

INTRODUCTION TO THE FRONTIER SYSTEM

Frontier Application Readiness Kick-Off Workshop

Oct. 2019

CRAY[®]

a Hewlett Packard Enterprise company



Glenski@cray.com

Technical Data Rights

All materials contained in, attached to, or referenced by this document that are marked Cray Confidential or with a similar restrictive legend may not be disclosed in any form without the advance written permission of Cray, a Hewlett Packard Enterprise company. These data are submitted with limited rights under Government Contract No. B626589 and Lease Agreement 4000167127. These data may be reproduced and used by the Government with the express limitation that they will not, without written permission of Cray, be used for purposes of manufacture nor disclosed outside the Government.

This notice shall be marked on any reproduction of these data, in whole or in part.

Copyright and Trademark Acknowledgements



©2016-2019 Cray, a Hewlett Packard Enterprise company, All Rights Reserved.

Portions Copyright Advanced Micro Devices, Inc. (“AMD”) Confidential and Proprietary.

The following are trademarks of Cray, and are registered in the United States and other countries: CRAY and design, SONEXION, URIKA, and YARCDATA. The following are trademarks of Cray: APPRENTICE2, CHAPEL, CLUSTER CONNECT, CLUSTERSTOR, CRAYDOC, CRAYPAT, CRAYPORT, DATAWARP, ECOPHLEX, LIBSCI, NODEKARE, and REVEAL. The following system family marks, and associated model number marks, are trademarks of Cray: CS, CX, XC, XE, XK, XMT, and XT. ARM is a registered trademark of ARM Limited (or its subsidiaries) in the EU and/or elsewhere. ThunderX, ThunderX2, and ThunderX3 are trademarks or registered trademarks of Cavium Inc. in the U.S. and other countries. The registered trademark LINUX is used pursuant to a sublicense from LMI, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis. Intel, the Intel logo, Intel Cilk, Intel True Scale Fabric, Intel VTune, Xeon, and Intel Xeon Phi are trademarks or registered trademarks of Intel Corporation in the U.S. and/or other countries. Lustre is a trademark of Xyratex. NVIDIA, Kepler, and CUDA are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and/or other countries.

Other trademarks used in this document are the property of their respective owners.

FORWARD LOOKING STATEMENTS

This presentation may contain forward-looking statements that involve risks, uncertainties and assumptions. If the risks or uncertainties ever materialize or the assumptions prove incorrect, the results of Hewlett Packard Enterprise Company and its consolidated subsidiaries ("Hewlett Packard Enterprise") may differ materially from those expressed or implied by such forward-looking statements and assumptions. All statements other than statements of historical fact are statements that could be deemed forward-looking statements, including but not limited to any statements regarding the expected benefits and costs of the transaction contemplated by this presentation; the expected timing of the completion of the transaction; the ability of HPE, its subsidiaries and Cray to complete the transaction considering the various conditions to the transaction, some of which are outside the parties' control, including those conditions related to regulatory approvals; projections of revenue, margins, expenses, net earnings, net earnings per share, cash flows, or other financial items; any statements concerning the expected development, performance, market share or competitive performance relating to products or services; any statements regarding current or future macroeconomic trends or events and the impact of those trends and events on Hewlett Packard Enterprise and its financial performance; any statements of expectation or belief; and any statements of assumptions underlying any of the foregoing. Risks, uncertainties and assumptions include the possibility that expected benefits of the transaction described in this presentation may not materialize as expected; that the transaction may not be timely completed, if at all; that, prior to the completion of the transaction, Cray's business may not perform as expected due to transaction-related uncertainty or other factors; that the parties are unable to successfully implement integration strategies; the need to address the many challenges facing Hewlett Packard Enterprise's businesses; the competitive pressures faced by Hewlett Packard Enterprise's businesses; risks associated with executing Hewlett Packard Enterprise's strategy; the impact of macroeconomic and geopolitical trends and events; the development and transition of new products and services and the enhancement of existing products and services to meet customer needs and respond to emerging technological trends; and other risks that are described in our Fiscal Year 2018 Annual Report on Form 10-K, and that are otherwise described or updated from time to time in Hewlett Packard Enterprise's other filings with the Securities and Exchange Commission, including but not limited to our subsequent Quarterly Reports on Form 10-Q. Hewlett Packard Enterprise assumes no obligation and does not intend to update these forward-looking statements.





U.S. DEPARTMENT OF
ENERGY



CRAY[®]

AMD 

News

U.S. Department of Energy and Cray to Deliver Record-Setting Frontier Supercomputer at ORNL

Exascale system expected to be world's most powerful computer for science and innovation

Topic: Supercomputing

May 7, 2019

Related Organizations

[Computing and Computational Sciences Directorate](#)



OAK RIDGE, Tenn., May 7, 2019—The U.S. Department of Energy today announced a contract with Cray Inc. to build the Frontier supercomputer at Oak Ridge National Laboratory, which is anticipated to debut in 2021 as the world's most powerful computer with a performance of greater than 1.5 exaflops.

