Hewlett Packard Enterprise

HPE 新世代處理器及技術的高性能計算 / 人工智慧解決方案

慧與科技 HPE

蔡德攸, HPC / AI APAC

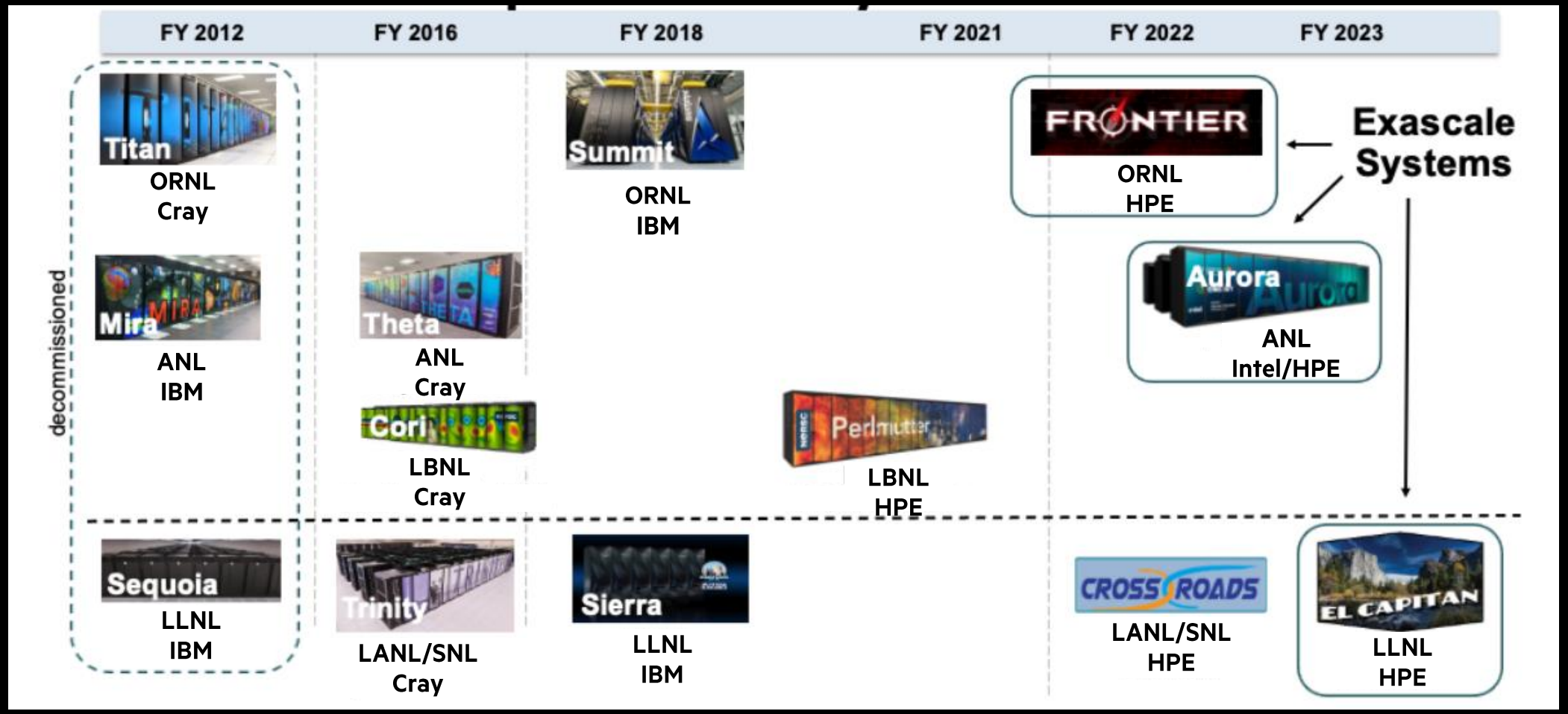
AGENDA

Intel's EXASCALE Systems

HPE HPC & AI Compute Portfolio

HPE Cooling Technologies

DOE HPC ROADMAP TO EXASCALE SYSTEMS



Report on the Oak Ridge National Laboratory's Frontier System - https://icl.utk.edu/files/publications/2022/icl-utk-1570-2022.pdf

INTEL'S PRE-EXASCALE SYSTEM – CROSSROADS (LANL/SNL)

Key Specifications

Peak Performance

Not disclosed, 165 PF peak estimated

Pure CPU System

Sapphire Rapids

All-to-All Connectivity with Node

HPE Slingshot - Low Latency, High Bandwidth

UNPARALLELED I/O SCALABILITY ACROSS NODES

2 fabric endpoints per node

INTEL'S FIRST EXA FLOPS SYSTEM - AURORA (ANL)

