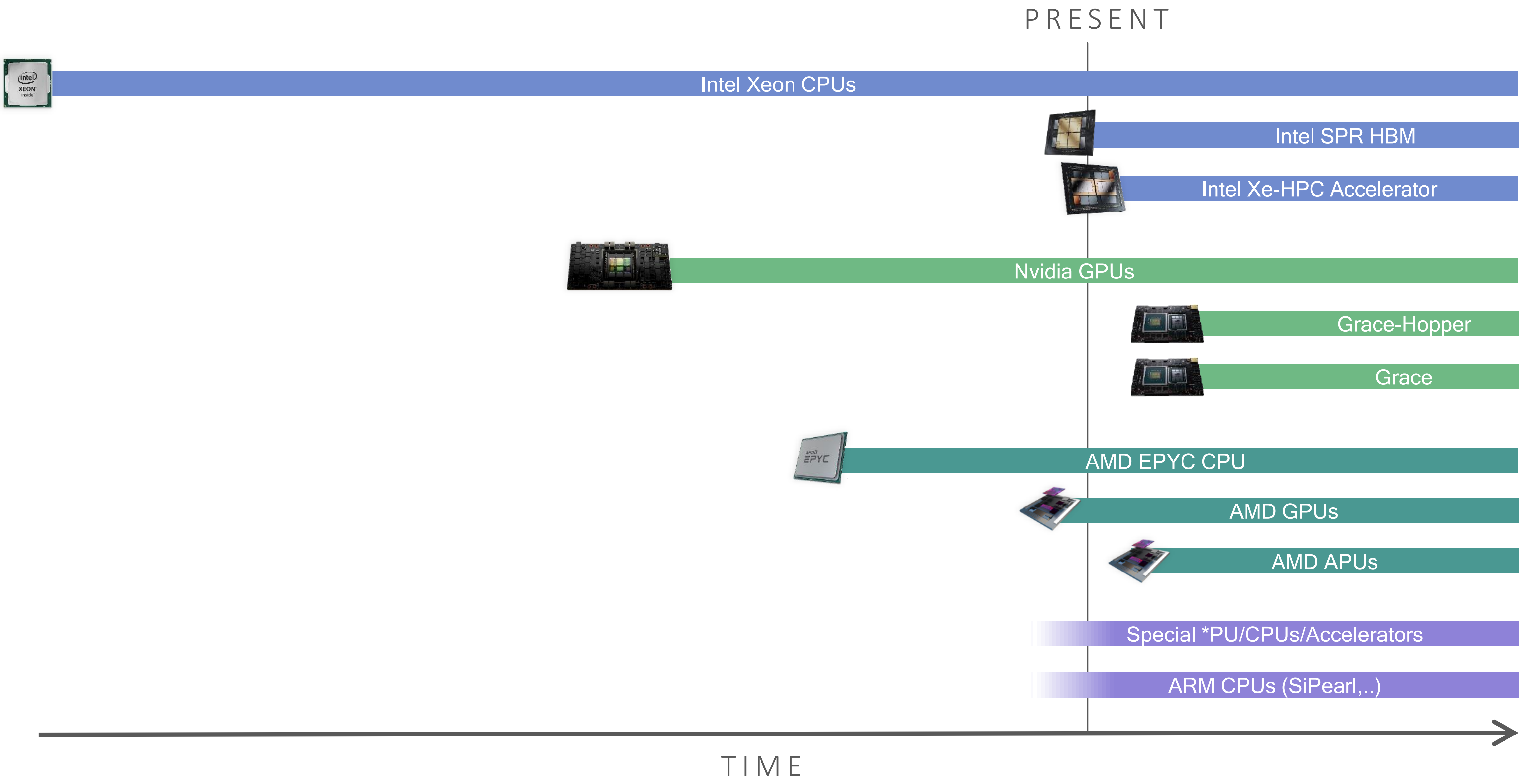
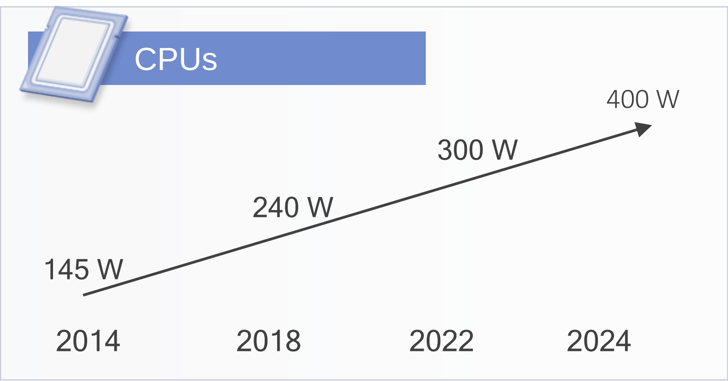
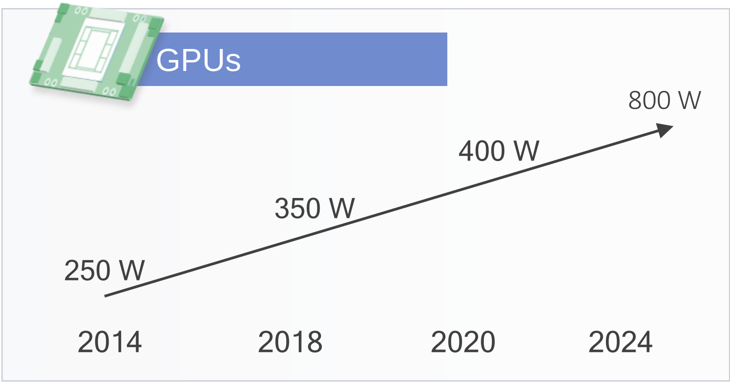
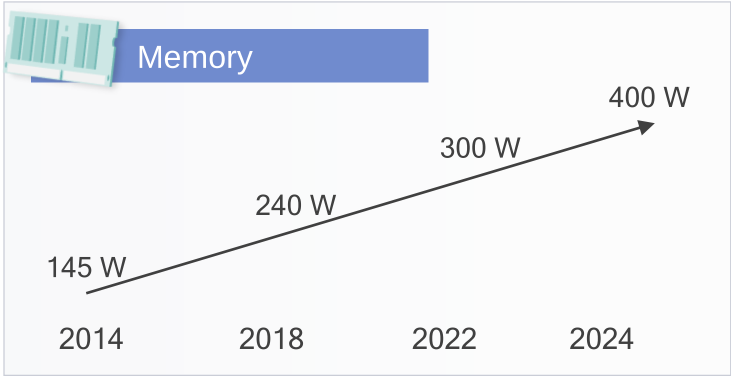
HPC Landscape - Technology Diversification