Frontier is a Shasta system

Shasta is Cray's platform for the Exascale Era



Flexible, Extensible, & Scalable
Hardware Infrastructure

Wide Diversity of Processors

Slingshot Interconnect

Standards-based (interoperable and open)

High-Performance, Tiered, Integrated Storage

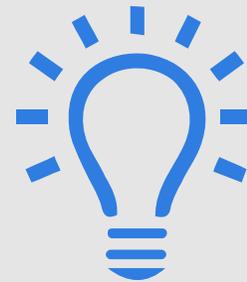
Dynamic, Cloud-like Environment
for Hybrid Workflows



HPC



AI



Analytics



Cloud



IoT

Shasta Flexible Compute Infrastructure

“Mountain”

Dense, scale-optimized Cabinet

- Up to 300KW with warm water cooling
- 512+ high-performance processors
- Flexible, high-density interconnect



“River”

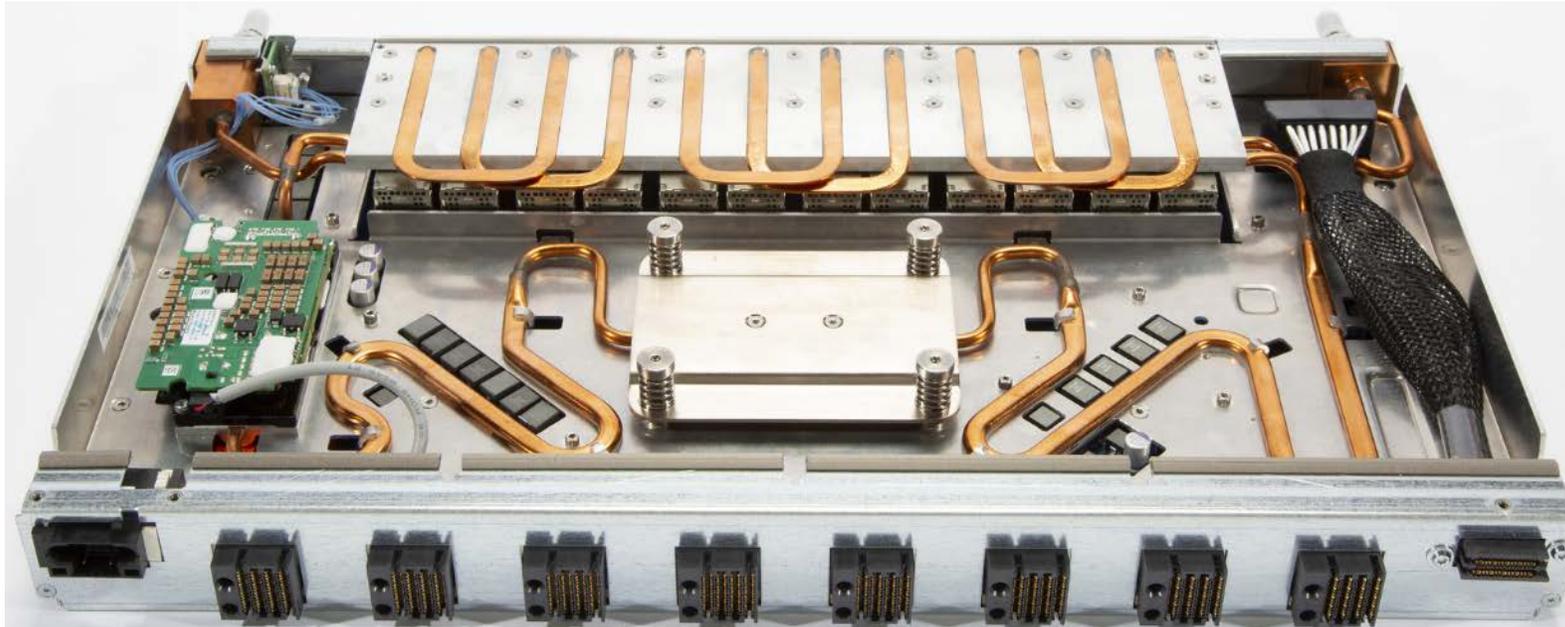
Standard 19” Rack

- Air cooled with liquid cooling options
- Wide range of available compute and storage



Same Interconnect - Same Software Environment

Slingshot Overview



Slingshot is Cray's 8th generation scalable interconnect

Earlier, Cray pioneered:

- ***Adaptive routing***
- ***High-radix switch design***
- ***Dragonfly topology***

64 ports x 200 Gbps

Over 250K endpoints with a diameter of just three hops

Ethernet Compliant

Easy connectivity to datacenters and third-party storage; "HPC inside"

