



**Hewlett Packard  
Enterprise**

# **HPE Update HPC User Forum**

---

Mike Vildibill  
VP and GM  
Slingshot High Performance Networking

April 2025



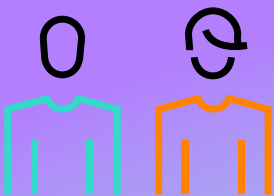
# HPE at-a-glance

## Advancing the way people live and work

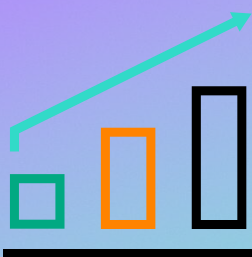
### By the numbers



HPE **founded In 2015** following the separation from HP, Inc.



**60,000+** team members



**\$30.1B** revenue in FY24

### Sustainability



**Net-zero GHG emissions** across our value chain by 2040



### Corporate citizenship



\*Total contract value (TCV) represents the total value of one-time and recurring revenue for the period defined by the contract

# HPC & AI Infrastructure Solutions

## Who we are

Focused on delivering fully integrated end-to-end high-performance computing systems to support customers that target complex scientific workloads and large-scale AI modeling needs.

- **1,770 team members across the globe**
- **Global scale in 140+ countries**
- **Worldwide expert services**
- **#1 HPC vendor worldwide**
- **Global leader in 100% fanless direct liquid-cooling**

