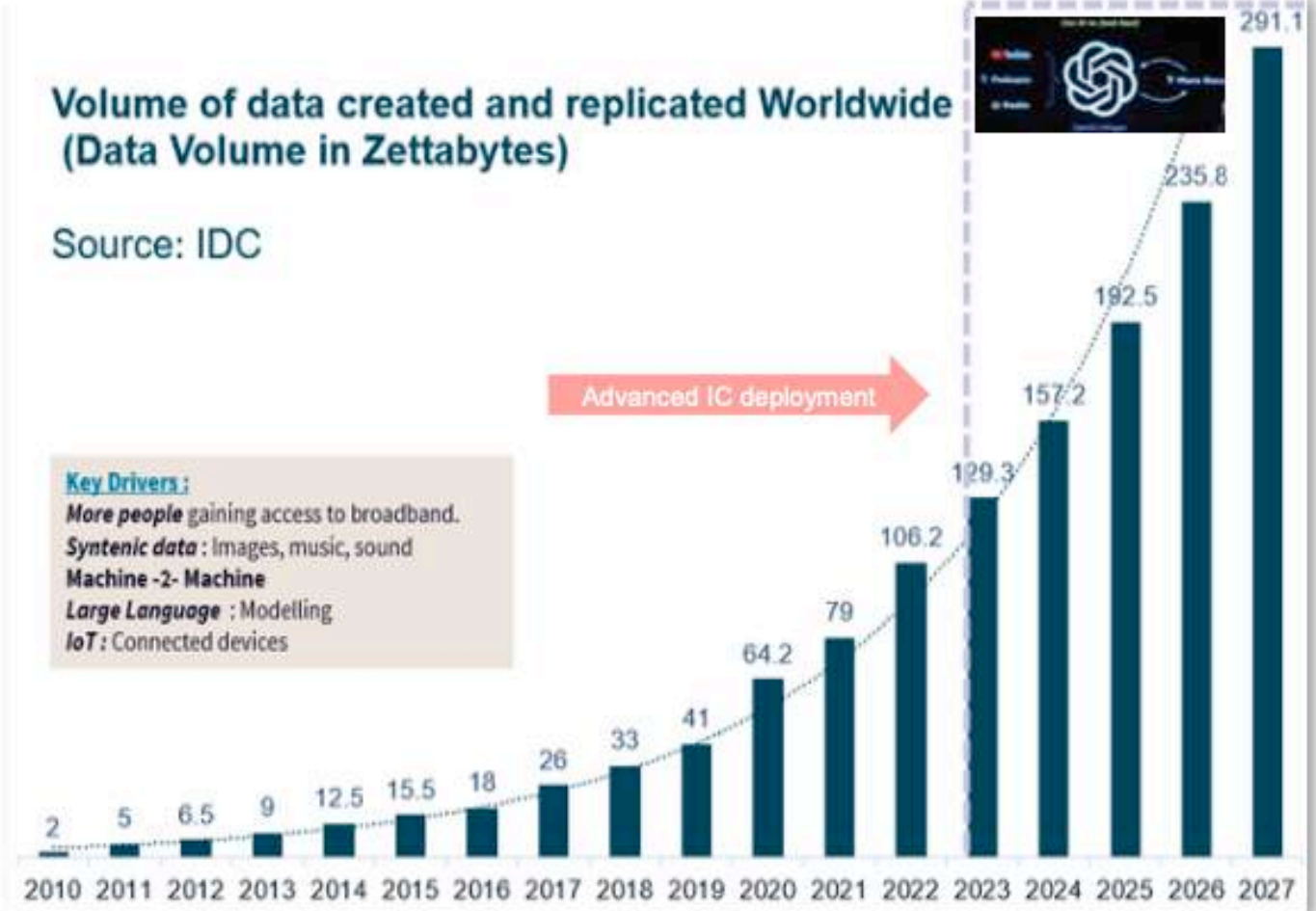
This IDC study presents a five-year forecast for the worldwide IDC Global DataSphere. The worldwide IDC Global DataSphere is a measure of how much new data is created, captured, replicated, and consumed each year. It is forecast by several segments including consumer/enterprise, region, data type, location (core, edge, endpoint), **and deployment type (cloud/noncloud) of the technology involved in data generation.**

"The Global DataSphere is expected to more than double in size from 2022 to 2026," according to John Rydning, research vice president, IDC's Global DataSphere. "The Enterprise DataSphere will grow more than twice as fast as the Consumer DataSphere over the next five years, putting even more pressure on enterprise organizations to manage and protect the world's data while creating opportunities to activate data for business and societal benefits."