World class Adaptive Routing and QoS

High utilization at scale; flawless support for hybrid workloads

Groundbreaking Congestion Control

Performance isolation between workloads

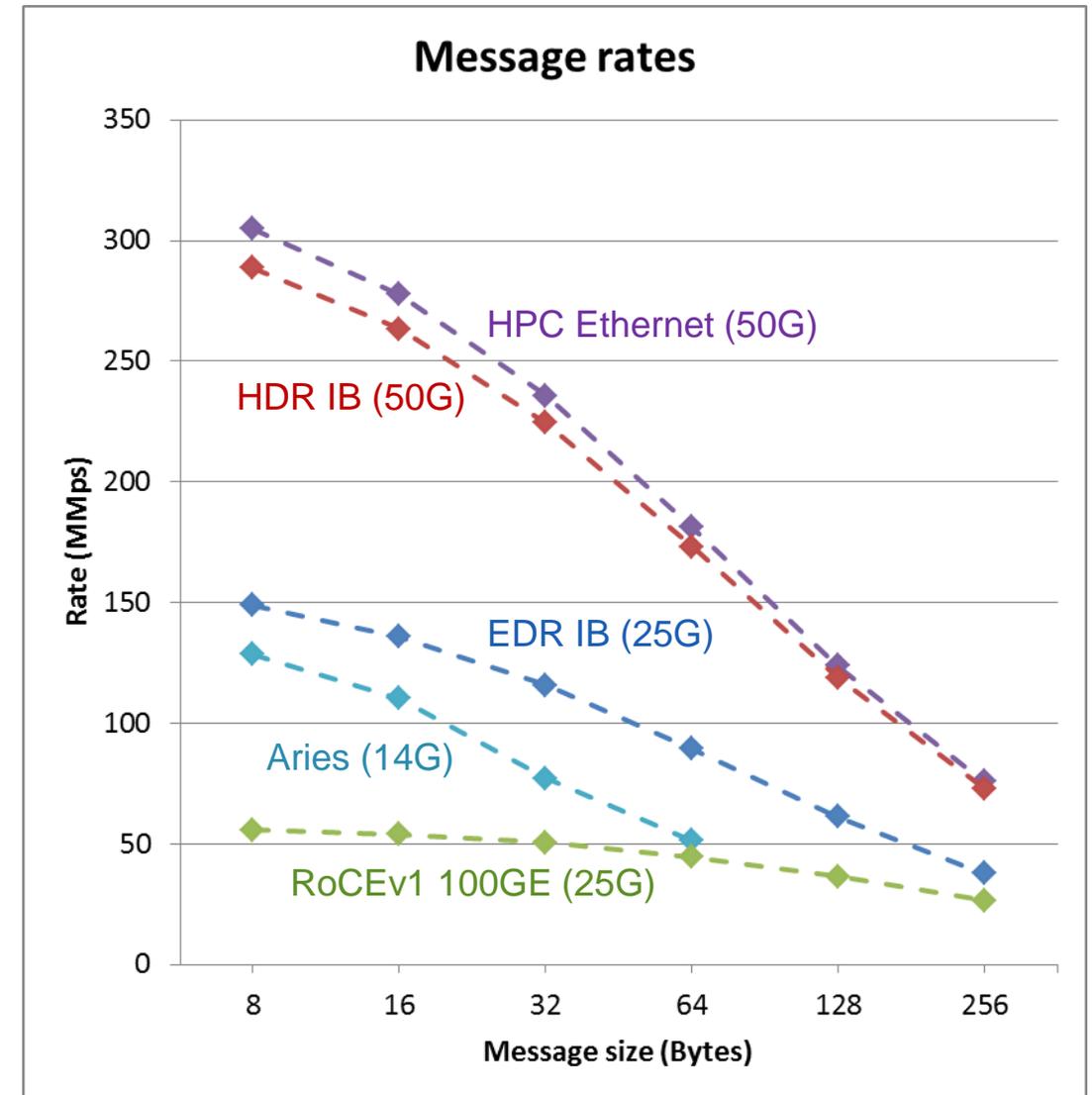
Low, Uniform Latency

Focus on tail latency, because real apps synchronize

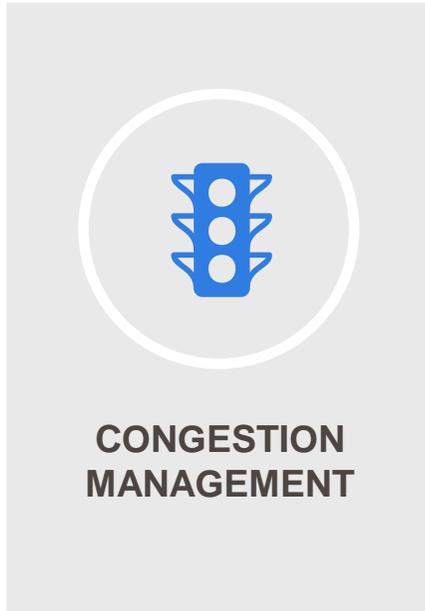
HPC Ethernet Protocol

Enhancements for Efficiency and Resiliency

- Slingshot speaks standard Ethernet at the edge, and optimized HPC Ethernet on internal links
- Reduced minimum frame size
- Removed inter-packet gap
- Optimized header
- Credit-based flow control
- **Protocol also provides resiliency benefits**
 - Low-latency FEC (see 25Gbit Ethernet Consortium)
 - Link level retry to tolerate transient errors
 - Lane degrade to tolerate hard failures