Key Specifications

Peak Performance

> 2 EXA FLOPS DP

Unified Memory Architecture

Across CPU and GPU

All-to-All Connectivity with Node

HPE Slingshot - Low Latency, High Bandwidth

UNPARALLELED I/O SCALABILITY ACROSS NODES

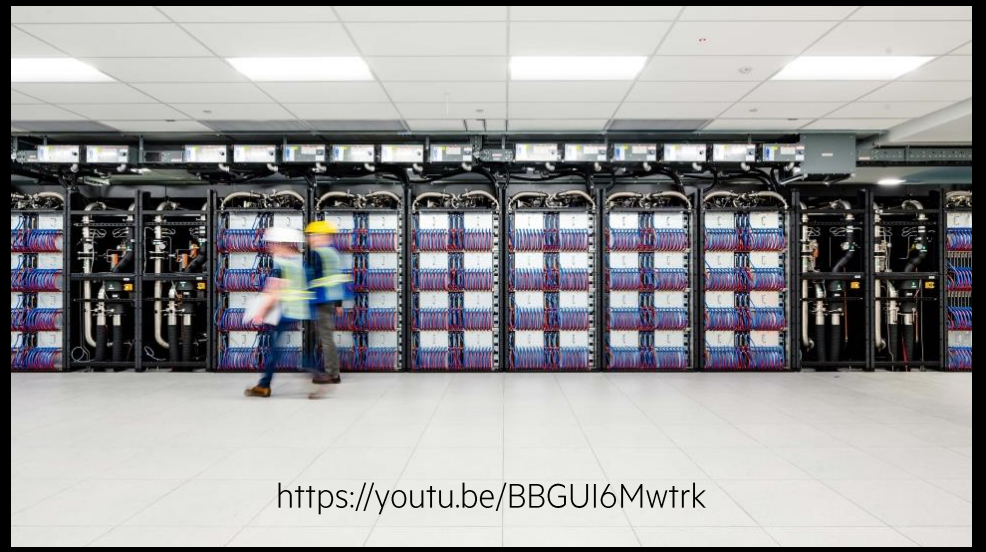
8 fabric endpoints per node, DAOS

INTEL XEON SCALABLE PROCESSOR

> 20,000 Sapphire Rapids CPUs

Xe ARCHITECTURE-BASED GPU

> 60,000 Ponte Vecchio GPUs



https://youtu.be/BBGUI6Mwtrk

AGENDA

Intel's EXASCALE Systems

HPE HPC & AI Compute Portfolio

HPE Cooling Technologies

HPE HPC & AI COMPUTE PORTFOLIO

Purpose-built supercomputing

HPE Cray EX supercomputers

HPE Cray EX2500 supercomputer

HPE Cray SC supercomputers

The next frontier of supercomputing systems redesigned for HPC, AI, and converged workloads

HPE Superdome Flex

HPE Superdome Flex 280

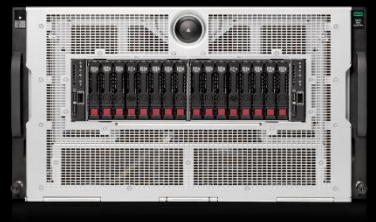
Scale-up, shared memory HPC

Accelerated HPC/AI

Accelerated compute platform for HPC and AI workloads

HPE Cray XD6500

HPE Apollo 6500 Gen10 Plus System



Mainstream HPC/AI

Density-optimized, scale-out compute for HPC and AI workloads

HPE Cray XD2000 System

HPE Apollo 2000 Gen10 Plus System

HPE ProLiant DL38x Gen10 Plus/Gen11 System

Edgeline

Range of platforms and capabilities engineered for the harsh edge environment

HPE Edgeline EL8000

Cray ClusterStor E1000

Parallel storage system purpose-engineered for HPC and AI

Integrated HPC & AI Software Portfolio including software development and acceleration environments, cluster management, data management, and system management

Everything available "as a Service" via HPE GreenLake

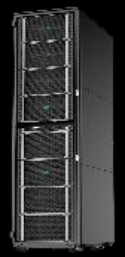
HPE PURPOSE BUILT HPC/AI SYSTEMS

Accelerated compute platform for HPC and AI workloads

HPE Apollo 2000 Gen10 Plus System

Density-optimized, scale out compute for HPC and AI workloads

HPE Superdome Flex



Near limitless scale-up compute/memory for HPC and AI

HPE Cray EX2500 Liquid-Cooled Supercomputer

Leadership Class technology in a smaller building block

HPE Cray SC Supercomputer

HPE Cray EX4000 Liquid-Cooled Supercomputer

Liquid-Cooled Converged HPC-AI

AI

HPC

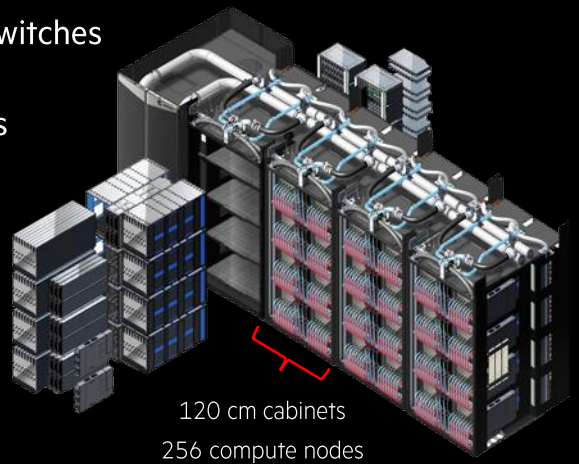
Scale

HPE CRAY EX AND SC SUPERCOMPUTER

Choice of enclosures for optimal density, efficiency, and cost per individual requirements

HPE Cray EX liquid-cooled optimized cabinet

- Up to 64 compute blades, and 512 processors per rack
- Flexible bladed architecture supports multiple generations of CPUs, GPUs, and interconnect
- Cableless interconnect between switches and nodes inside chassis
- 100% direct liquid-cooling enables 400 KW capability per cabinet
- Scales to hundreds of cabinets



120 cm cabinets
256 compute nodes
Up to 400 kW and 4 tons

HPE Cray Standard 19" Rack System

- Ready for any data center
- Full HPE Cray software and HPE Slingshot support
- Compute nodes based on HPE Apollo 2000 Gen10 Plus or Apollo 6500 Gen10 Plus
- Typical Apollo 2000 configuration is 4 HPE Slingshot switches and 64 compute nodes per rack
- Scale from one to hundreds of cabinets
- Can be integrated with HPE Cray EX supercomputers to meet customer system requirements for support of industry standard third-party equipment