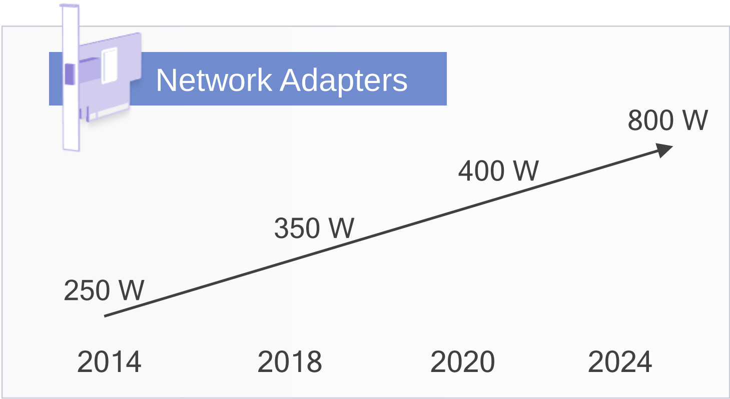


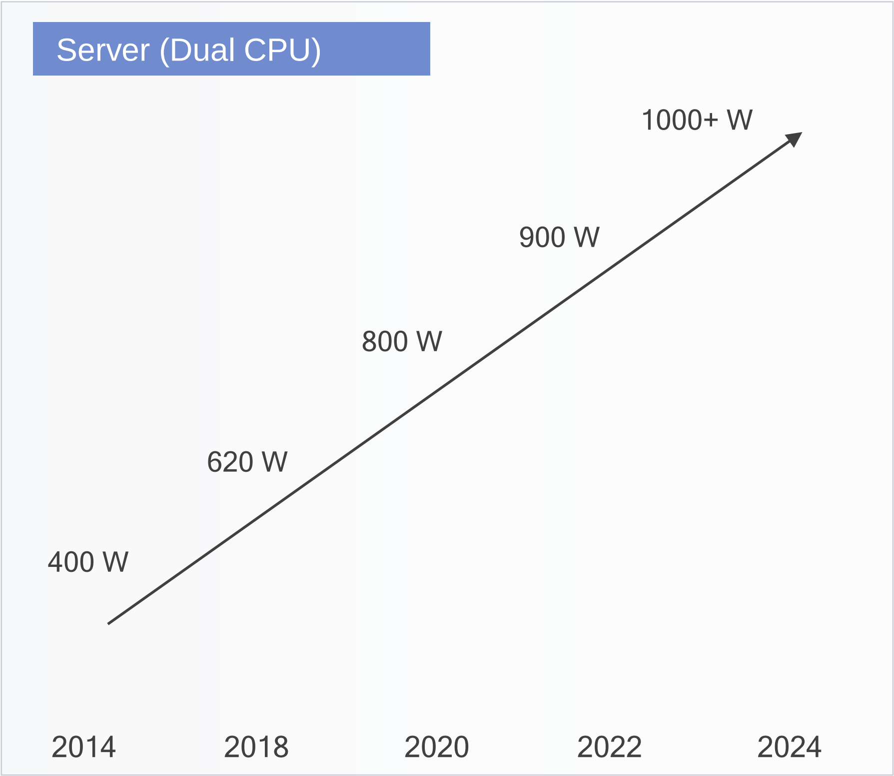
Traditional Cooling Approaches Reached Critical Limits