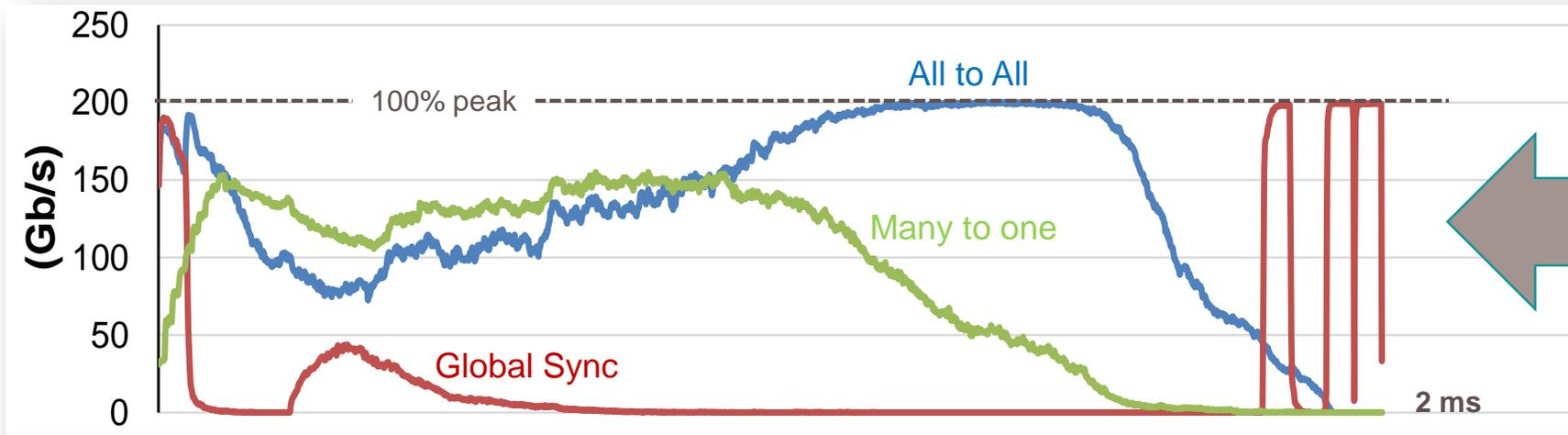
Slingshot Congestion Management



- Hardware automatically tracks *all* outstanding packets
 - Knows what is flowing between **every** pair of endpoints
- Quickly identifies and controls causes of congestion
 - Pushes back on sources... *just enough*
 - Frees up buffer space for everyone else
 - Other traffic not affected and can pass stalled traffic
 - **Avoids HOL blocking across entire fabric**
 - **Fundamentally different than traditional ECN-based congestion control**
- Fast and stable across wide variety of traffic patterns
 - Suitable for *dynamic HPC traffic*
- Performance isolation between apps on same QoS class
 - Applications much less vulnerable to other traffic on the network
 - Predictable runtimes
 - Lower mean *and tail* latency – a big benefit in apps with global synchronization

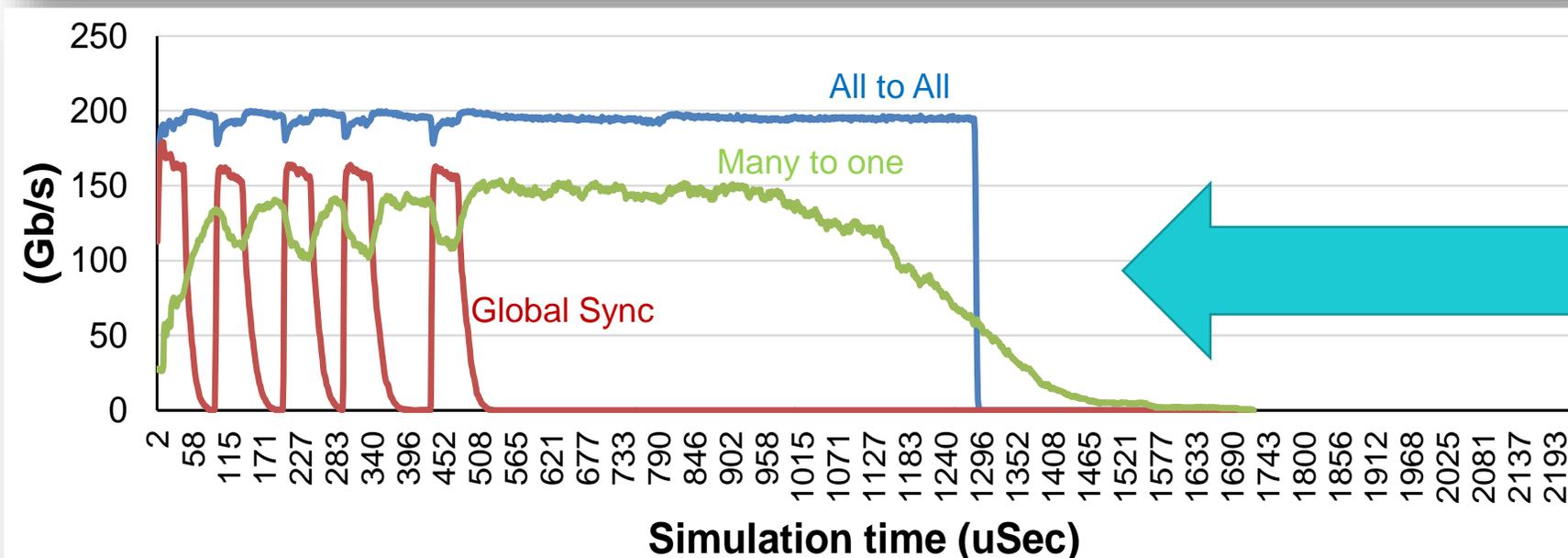
Congestion Management Provides Performance Isolation