60 or 80 cm racks
64 compute nodes
~65 kW

Same interconnect—same software environment

AURORA COMPUTE BLADE

2x Intel Xeon Scalable Processors

Sapphire Rapids

6x X^E Architecture Based GPU

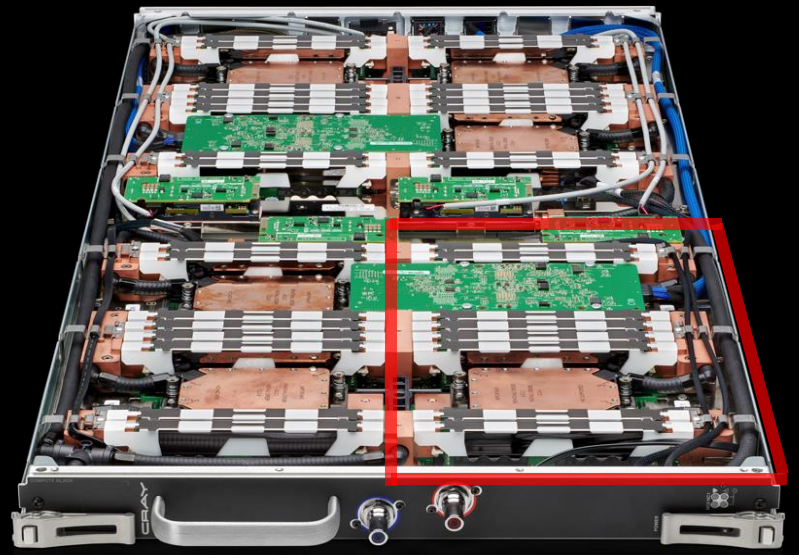
Ponte Vecchio

oneAPI

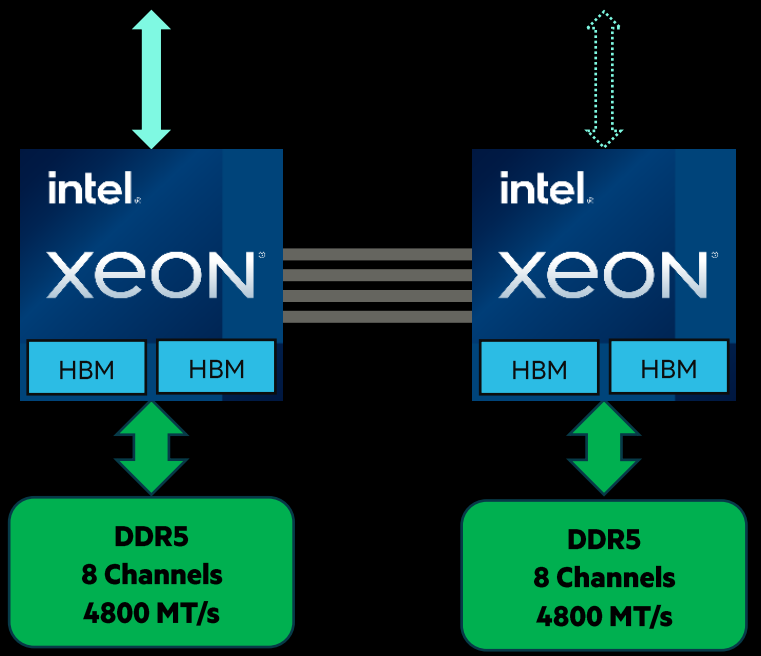
Unified Programming Model

HPE CRAY EX INTEL SAPPHIRE RAPIDS COMPUTE NODES

Cray EX420



To Slingshot 11



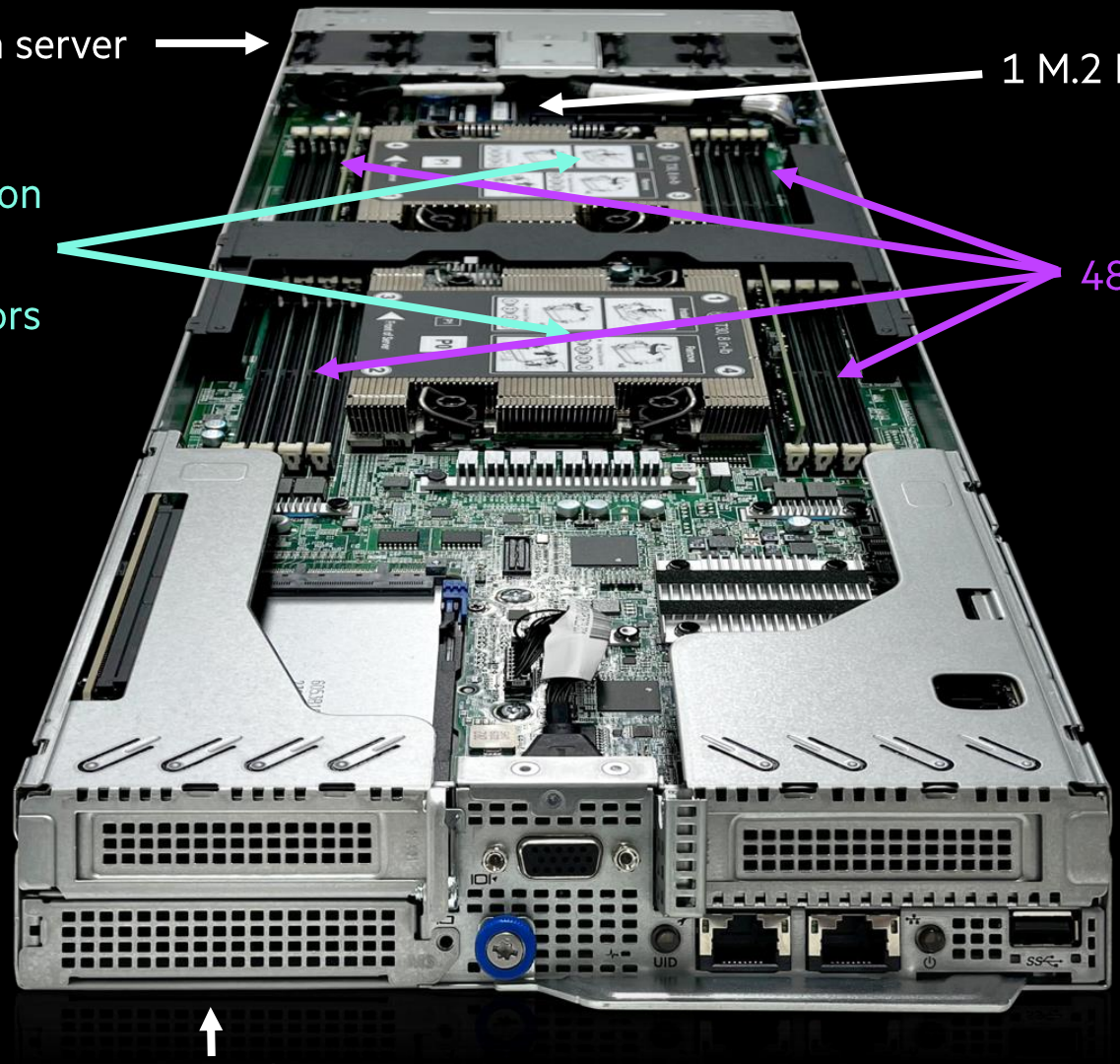
Slingshot 11 Configurations:
200Gpbs per node
400Gpbs per node