T-case temperature reducing for high TDP CPUs

GPU servers soon to reach ~4-5kW

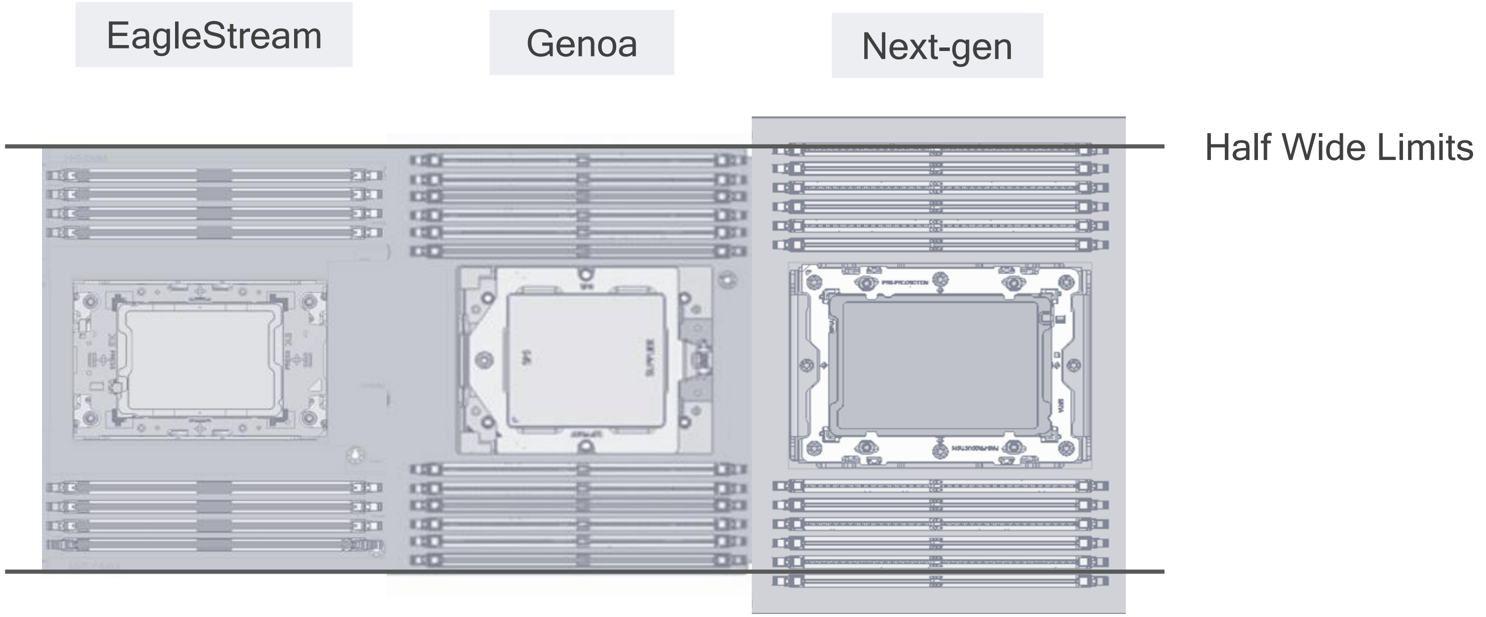
Greater CFM -> more fans, higher speeds

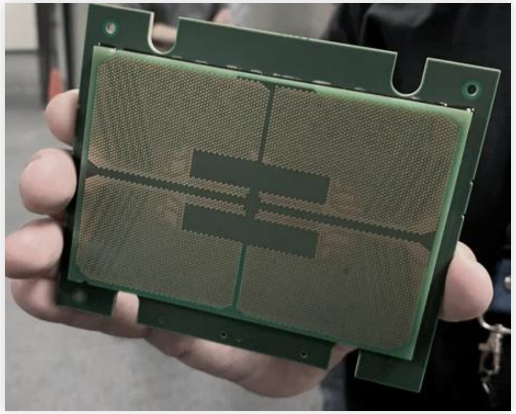
Higher performing heat sinks (needs liquid)

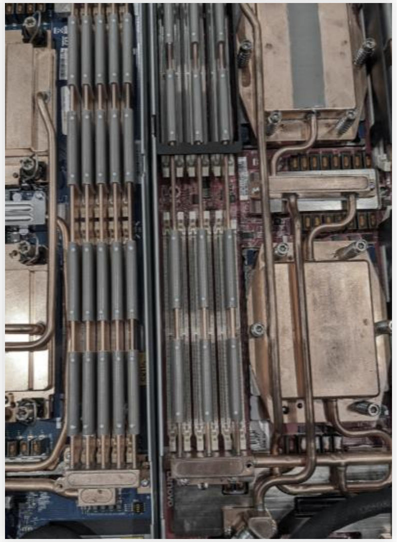
Increase in datacenter air-handling capabilities

High-capacity power supplies

Mechanical challenges with dense nodes



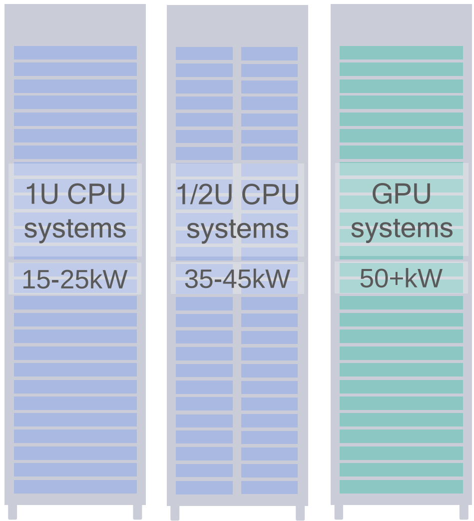




The number of channels per socket is increasing.
Expecting 12 per socket at minimum

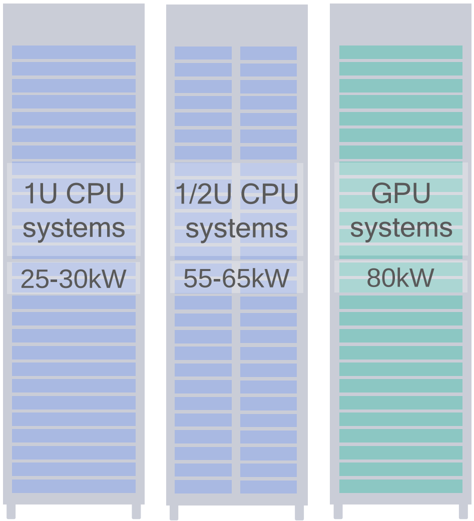
DDR5 larger DIMMs due to die-stacking

HPC Computing Characteristics Timeline – High Level



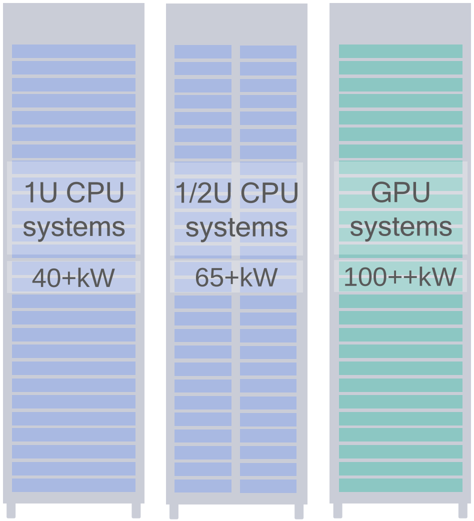
Max 205W air-cooled dense
Dense Half-wides "standard" for HPC
72 nodes per rack
4 x 32amp drops/rack
High heat capture using 50C water