Average egress
BW per endpoint



Job Interference in today's networks

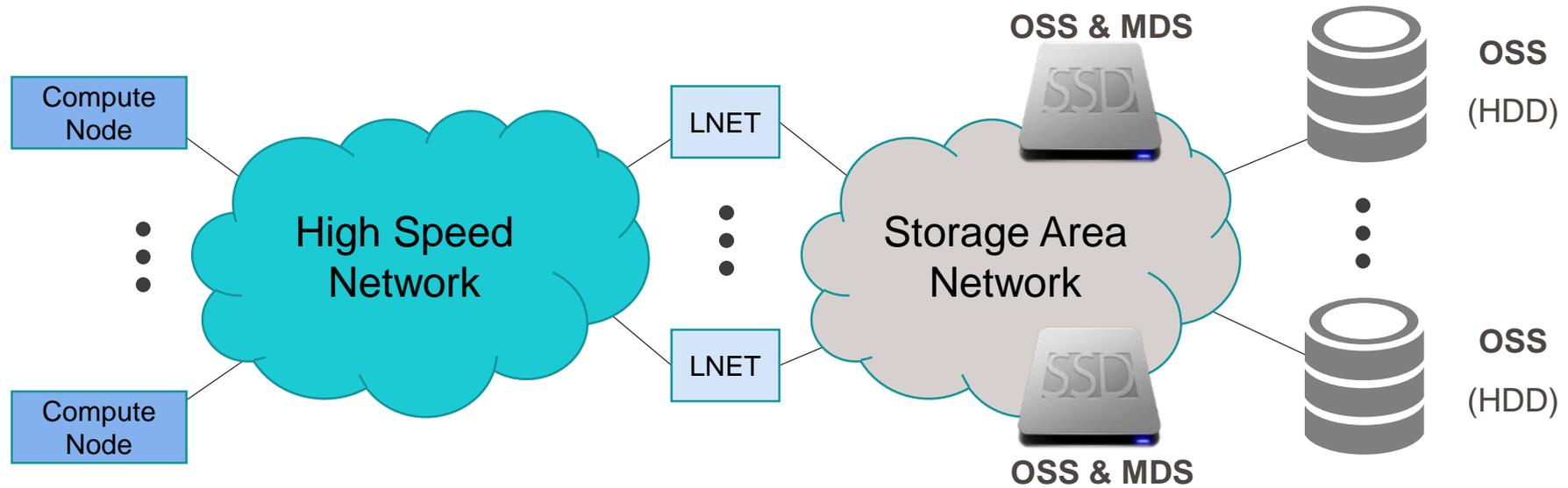
Congesting (green) traffic hurts well-behaved (blue) traffic, and *really* hurts latency-sensitive, synchronized (red) traffic.



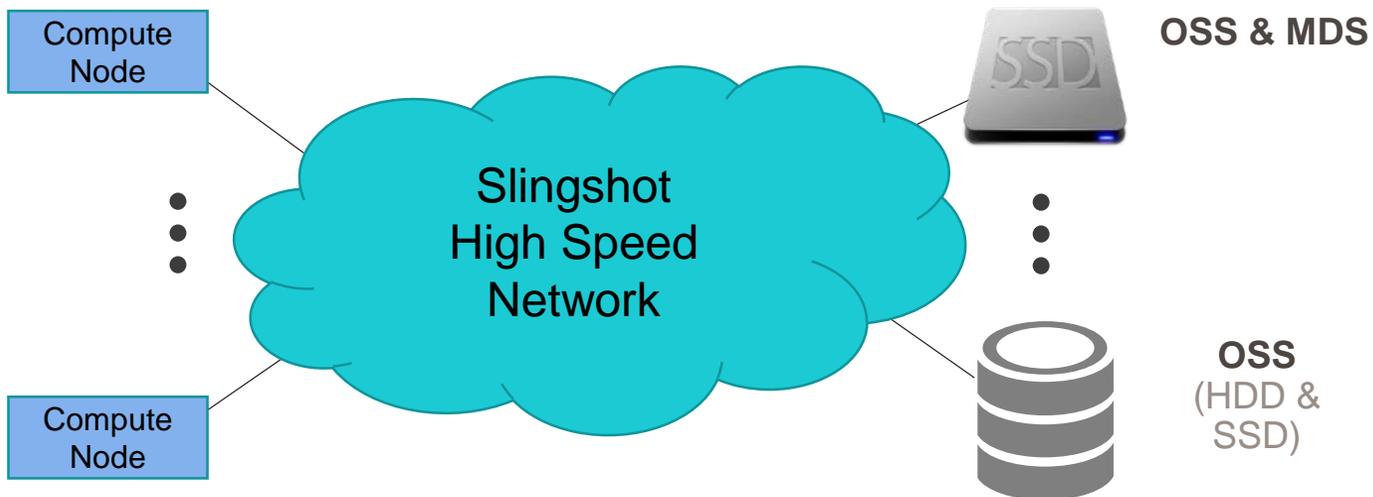
*With Slingshot
Advanced
Congestion
Management*