M.2 NVMe
x 4 per blade

CRAY EX 2500

- Cray EX 2500 provides DLC (Direct Liquid Cooling) to all components in the server, including CPU/GPU, memory, voltage regulators, etc.
- Supports the latest top bin CPU and GPU technology
- Racks can be populated one chassis at a time and scale over time. Pre-filled with coolant and ready to connect to existing system.
- Dedicated in-rack CDU, allows additional Cray EX 2500 cabinets to be added to the existing cluster over time
- Up to 96 nodes per cabinet
- Integrated liquid cooled switches
- 100% liquid cooling, up to 136kW
- Supports top or bottom feed facility water
- 380V DC internally

HPE CRAY XD220V 1U SERVER NODE

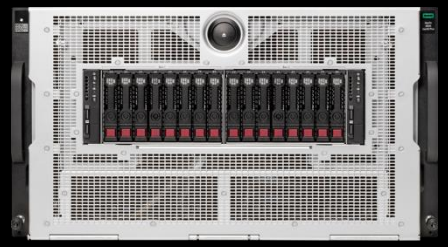


HPE APOLLO 6500 GEN10 PLUS -> CRAY XD 6500

Accelerated HPC and AI Servers

CY 2022 → **CY 2023**

Apollo 6500 Gen10 Plus



HPE CRAY XD 6500

Models tailored for HPC, AI, Mixed Workloads

XD670

X8 GPU - Standard x86 Architecture

- 2P Intel Sapphire Rapids + 8x NVIDIA H100 GPU
- AI Training workloads

HPE PROLIANT GEN11 SERVERS FOR EVERYWHERE YOUR APPS AND DATA LIVE

Orderable from March 2023

Accelerator optimized

2U, 2P

4th Gen Intel Xeon Scalable processor

Next Gen Enterprise Workloads

Extended GPU scalability
Up to 4 Double-wide or 8 single-wide accelerators

Cloud optimized

Blade, 2P

HPE Synergy 480 Gen11
4th Gen Intel Xeon Scalable processor

Private Cloud for Enterprise

Software-defined data center
High density compute
Flexible storage

Orderable from May 2023

SMB optimized

Tower, 1P

HPE ProLiant ML110 Gen11
4th Gen Intel Xeon Scalable processor

Compute for growing SMB

Office / branch footprint
Optimized for affordability
Expanded memory & IO

Big Data optimized

2U, 4P

HPE ProLiant DL560 Gen11
4th Gen Intel Xeon Scalable processor

High Performance Analytics

Quad socket density
In-memory storage
Advanced scalability

Orderable - Aug'23

Telco optimized

1U, 1P

HPE ProLiant DL110 Gen11
4th Gen Intel Xeon Scalable processor

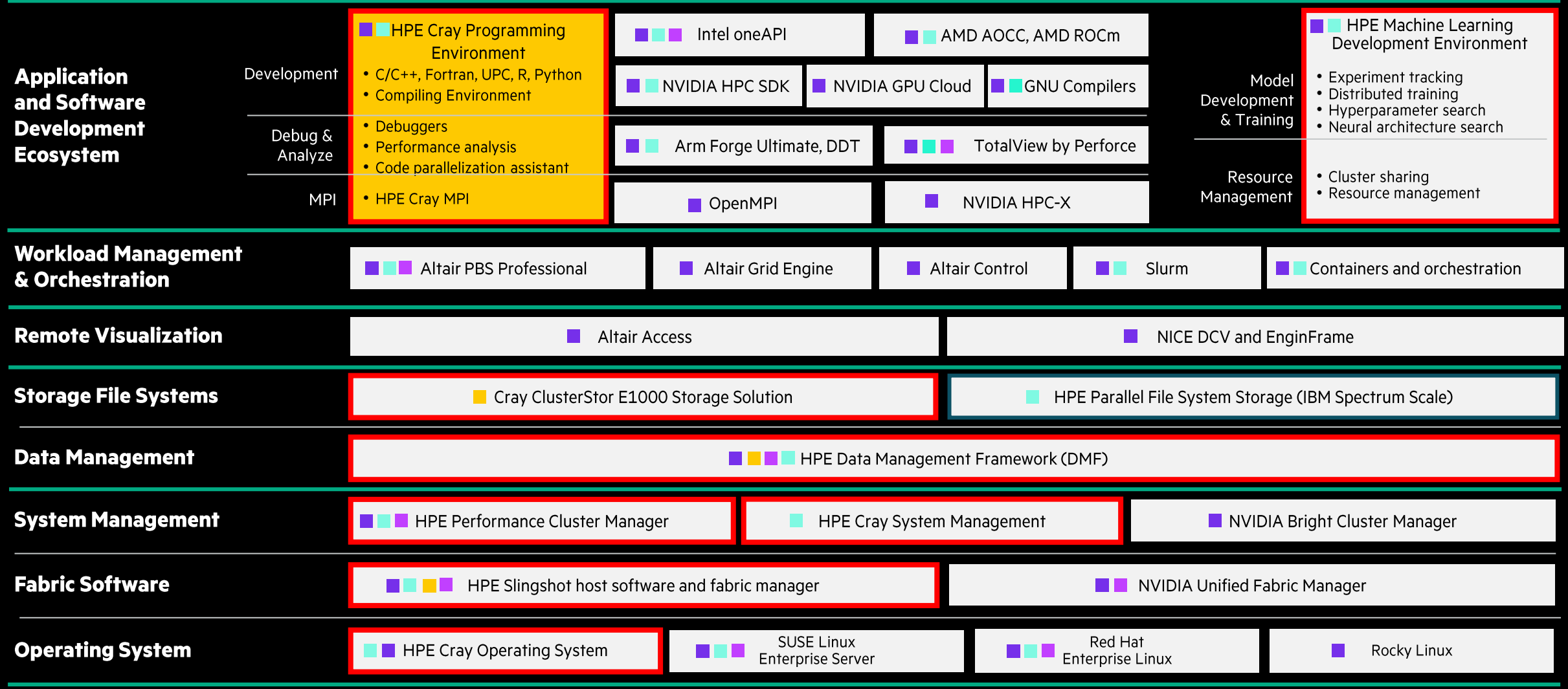