Important to understand the newer disruptive technologies of the future (AL, ML, AR, IoT,,,) this workload / data profile is very much different than previous. There is an expected step-change once the latest IC devices become pervasive.

Volume of data created and replicated Worldwide (Data Volume in Zettabytes)

Source: IDC



ZB to YB : Zettabyte 10_{21} to Yottabyte 10_{24}

WHAT CAN USystems DO TO
HELP??



DATA CENTER COOLING: RDHx

REAR DOOR COOLING FUTUREPROOFING YOUR DC

- Delivering 156kW per Rack of cooling (site specific)
- Trusted tech since 2006
- No fundamental design change unlike other new cooling technologies.
- Manages the rack heat load with the added benefit of controlling the entire room temperature, without the need for supplementary cooling

Delivers more than a traditional design

- Removal of heat closer to the heat source means higher efficiency
- Warmer water temperatures (20+) = FC
- Zero water wastage
- Increases usable real estate
- Retrofittable to traditional DCs
- Heat capture and reuse
- Allows for varied external water sources



Picture Credited to ECMWF

Early adopters are already ahead of the game, using higher water temperatures gaining more free cooling and maximizing kW Rack Cooling.

These customers have already futureproofed their sites to deploy today's AI Hardware architecture.



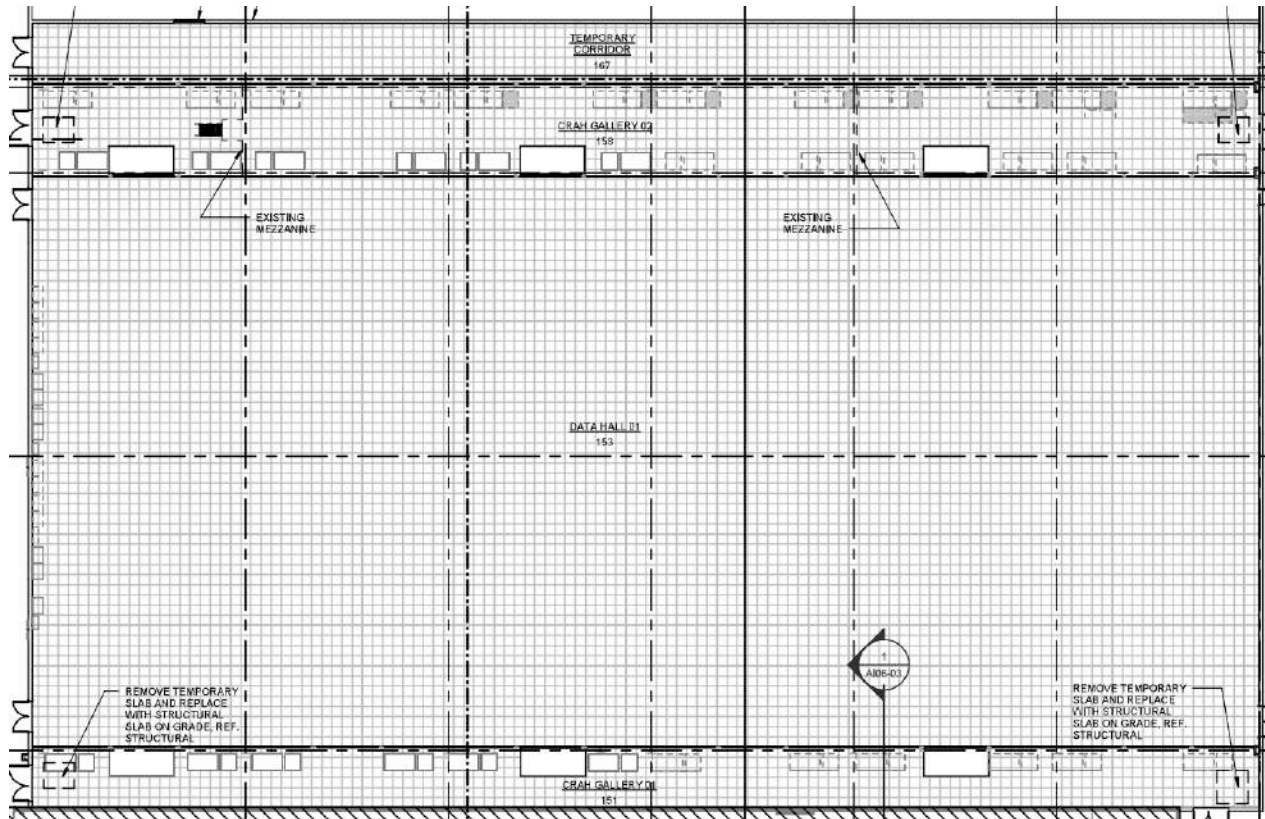


IT'S A POWER PROBLEM, NOT A COOLING ONE

IT'S A POWER PROBLEM, NOT A COOLING ONE

There has been a shift change..... Traditionally it was always a cooling limitation on rack cooling capability.

Instead of changing the design, we added more and more components to the architecture to try and solve the issues