<2021



260W -> 400W TDP
EOL for high-end air-cooled dense
1U/2U re-emerging
Rack power approaching 100kW
Exascale-level node building blocks
High heat capture, increased flow rates

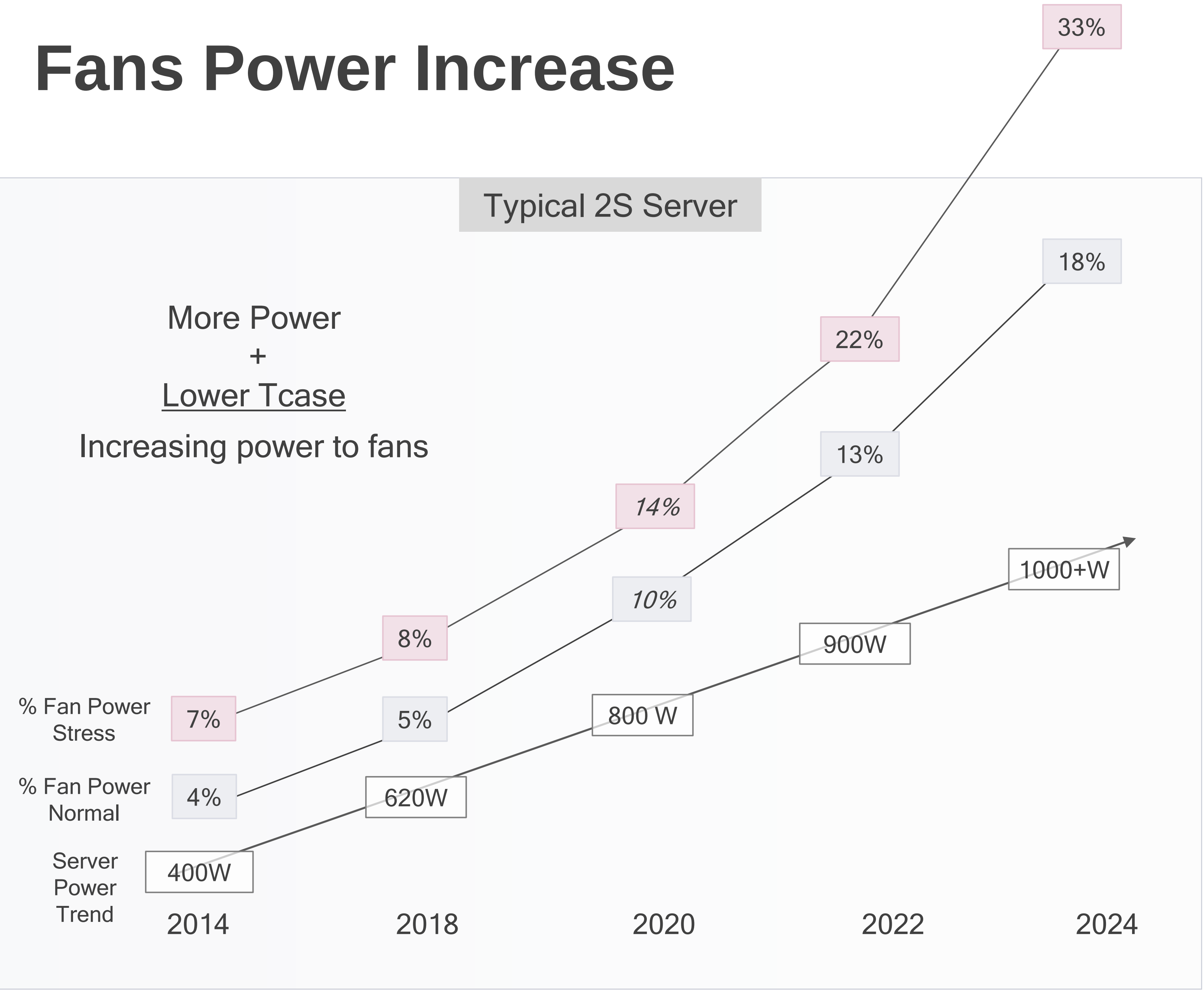
2022-23



>400W TDP CPUs,>800W GPUs
x, y and z form-factor dimension increase
Increased flow rates, lower water temps
> 6 power drops on 32A

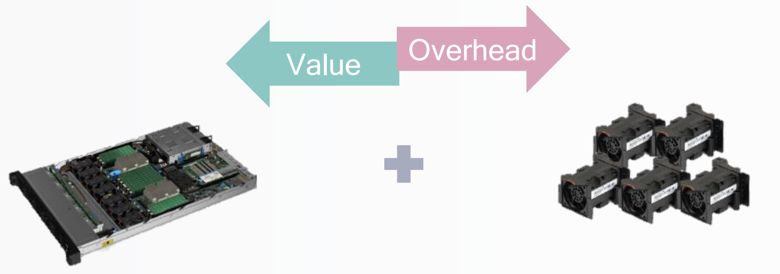
>2024

Fans Power Increase



Typical 700W Server

Server Components **Server Fans**



Component	Power
Procs (2 x 205W	400W
64GB DIMMs (16)	130W
Networking	40W
Mobo (VR, BMC, etc)	70W
Computing Value	**640W**