5G Services and vRAN density

IO Density
Small rack footprint
Single socket density

Workload ecosystem for HPE ProLiant Gen11 with 4th Gen Intel® Xeon® Scalable processor

Private and Hybrid Cloud	Containers	AI/ML and Data	VDI
VMware vSphere, VCF Microsoft Azure Stack HCI	HPE Ezmeral, Red Hat, SUSE	Microsoft, Oracle, PostgreSQL, MySQL, MongoDB, Cassandra	Citrix VAD, HP Anyware, VMware Horizon

HPE HPC AND AI SOFTWARE PORTFOLIO



AGENDA

Intel's EXASCALE Systems

HPE HPC & AI Compute Portfolio

HPE Cooling Technologies

LIQUID COOLING – CUSTOMER JOURNEY

1

10kw - 20kw racks

TCO is top of mind, while also considering environmental legislation

Air cooling remains major part of cooling in datacenter overall

Looking to the industry to see the technology working without issue and hear from peers about successes/challenges

Relevant HPE Solutions: RDHX, ARCS

Example customers: Finance, Oil and Gas

2

20kw - 50kw racks

Investing in the future to keep up with new CPU/GPU technology

Prefer familiar and flexible infrastructure such as standard racks and air cooling, coupled with liquid cooling for power systems

Relevant HPE Solutions: Apollo DLC with or without RDHX/ARCS

Example customers: Mid sized university research centers, Automotive

3

60kW+ racks

Driven by ambitious agenda to achieve best performance possible

Prepared to build supporting infrastructure to ensure product succeeds (water supply, new building, etc)

Relevant HPE Solutions: Apollo 2000 DLC + RDHX/ARCS, Cray EX 2500, Cray EX

Example customers: Large university research centers, National Agencies

COOLING TECHNOLOGY DIFFERENCES

Liquid to Air Cooling

Chilled water supply from the facility cools down the air-cooling system positioned close to the servers.

RDHX
Rear Door
Heat Exchanger

ARCS
Adaptive Rack
Cooling Solution

Hybrid Cooling

Combined direct liquid cooling and air cooling

Apollo and ProLiant

Direct Liquid Cooling

Coolant flows through a network of tubes and coldplates to extract heat directly from all components on the server