Shasta Pulls Storage onto Slingshot HSN

Traditional Model



Shasta

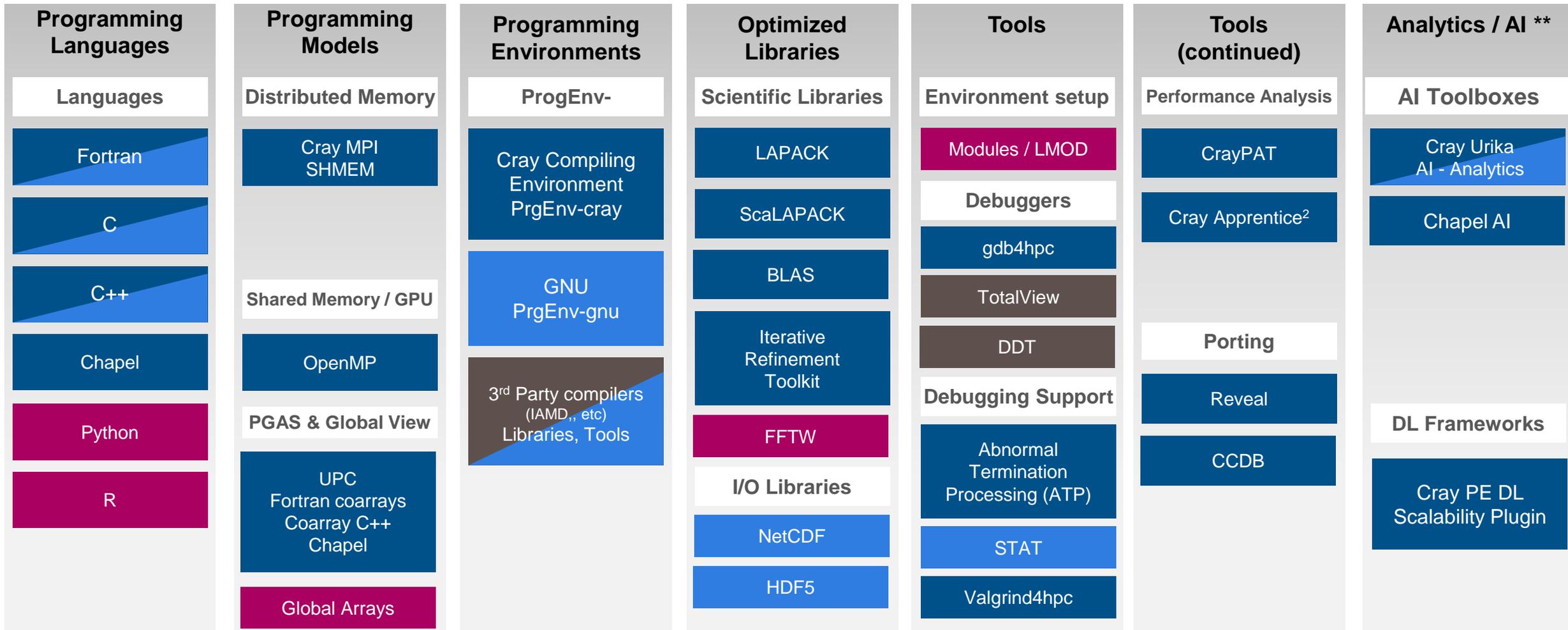


Tiered Flash and HDD Servers

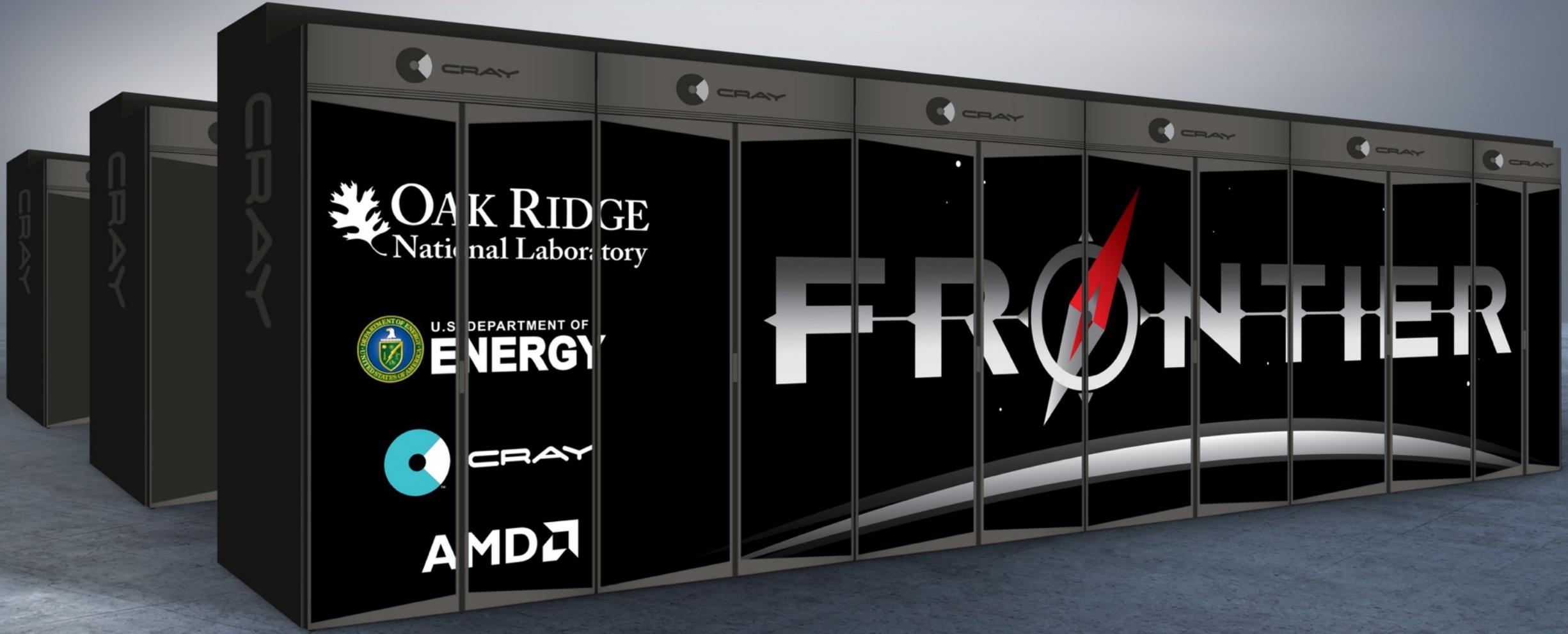
Benefits:

- Lower complexity
- Lower latency
- Improved small I/O performance

Shasta Developer Environment



Cray Developed
 3rd party packaging
 Cray added value to 3rd party
 Licensed ISV SW



 **OAK RIDGE**
National Laboratory

 U.S. DEPARTMENT OF
ENERGY

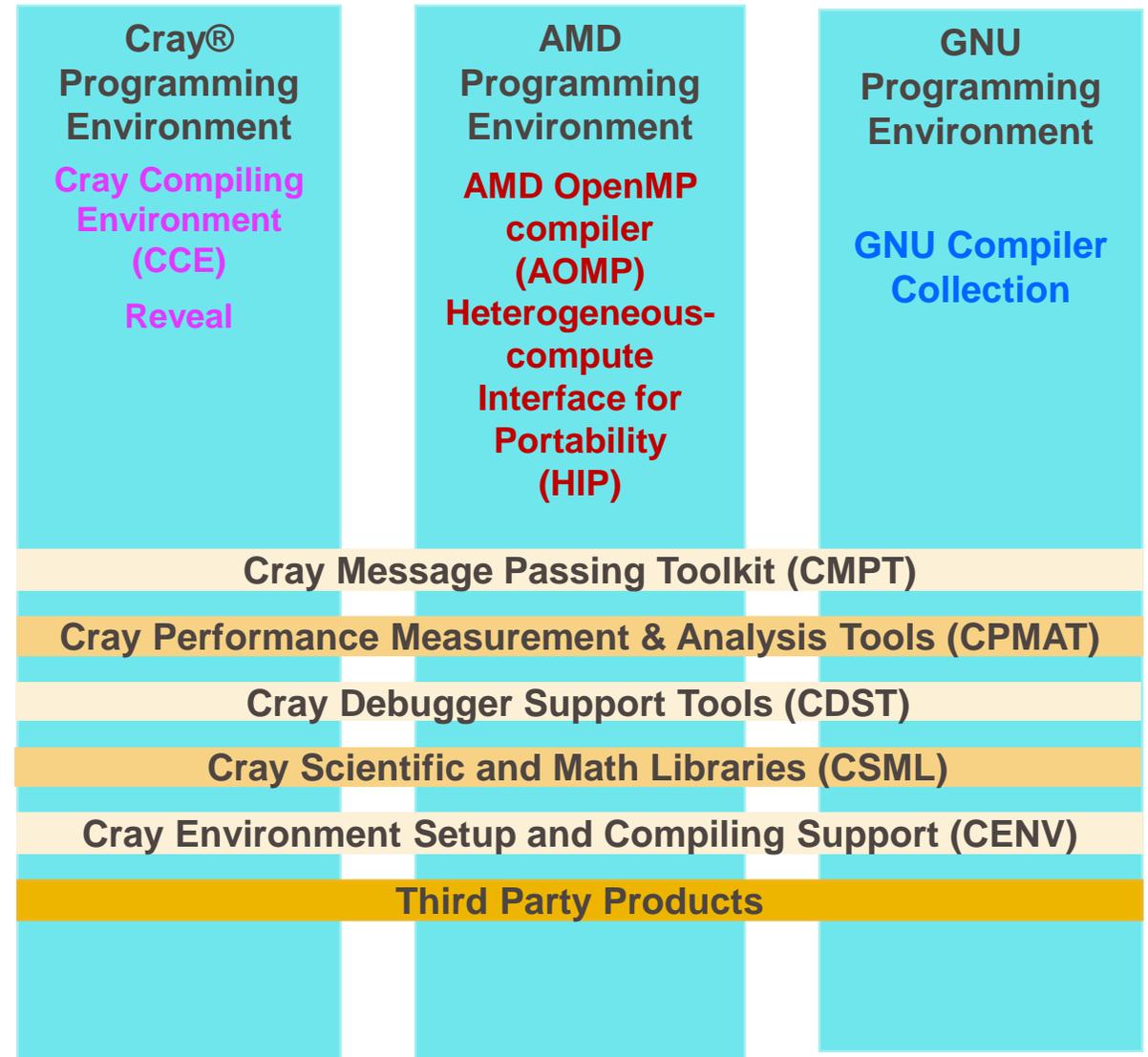
 **CRAY**

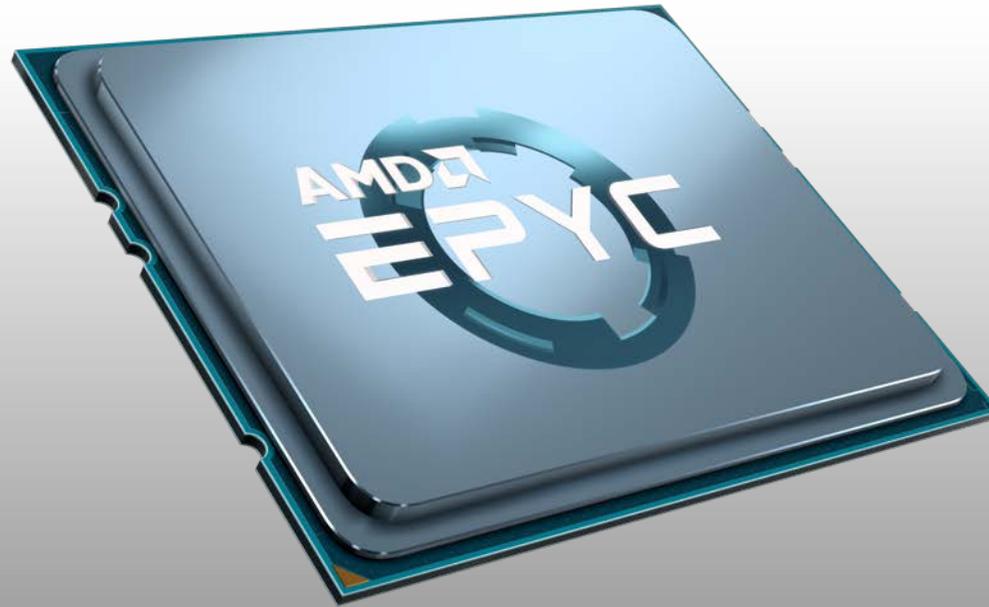
AMD 

FRONTIER

Frontier Application Software Stack

- Cray, and AMD are working together with ORNL and other Labs to deliver a full software stack
 - Provide Compiler and library choice
 - Includes:
 - Multiple programming environments
 - Performance and correctness tools
 - Will Include Optimizations such as:
 - Cray MPI GPU-to-GPU data movement
 - libsci_acc
 - Cray PE DL Plugin





HIGH PERFORMANCE CPU CUSTOMIZED FOR HPC

Custom AMD EPYC processor
optimized for HPC and AI

Utilizes Future “Zen” Core High-
Performance Architecture

AI-Optimized for
Supercomputing Workloads

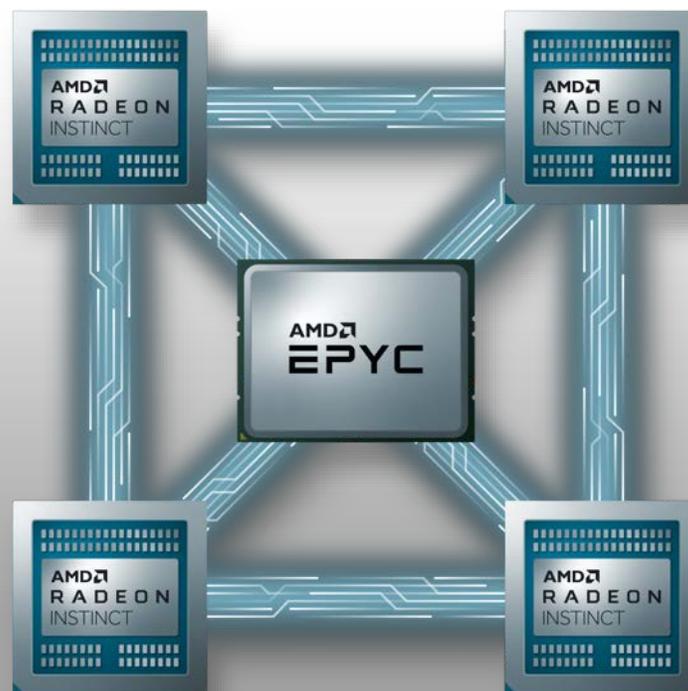


HIGH PERFORMANCE GPU OPTIMIZED FOR HPC AND AI

HPC-Customized
Compute Engines

Extensive Mixed Precision Ops for
Optimum Deep Learning Performance

High-Bandwidth Memory (HBM)
for Maximum Throughput



Infinity Fabric

High-Bandwidth, Low-Latency
Connection Between
CPU and GPU

Custom Coherent
Fabric

Connects 4:1
GPU to CPU Per Node

Shasta Blades, Cabinets & Slingshot Network



Frontier - System Summary

Hardware Element	Details