Component	Power
Fans / Overhead	**60W**

+10%
for cooling

Lenovo

The journey to zero emission computing

‘The Status Quo’

Carbon Positive

More CO_2 released than removed

Vast majority of data centers are here today

GOAL: Reduce as much as we can

‘We Can Do This’

Carbon Neutral

CO_2 released = CO_2 removed

We can drive dramatic reduction with existing technologies

GOAL: Use savings from efficiency gains to justify move to green power

‘The End Game’

Carbon Negative

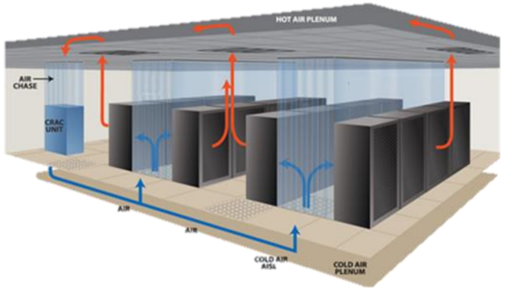
More CO_2 removed than released

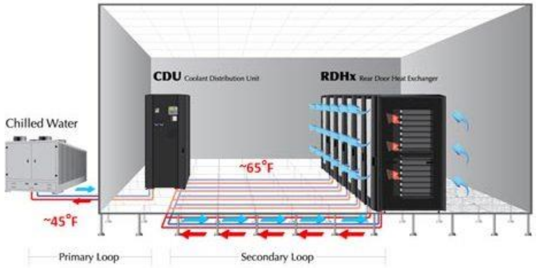
The end game we will deliver with technologies available today

GOAL: Recycle the heat from computing and use it to replace other CO_2 emitting power uses

The journey to zero emission computing

‘Classical’ air-cooled datacenter	Cold water Rear Door Heat Exchanger	Warm water Direct Liquid Cooling	Warm water Direct Liquid Cooling + heat reuse