Cray EX 2500

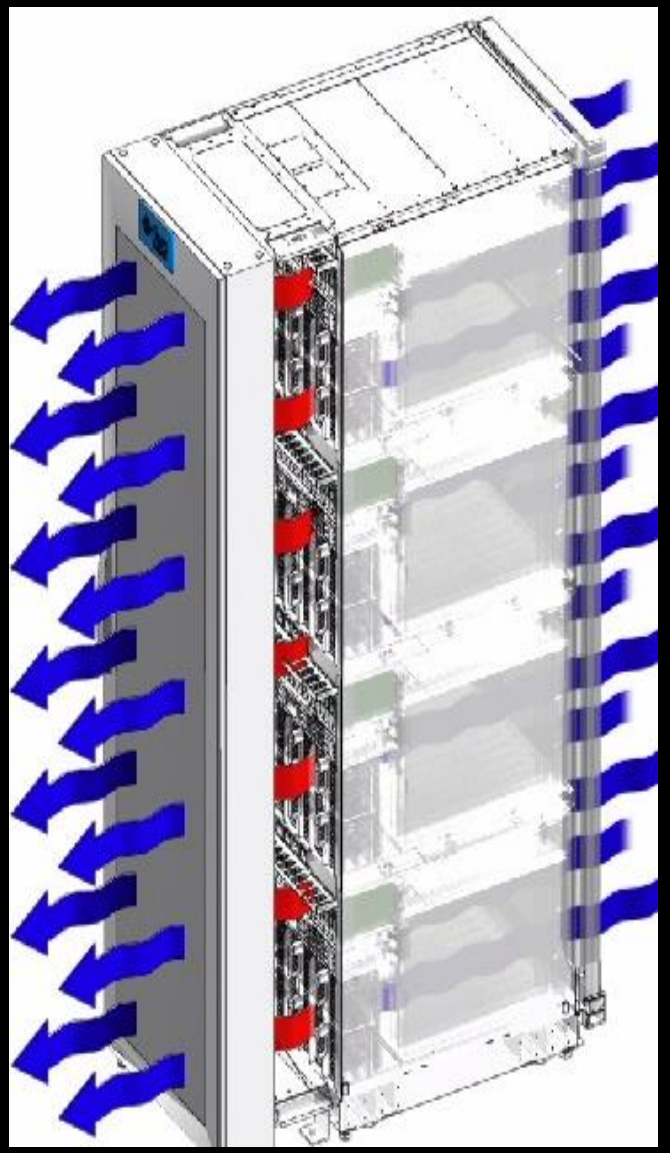
Cray EX

Cooling efficiency and capacity (kW/rack) increases from left to right

HPE CRAY XD/DL + RDHX

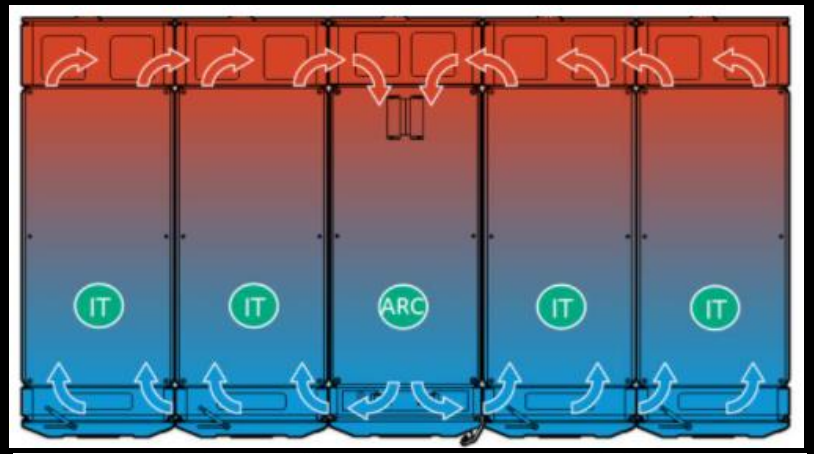
The HPE RDHX uses cooled facility water to neutralize the hot exhaust air from the servers, reducing the amount of air cooling required in the facility.

- Many door SKUs to match heat load generated, rack fit & other options
- Typically configured room-neutral cooling solution
- Can deploy with Cray XD 2000/6500 and DL DLC to meet room-neutral configuration
- Direct facility water feed or with a CDU
- If facility water supply:
  - Below dewpoint = CDU required
  - 2C above dewpoint = Direct facility connection



HPE CRAY XD + ARCS (ADAPTIVE RACK COOLING SYSTEM)

- The HPE ARCS allows for in-rack IT cooling without adding to the heat load in the data center.
- Using a closed-loop, room-agnostic design, ARCS is capable of cooling fully populated racks. The horizontal airflow of the HPE Adaptive Rack Cooling system fully supports Apollo, DL, Cray XD and ClusterStor, and industry-standard front-to-back hardware.
- Implementation of variable speed fans within ARCS enables improved energy efficiencies by providing the right volume of airflow to all IT regardless of mounting position or workload.
- Acoustic benefits and reduced number of water taps bring advantages over other rack-level cooling products.
- ARCS enables customer datacenter power density to increase without need for facilities upgrades.



Top View

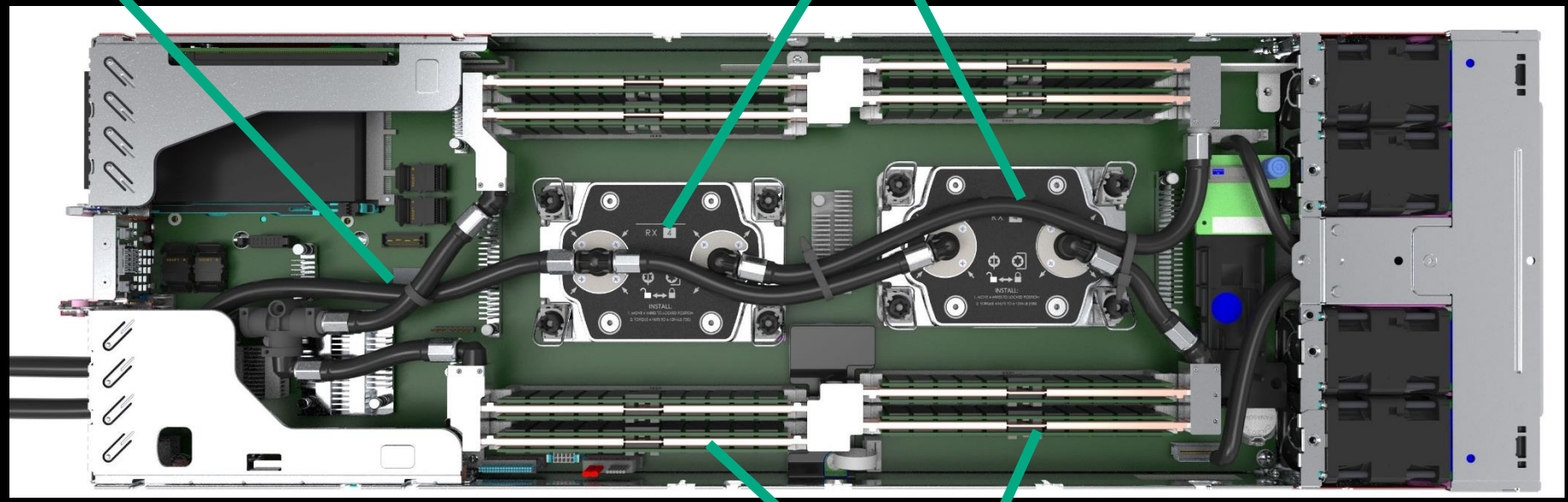
HPE Adaptive Rack Cooling System with 42u Rack Options

