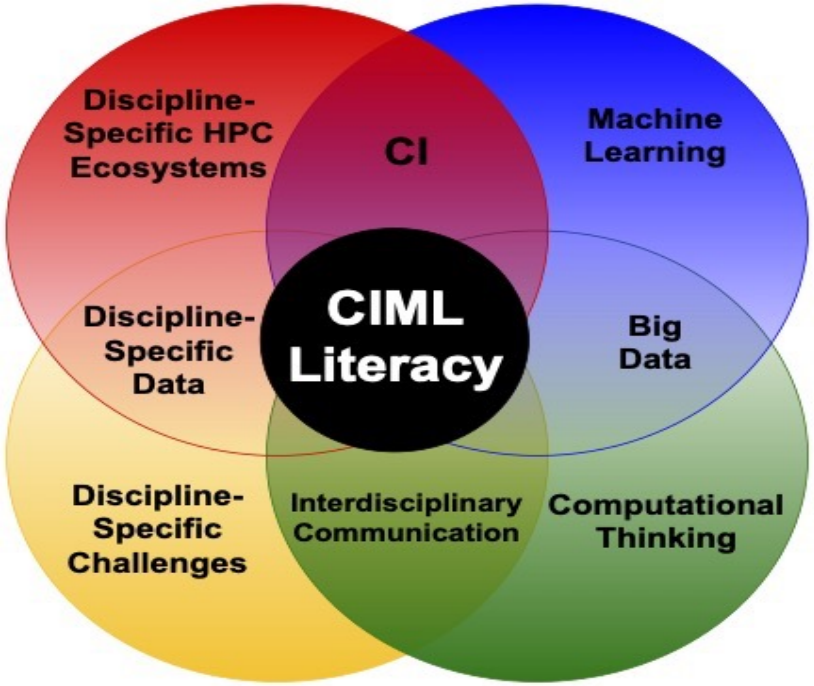
CIML Summer Institute:

CPU Computing - Hardware, Architecuture, Software Concepts

June 27, 2022



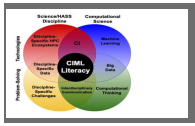
Mary Thomas
(SDSC)

NSF Award 1928224

SDSC SAN DIEGO SUPERCOMPUTER CENTER
UC San Diego

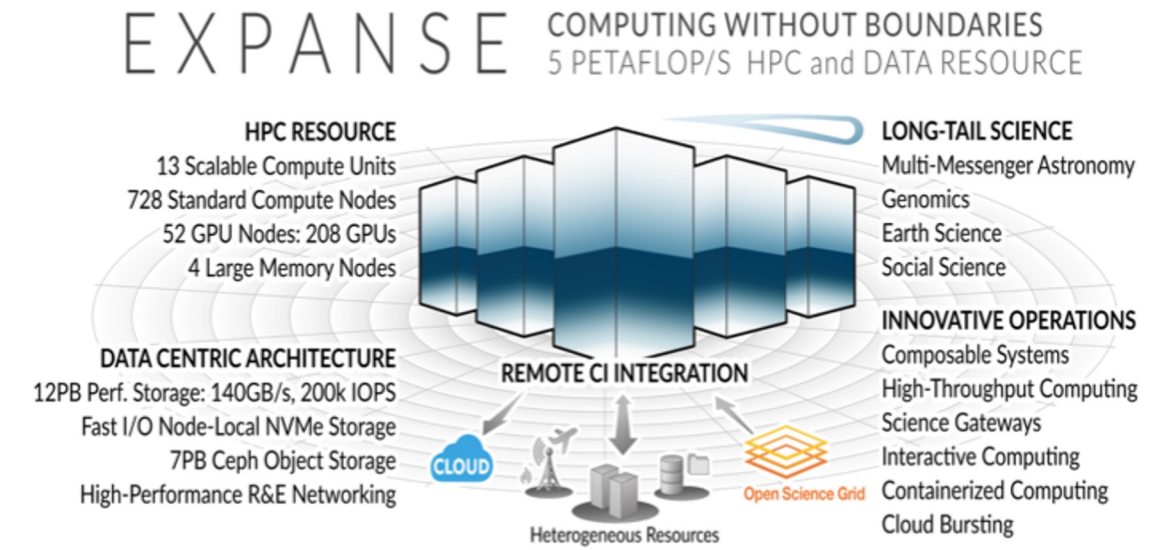
Outline

- Basic HPC/CI Concepts
- HPC/CI Hardware and Architecture: Expanse
- Final Comments



What is High Performance Computing (HPC) and Cyberinfrastructure (CI)?