PUE 1.4+

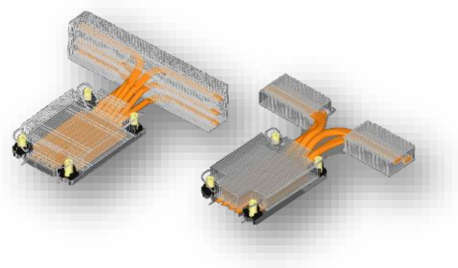
PUE 1.3

PUE >1.1

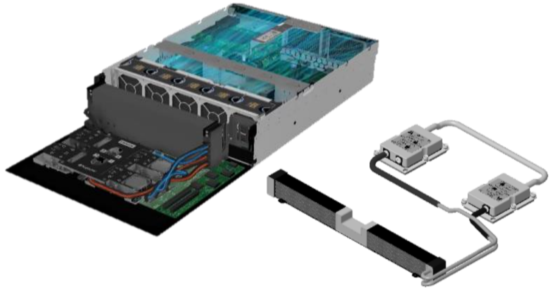
PUE >1.1
+heat reuse -> carbon negative

Liquid Assist Cooling

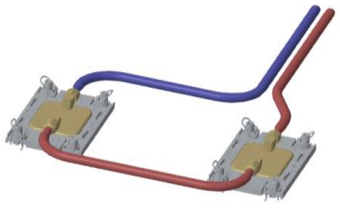
Thermal Transfer Module

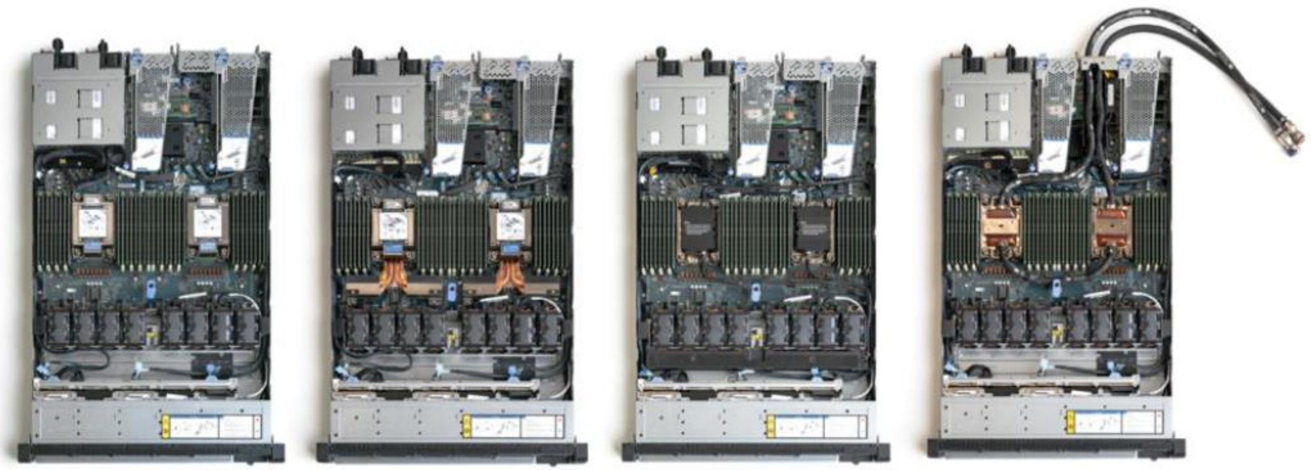


Liquid to Air Module



CPU direct liquid cooling





Support for high-end CPUs

Support for higher env temperature

Reduce fan speed and power

Eliminate many config restrictions coming from thermal limitations

Direct Water Cooling

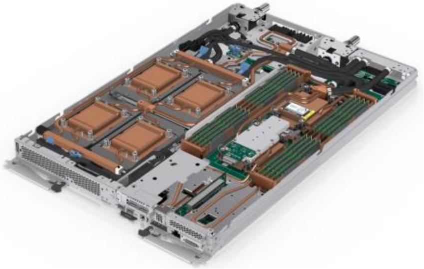
Dual Intel CPU

SD650v3

Dual AMD CPU

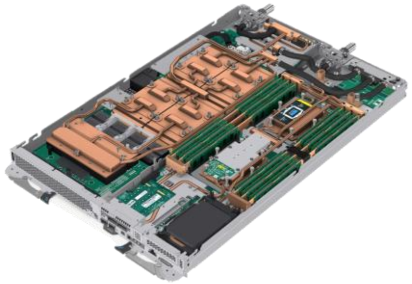
SD665v3

Dual Intel CPU
4 x PonteVecchio



SD650v3-I

Single/Dual AMD CPU
4 x H100 SXM 700W



SD665v3-N

6U enclosure
6 trays
12 CPU nodes
6 GPU nodes

Air-cooled
PSUs

DW612S

Direct-Liquid
cooled PSUs

Standard 19'' rack

Up to 50c inlet water

High water temp enables 'free cooling'