HPE CRAY XD DIRECT LIQUID COOLING – CPU AND MEMORY COOLING

Flexible Tubing for Component Serviceability

CPU Coldplates

Dripless Quick Disconnects

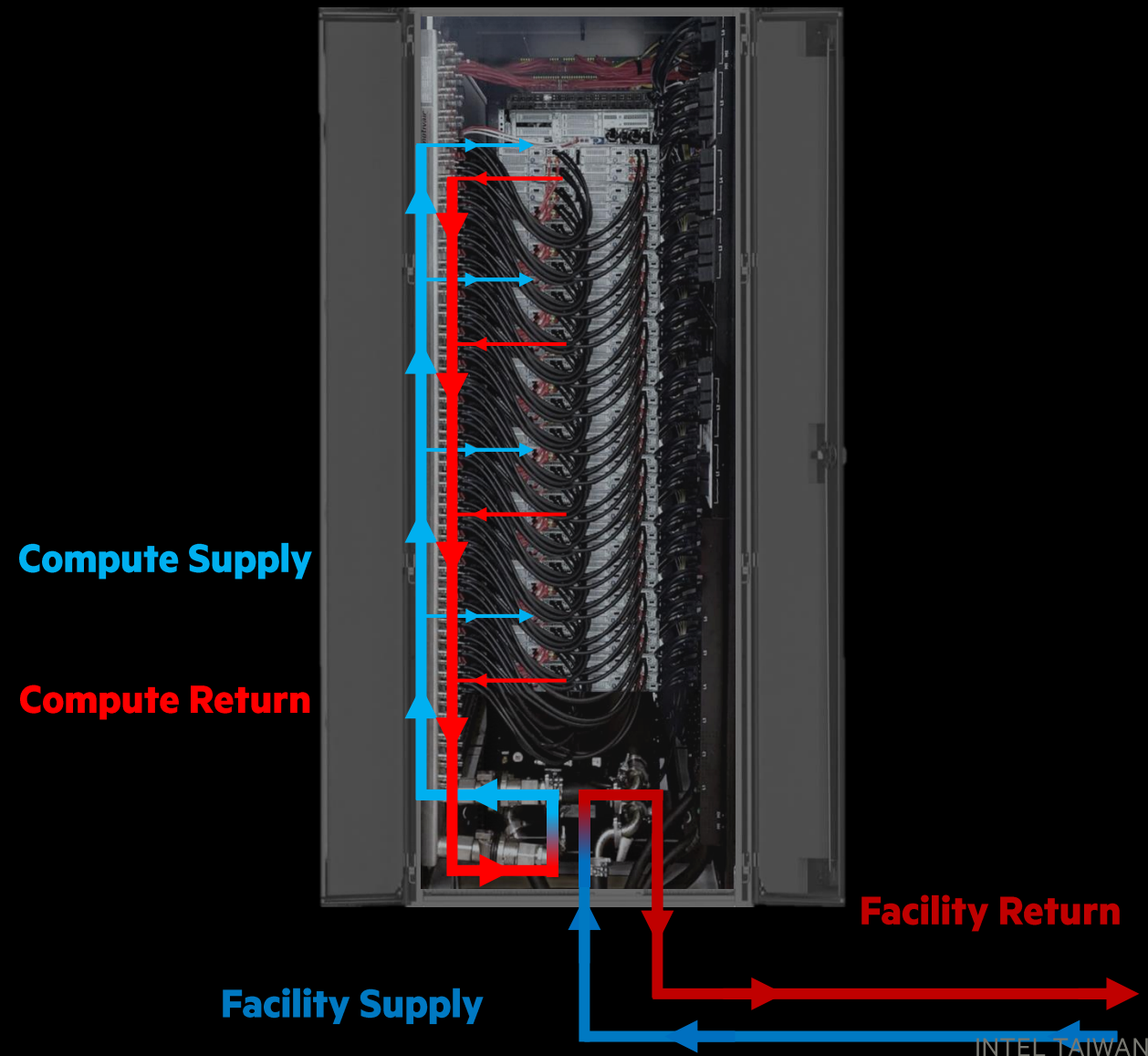


Memory Cooling (optional)