- *HPC: Aggregating computing power*
- *CI: Connect resources with distribute emerging technologies*
- *Deliver much higher performance than desktop computer or workstation*
- *Solve large problems in science, engineering, or business.*



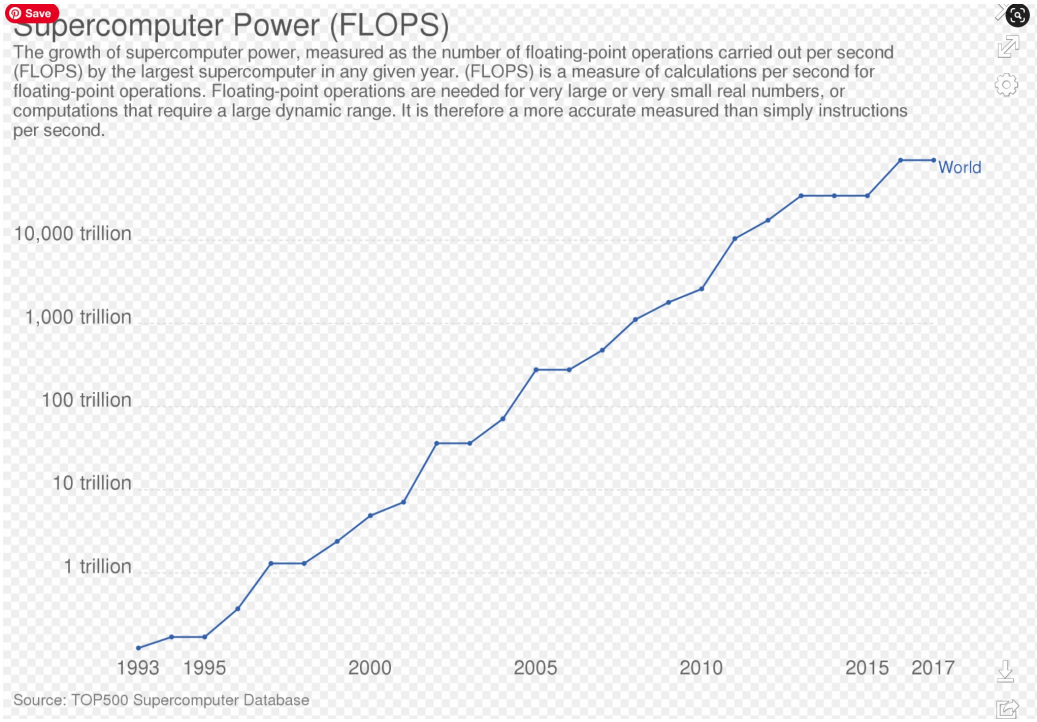
For more details see the Expanse user guide @ https://www.sdsc.edu/support/user_guides/expanse.html
and the "Introduction to Expanse" webinar @ https://www.sdsc.edu/event_items/202006_Introduction_to_Expanse.html

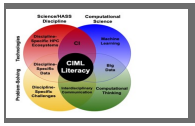
SDSC SAN DIEGO SUPERCOMPUTER CENTER UC San Diego

HPC Scaling: Currently at Exascale

Computer performance

Name	Unit	Value
kiloFLOPS	kFLOPS	10^3
megaFLOPS	MFLOPS	10^6
gigaFLOPS	GFLOPS	10^9
teraFLOPS	TFLOPS	10^{12}
petaFLOPS	PFLOPS	10^{15}
exaFLOPS	EFLOPS	10^{18}
zettaFLOPS	ZFLOPS	10^{21}
yottaFLOPS	YFLOPS	10^{24}