Up to ~98% heat removal to water

Heat reuse with high return water temp

Drip-less quick connects for ease of service and node removal

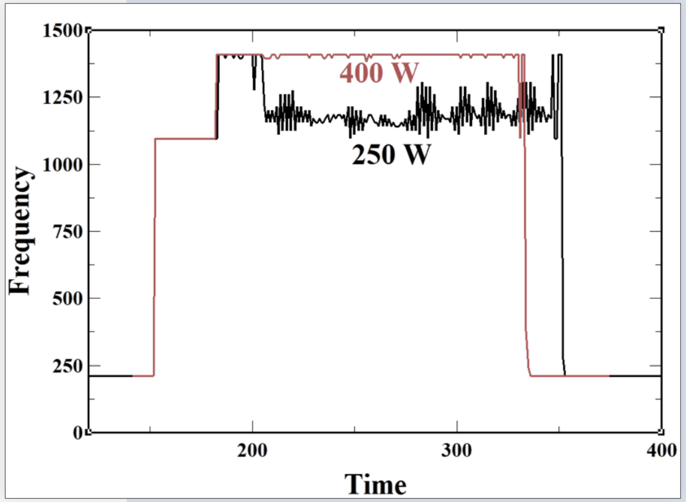
Performance Benefits of Liquid Cooling

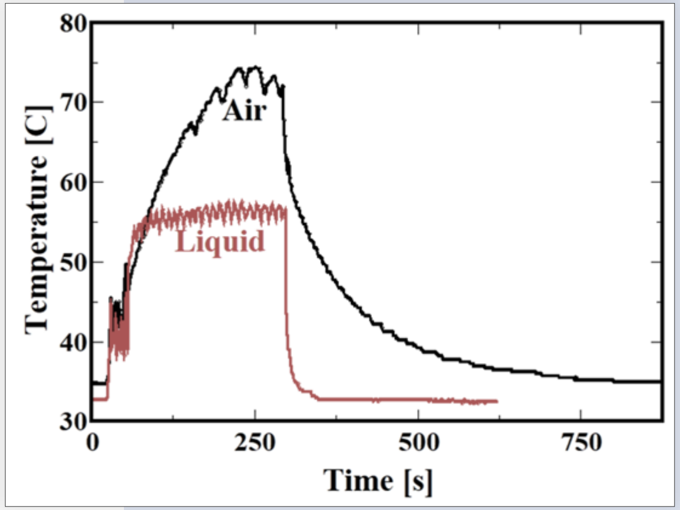
Highest absolute performance

Reliable clock speeds

Stay well below throttling temperatures

No fan power required



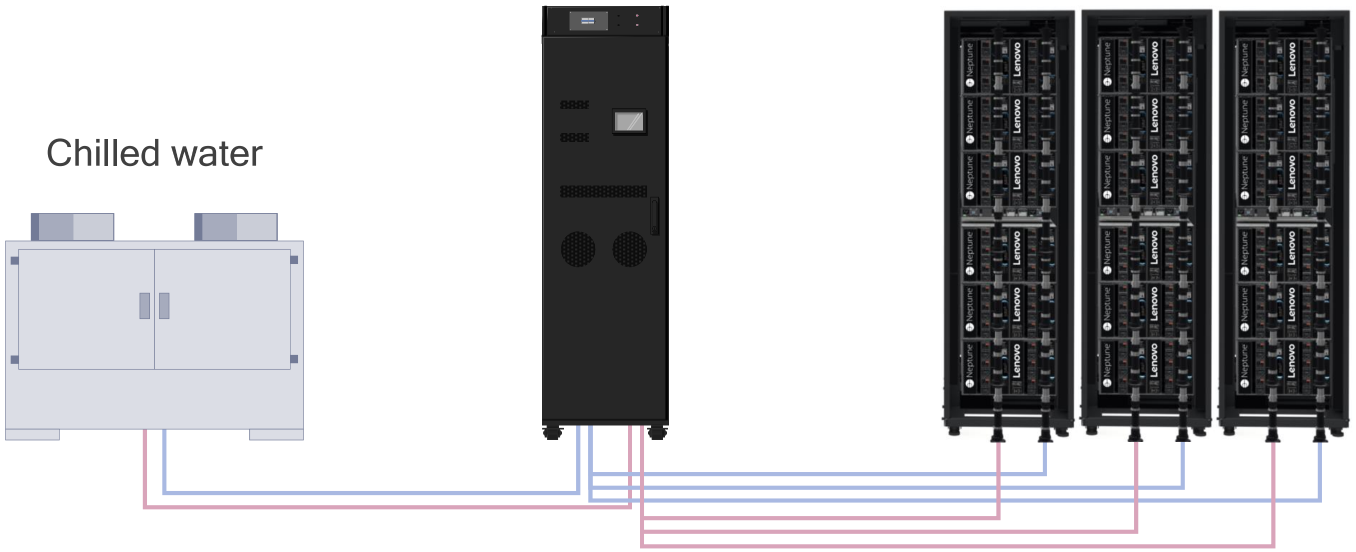


Direct Water Cooling is not complicated

CDU
Coolant Distribution Unit

Compute racks

Chilled water

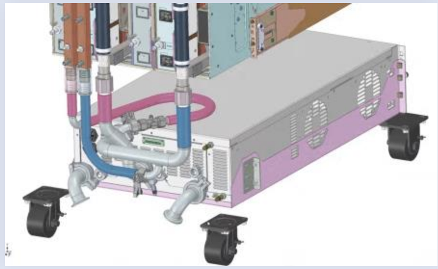


Primary loop
Building water at any temp

Secondary loop
>dew point

In-row CDUs
(450kW,1.35MW)

in-rack CDU (~100kW)



Direct Water Cooling TCO

DWC purchasing cost is higher (well, depends..), but TCO isn't

100kW example

20% for fans -> 20kW overhead

PUE

1.5 for air-cooled data center -> 50kW overhead

Price for three years

Overall **$670k**

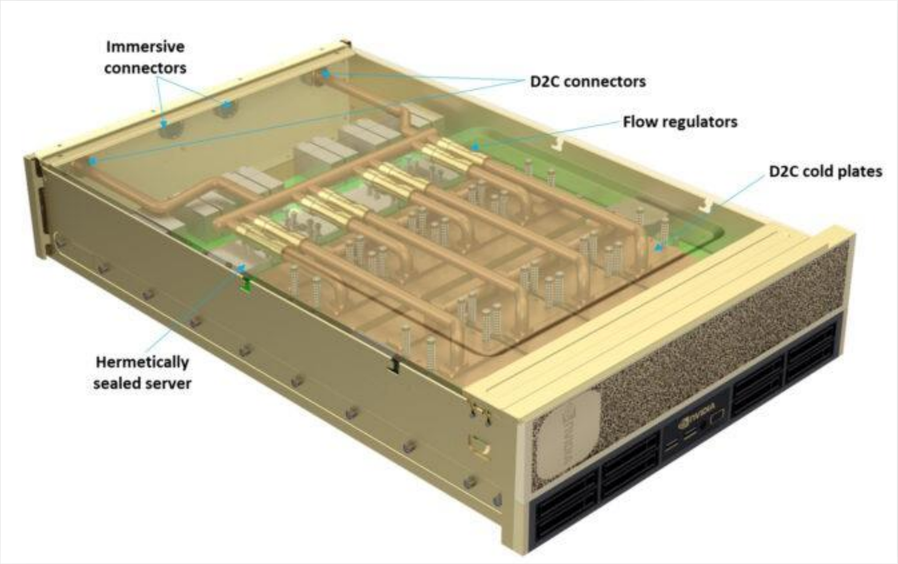
Overhead **$312k**

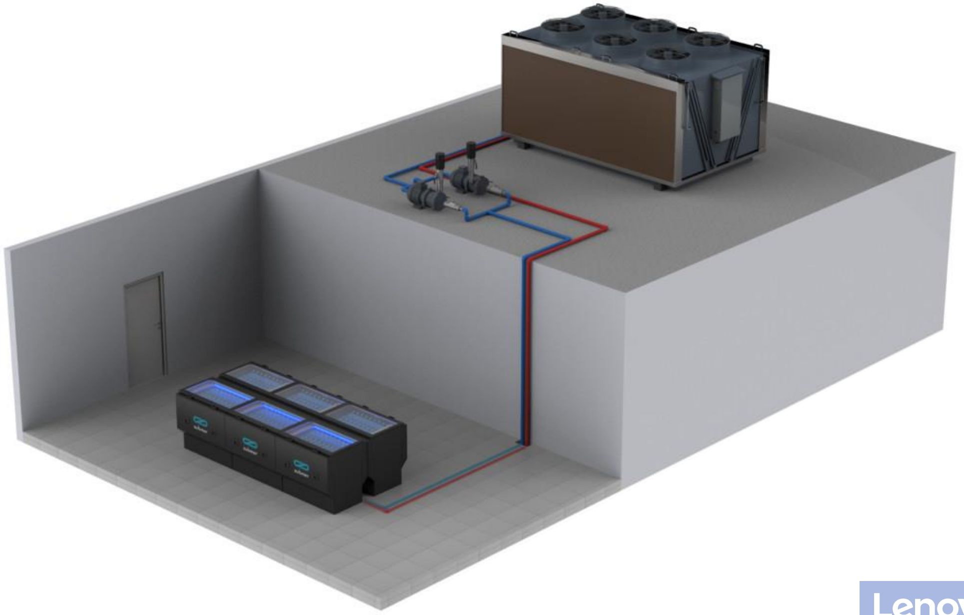
$0.17 per kWh, based on 0.6 NIS for kWh industrial use. Actual numbers might be different

Immersion Cooling in the Datacenter

Not efficient with the dense and high TDP devices even for current generation of GPUs/APUs/XPUs

Nvidia’s recent suggestion looks interesting but far from production yet. Economics not clear