HPE CRAY XD2000 – THERMAL EXAMPLE

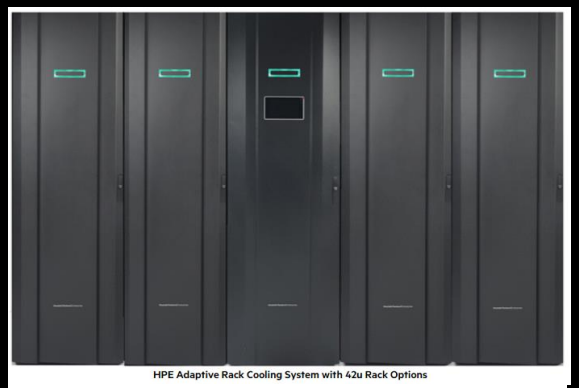
Example Parameters:

- 72 HPE Cray XD2000 servers with 280W CPU
- 26GPM Facility Supply, 27GPM Compute Supply
- +70% server heat into water



CRAY XD2000 DLC IN ROOM-NEUTRAL CONFIGURATION

+



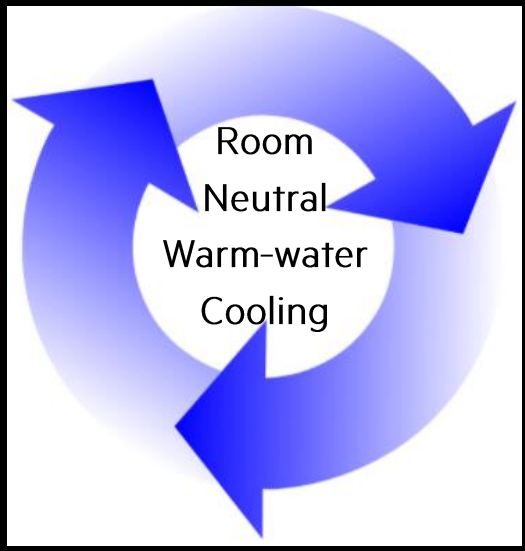
HPE Adaptive Rack Cooling System with 42u Rack Options

ARCS

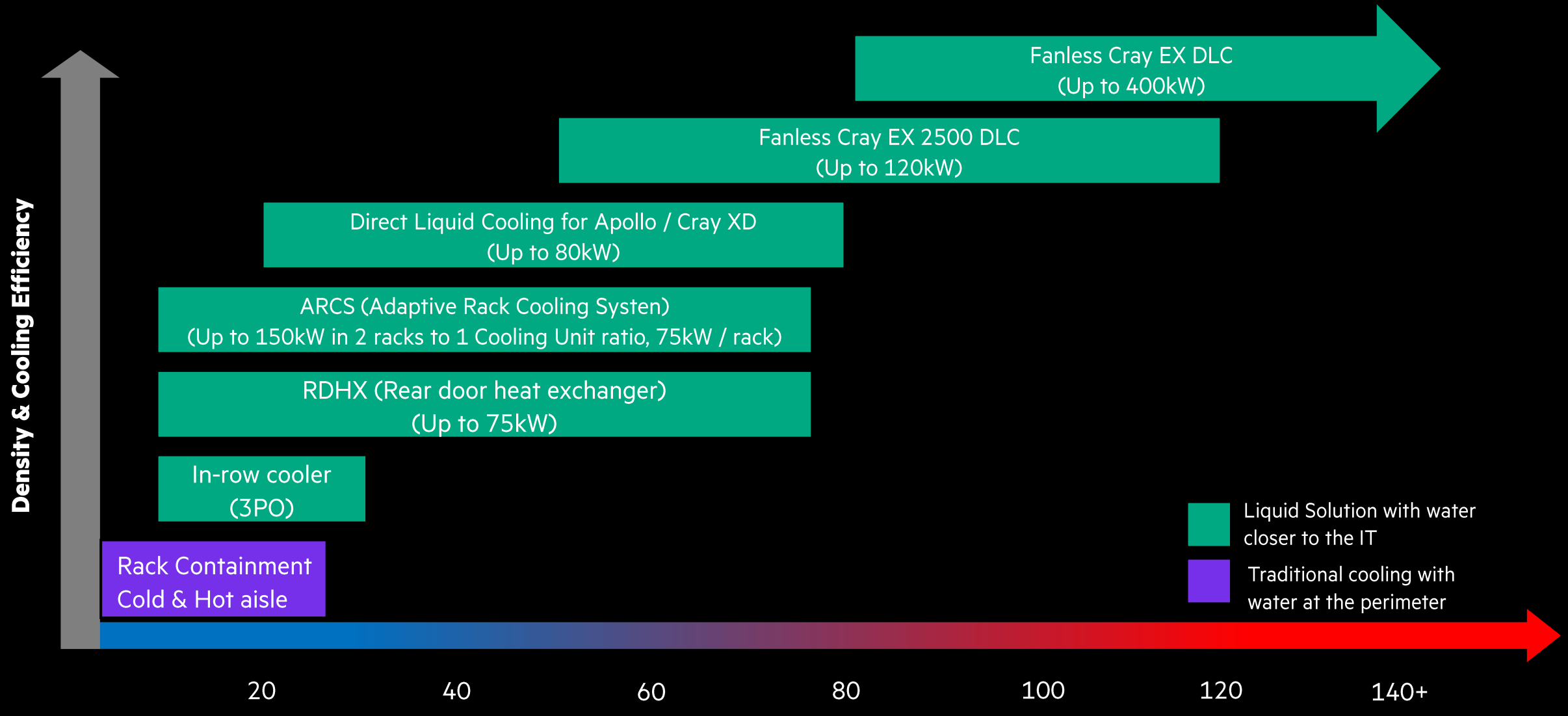
OR

RDHx

=



COOLING SOLUTIONS THERMAL RANGE



THANK YOU