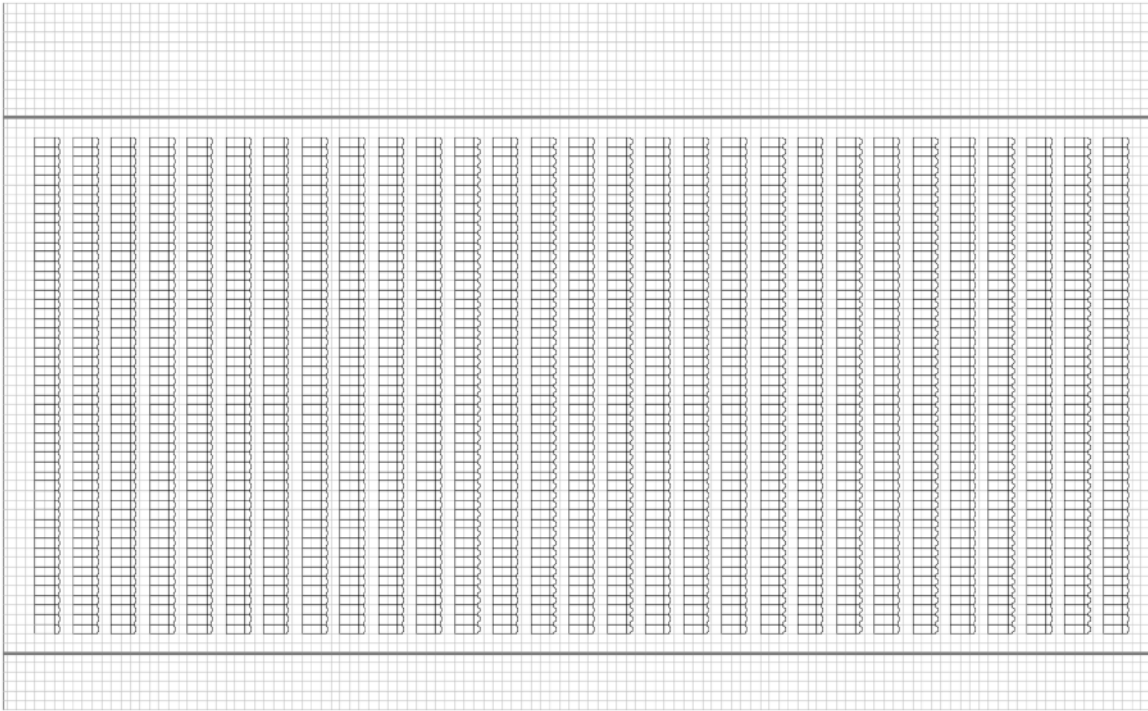
9MW Data Hall

1506 Racks

Design kW per rack = 6kW

CRAC based cooling using technical corridor

IT'S A POWER PROBLEM, NOT A COOLING PROBLEM

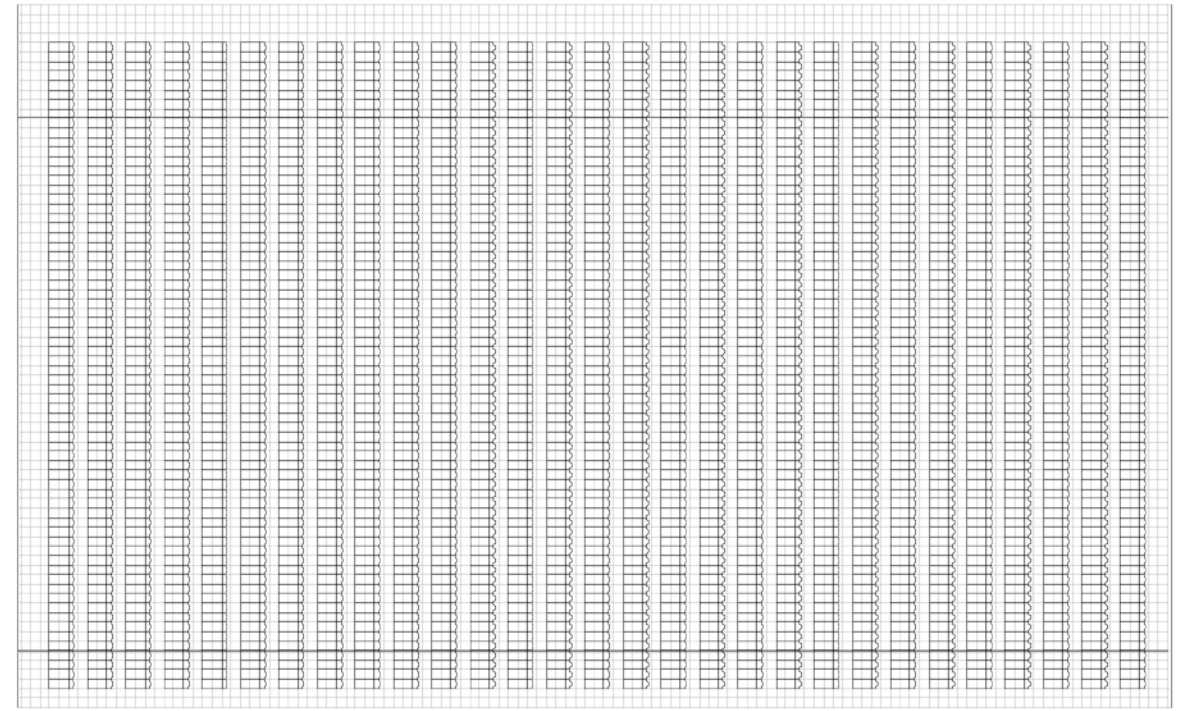


37.6MW Data Hall

1506 Racks

Design kW per Rack = 25kW

RDHx Design - Soldier formation



46.4MW Data Hall

1856 Racks

Design kW per Rack = 25kW

RDHx Design - Soldier formation

92.8MW Data Hall

1856 Racks

Design kW per Rack = 50kW

RDHx Design - Soldier formation

IT'S A POWER PROBLEM, NOT A COOLING PROBLEM SUMMARY

It's all about your DC Design.



SUMMARY

IN SUMMARY

- **IT Hardware is driving change in DC requirements**
- **Meaning overall DC design has to change as are not fit for purpose or able to meet the demands of the new hardware**
- **You need look at your Power strategy as well as your Cooling strategy to achieve success for the AI Boom!**
- **To Futureproof your facilities to cope with this demand, RDHx technologies should be adopted.**
- **It's still all about your Design!!!**