Peak Performance	> 1.5 ExaFlops
Footprint	> 100 cabinets
Node	1 HPC and AI Optimized AMD Future EPYC CPU 4 Purpose Built AMD Radeon Instinct GPU
CPU-GPU Interconnect	AMD Infinity Fabric Coherent memory across the node
System Interconnect	Multiple Slingshot NICs per node providing 100 GB/s network bandwidth Slingshot dragonfly network which provides adaptive routing, congestion management, and quality of service.
Storage	2-4x performance and capacity of Summit's I/O subsystem. Frontier will have near node storage.

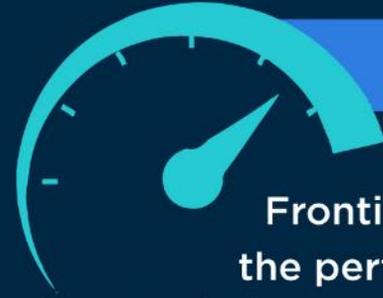
BANDWIDTH

The network bandwidth of the Frontier system is 24,000,000 times greater than the top home internet connection. With it, you could download



100,000
HD MOVIES IN
ONE SECOND

HORSEPOWER



Frontier will have
the performance of the
TOP 160 fastest supercomputers
in the world **COMBINED**

SIZE

Frontier will cover over
7,300 square feet.

That's
almost

2

**BASKETBALL
COURTS**

SPEED

If all 7.7 billion people on earth
each completed one calculation
per second, it would take over

6 YEARS

to do what the Frontier
system can do in

1 SECOND



CABLING

THE
90 MILES OF
CABLES
IN THE FRONTIER SYSTEM
would span the
distance from
Philadelphia to
New York City



QUESTIONS?

CRAY[®]
a Hewlett Packard Enterprise company

cray.com



[@cray_inc](https://twitter.com/cray_inc)



[linkedin.com/company/cray-inc/](https://www.linkedin.com/company/cray-inc/)