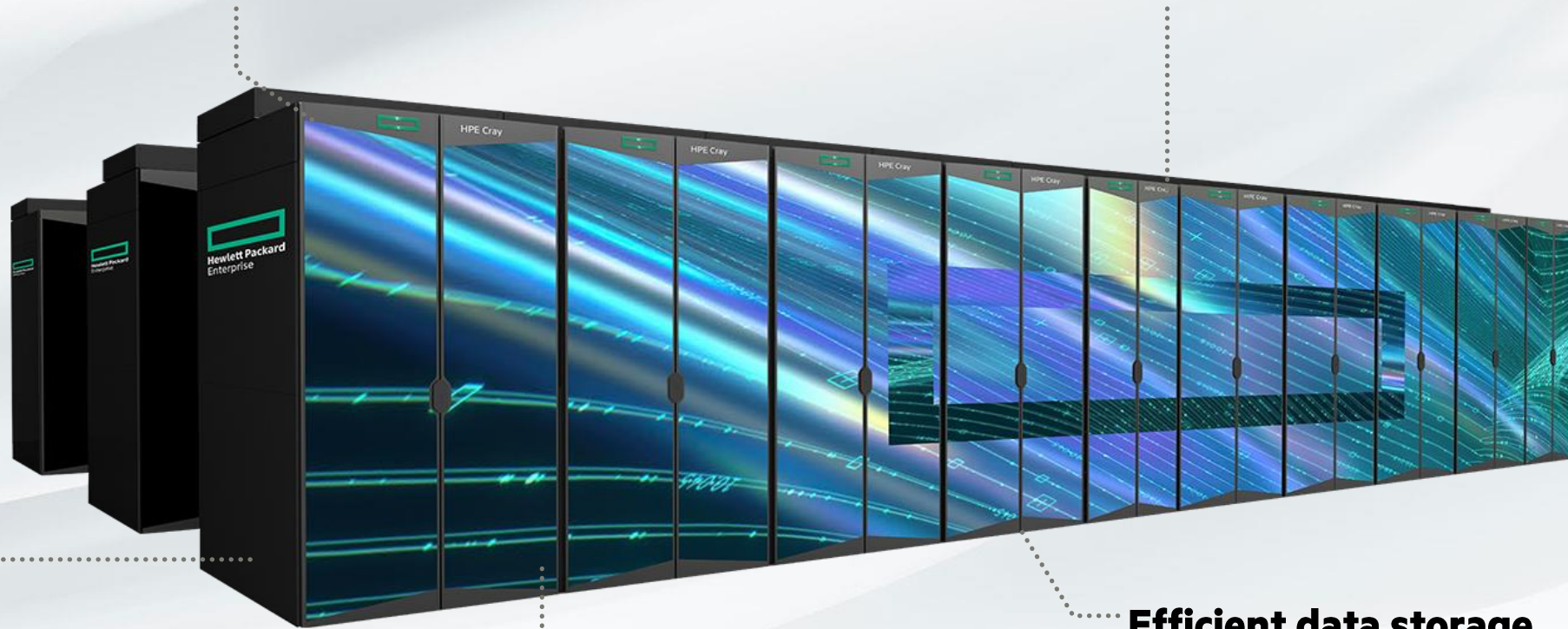
# HPE Cray EX supercomputer

End-to-end supercomputing solution built for the exascale era

System management software for resiliency & efficiency

Interconnect built for HPC and AI performance at-scale

Sustainable solutions with 100% fanless Direct Liquid-Cooling



Optimize performance for HPC & AI workflows with developer software

Efficient data storage and accessibility



# Research and Production. CPU/GPU breadth. Clusters to Supercomputers. HPE Slingshot and Industry NICs.

November 2024

HPE Slingshot in Top 500



NERSC



OAK RIDGE  
National Laboratory



CSCS



Argonne  
NATIONAL LABORATORY



Lawrence Livermore  
National Laboratory



pawsey



EuroHPC  
Joint Undertaking



NREL  
NATIONAL RENEWABLE ENERGY LABORATORY



Los Alamos  
NATIONAL LABORATORY



KAUST



US Air Force Weather (ORNL)



NCAR



INES

7 of the Top 10

61% of the aggregate compute power of the Top 100

97% of the aggregate compute power of the most Energy-Efficient Top 25

100% of verified systems to break exaFLOP barrier

# EXTREME PERFORMANCE BEYOND TRADITIONAL HPC SUPERCOMPUTING

- Isambard-AI is UK's fastest AI supercomputer
  - 5,448 Grace-Hopper GPUs
  - over **21 ExaFLOP/s** of 8-bit FP performance for large language model (LLM) training
  - **< 5MW of power**
- Isambard-AI's software stack is designed for users accustomed to **cloud-based platforms**
  - accessed via Jupyter notebooks, MLOps, and other web-based interfaces with full support for containers
- Cloud capabilities support leading edge GenAI LLMs
  - enabled via a **container-based platform** with a **Kubernetes (K8s)** infrastructure and **high speed RDMA**
  - ... all running on the **HPE Slingshot** fabric

13 May 2024

The most sustainable, powerful and fastest AI supercomputer in the UK, located at the University of Bristol, is now officially online, with pioneering technology helping to make the UK a world leader in artificial intelligence



# Open Industry Ecosystems for High Performance Fabrics

## Industry Driving Standards-based Scale-out and Scale-up Networks

### Ultra-Ethernet Consortium (UEC)

Scale-out High-Performance Ethernet



Founded late 2021; Public announcement July 2023

Over 70 industry partners joined since going public;  
the fastest growing Linux Foundation org ever



#### **Consortium Motivations:**

- Industry desire to expand Ethernet for broad AI and HPC use cases
- UEC Focus: Performance and Switch  $\leftrightarrow$  NIC interoperability:
  - **Accelerator and Network vendors:** Discrete & integrated NICs, switches with AI/HPC features (Congestion Mgmt., Collectives, etc) in an open Ethernet model
  - For users, reduced risk, no vendor lock-in, familiar datacenter tech

#### **HPE goals:**

- Expand addressable market with 3<sup>rd</sup> party and integrated NICs
- Enable Ethernet as **Leading Choice For AI & HPC Fabrics**
- Industry leadership with HPE Slingshot IP integrated into the standard

### Ultra Accelerator Link Consortium (UALink)

Scale-up High-Performance Accelerator Interconnect



Public announcement May 30, 2024

9 founding members as BoD responsible for driving the technology;  
ecosystem adopters to be added after initial spec release

#### **Consortium Motivations:**

- Address the industry gap in scale-up accelerator/GPU interconnects created as models and workloads get larger and are distributed across multiple accelerators; NVLink does this today
- Move fast to catch competition and enable an open datacenter ecosystem