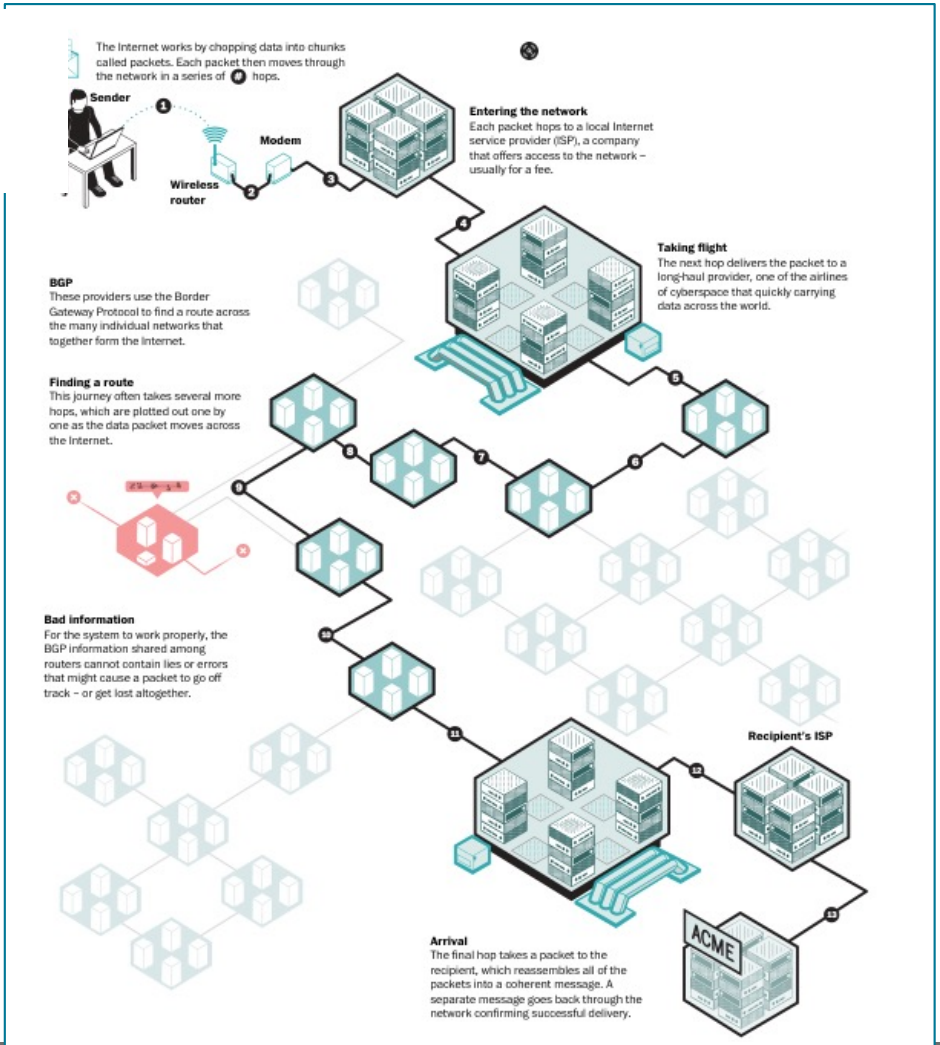


HPC/CI Is Everywhere: You use it daily!

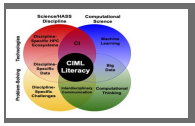
- The “Internet”
- Reaches 65% of the world (5.6 billion)
- 93% access via mobile technologies
- Data consumed globally in 2021: 79 zettabytes* (2021), projected to be 180 zettabytes by 2025

*Zettabyte=1,000,000,000,000,000,000,000 [10^{21}] bytes

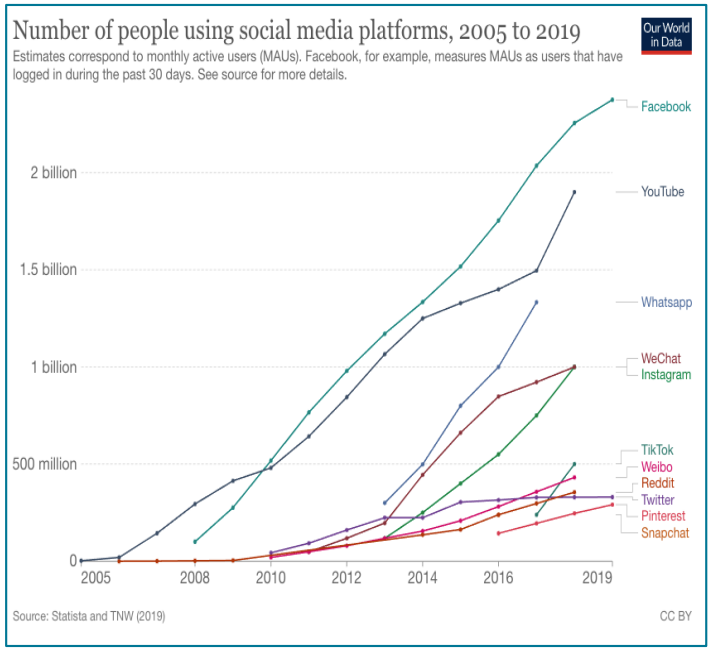
Source: https://www.visualcapitalist.com/from-amazon-to-zoom-what-happens-in-an-internet-minute-in-2021/