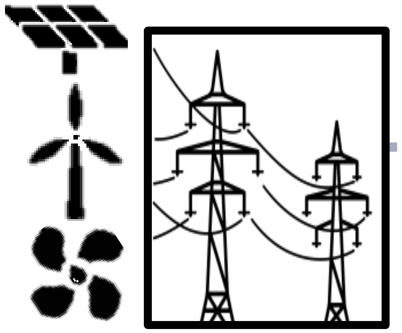


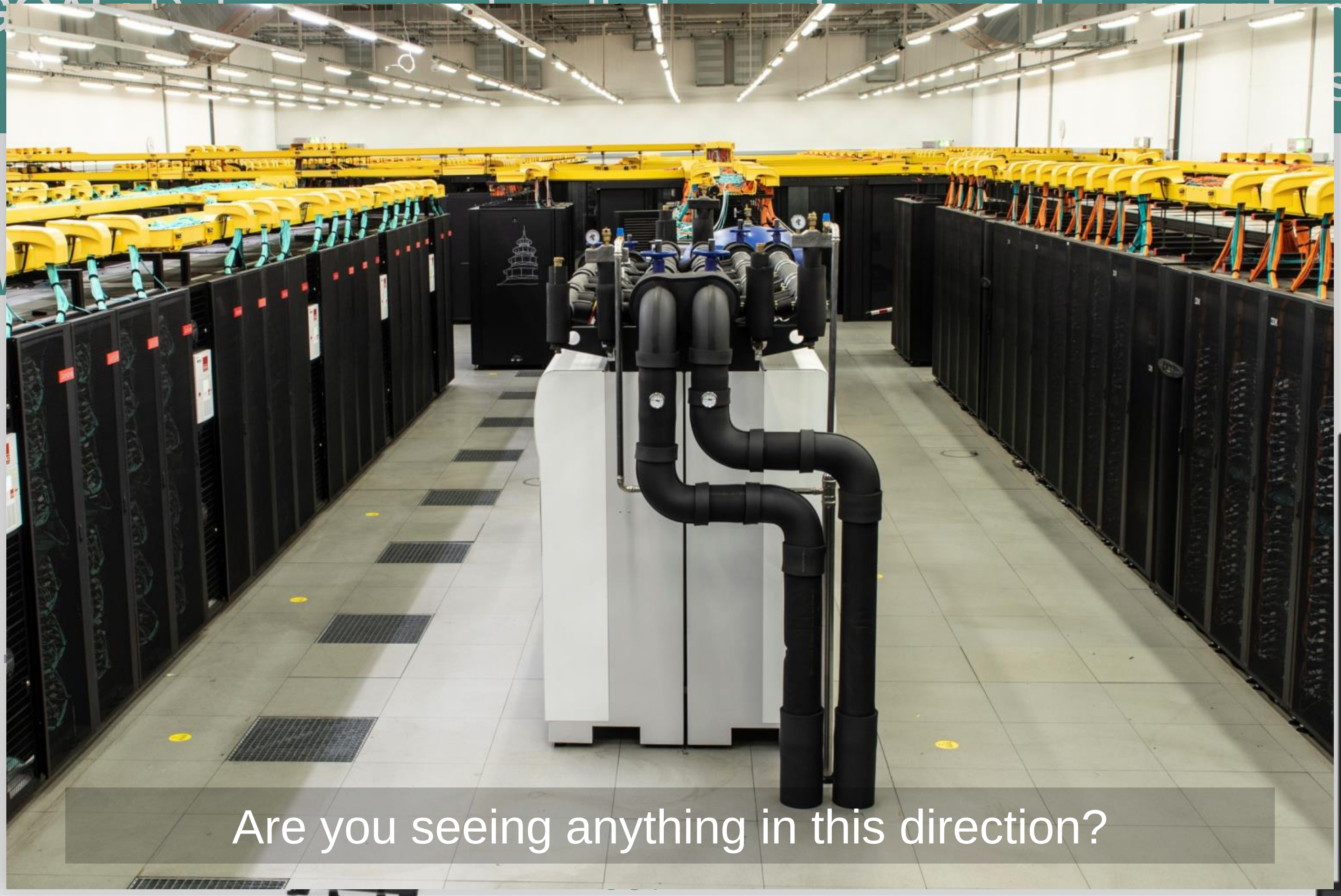


Beyond carbon neutral to carbon negative

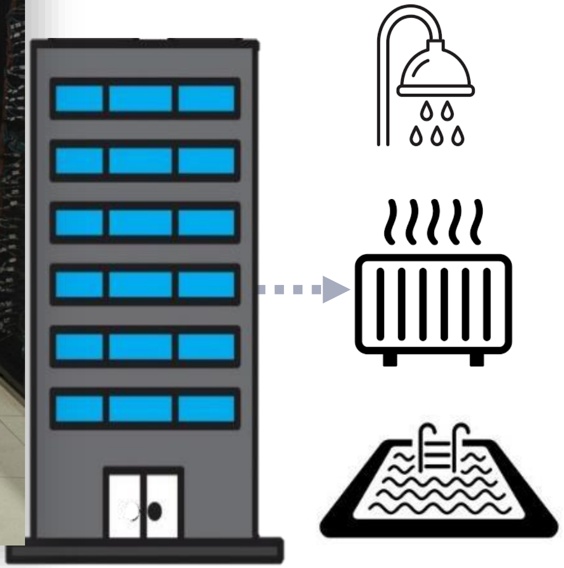
G … nally
se

No/Low Emission Pov
Now more affordable than before





Facility Reuse:
Hot Water, Heat, Pool



Discussion points, if you can share some info

How important is node density?

Node density? Power density? Performance density?

Dense nodes still the preffered choice or limitation force the move to 1U/2U?

Are you seeing configuration compromise due tp power and cooling?

Heat reuse being discussed?

New datacenter power density requirement is a challange?

Move to 63A? Number of connections per rack?

TCO is an important factor? Or just purchasing cost?

New datacenters planned?

With warm water cooling? free Cooling?

Lenovo