#### **HPE goals:**

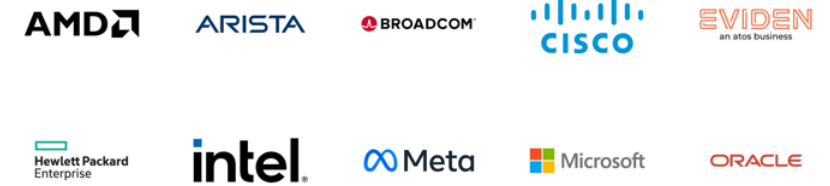
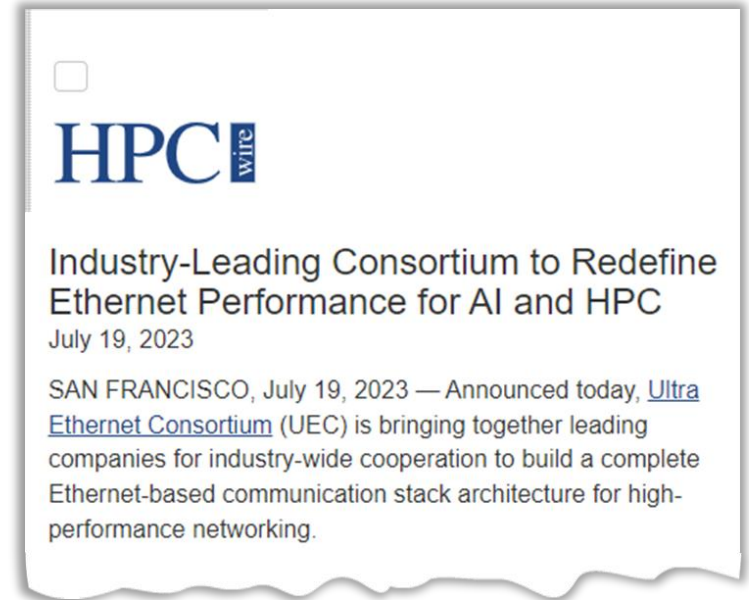
- Expand HPE's focus on open networks and protocols
- Deliver the short and long-term performance needed for compute intensive workloads; a memory semantic fabric has advantages
- Expand Slingshot products and technology into a fast growing, adjacent market; Enable greater adoption of scale-up Interconnects

# The Industry is Moving to High-Performance Ethernet

HPE Slingshot is there TODAY



- HPE Slingshot technology is the foundation for many of the UEC transport and link level innovations, **Slingshot gives you UEC performance today** and compatibility for tomorrow.
- HPE Slingshot, combined with the forward-thinking innovations of the Ultra Ethernet Consortium, provides an unmatched network solution for AI and HPC.
- Slingshot ensures that organizations can meet the demands of tomorrow's computational challenges while maximizing their infrastructure investments.
- This positions Slingshot as a leading solution for modern workloads, emphasizing its compatibility with future innovations in Ultra Ethernet.



Steering Committee Members



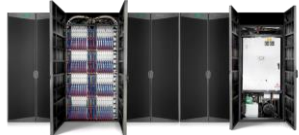
# HPC and AI systems portfolio

## Leadership-class supercomputing

100% fanless direct liquid cooling

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

### HPE Cray Supercomputing EX4000



### HPE Cray Supercomputing EX2500



**HPE Slingshot** combines the performance of a supercomputing interconnect with the cost-effectiveness of Ethernet



## Accelerated AI

70% direct liquid cooling, liquid to air cooling

Purpose-built, 8-way AI servers for AI model training, tuning and inference

### HPE ProLiant Compute XD680



### HPE ProLiant Compute XD685



### HPE Cray XD670



### HPE Cray XD675



HPE ProLiant Compute accelerating AI applications for Enterprises

### HPE ProLiant Compute DL380a Gen12



### HPE ProLiant Compute DL384 Gen12



## Mainstream HPC/AI

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

### HPE Cray XD2000



### HPE Cray XD665



Integrated HPC and AI software portfolio, including application and software development ecosystem, system management suite, orchestration tools, enhanced compute environment & more

HPE Services Experts available globally to accelerate your strategic HPC and AI initiatives

## Purpose-built storage

Unprecedented data storage price/performance for HPC, AI, and converged workloads

### HPE Cray Supercomputing Storage Systems E2000



### Cray ClusterStor E1000 Storage Systems



### HPE Cray Storage Systems C500



# Slingshot for Open Standards-based HPC and AI Ecosystems

## Form Factors & Networking



64 port 200GbE ToR Switch  
64 port 400GbE ToR Switch (2Q2025)



1x200GbE PCIe NIC  
1x400GbE PCIe NIC (2Q2025)  
2x400GbE OCP-W NIC (2026)

Ultra Ethernet Consortium

## Host OS & EcoSystem

HPE Cray OS

SUSE Linux Enterprise Server 15



OS Support



NVIDIA/nccl

Optimized primitives for collective multi-GPU communication



ROCm/aws-ofi-rccl



intel/mpi

Intel MPI Library



OPEN MPI

RDMA/Messaging Middleware

## Cluster & Workload Management

HPE Performance Cluster Manager

OpenCHAMI

HPC System Management for Cloud Engineers and HPC Sysadmins



Altair | PBS Works™



## Telemetry & Monitoring

LDMS



HPE Slingshot Monitoring & AIOps

# Thank You