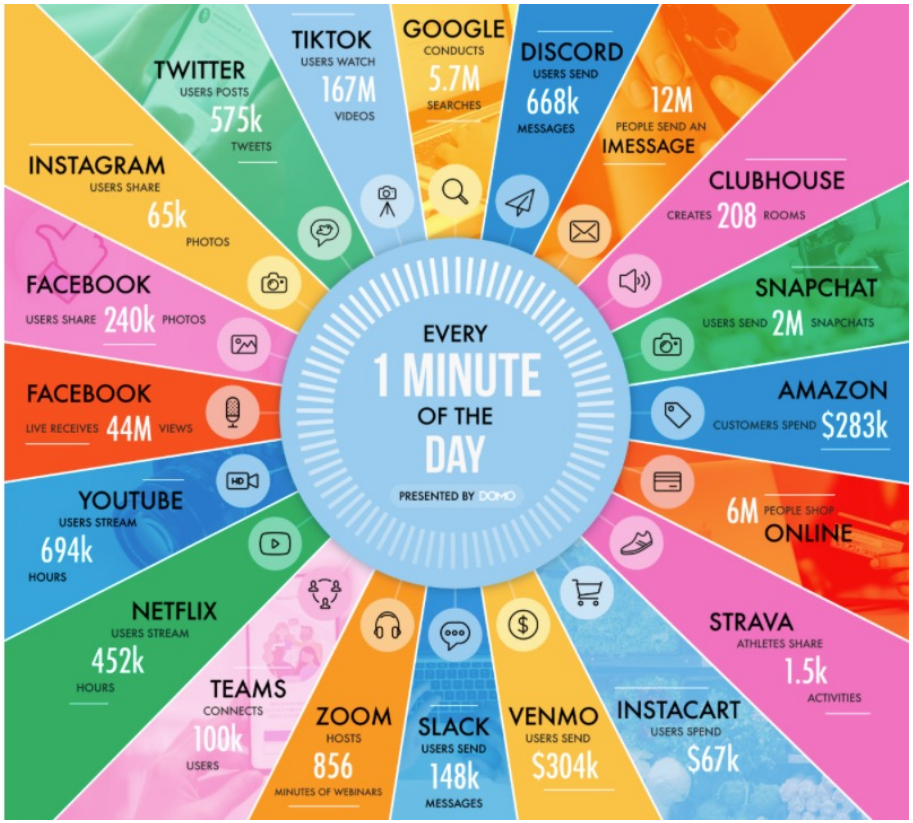
Src: https://www.washingtonpon.com/graphics/national/security-of-the-internet/bgp/



HPC/CI Is Everywhere: billions of users performing billions of transactions sending billions of data packets



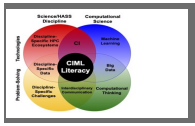
Source: https://ourworldindata.org/exports/users-by-social-media-platform_v17_850x600.svg



Source https://www.visualcapitalist.com/from-amazon-to-zoom-what-happens-in-an-internet-minute-in-2021/

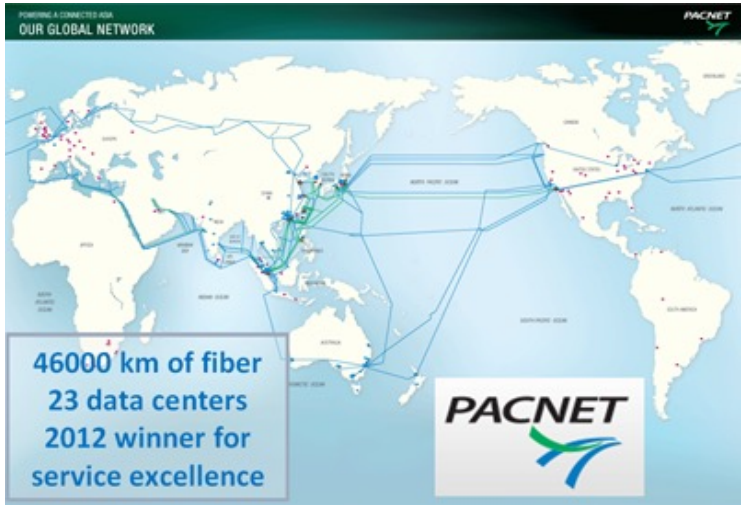
HPC/CI Is Everywhere:
Where do we house all this data?

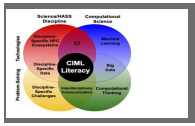
- Want to keep it “localized” for efficient access:
  - E.g. Luster file systems
- For very large data sets, global solution:
  - locate it where it is easy/cheap to keep cool (e.g. Iceland)
  - Current Idea: build skyscraper sized data storage facilities



HPC/CI Is Everywhere: How do we send all this data?

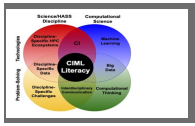
- Gigabit+ networks



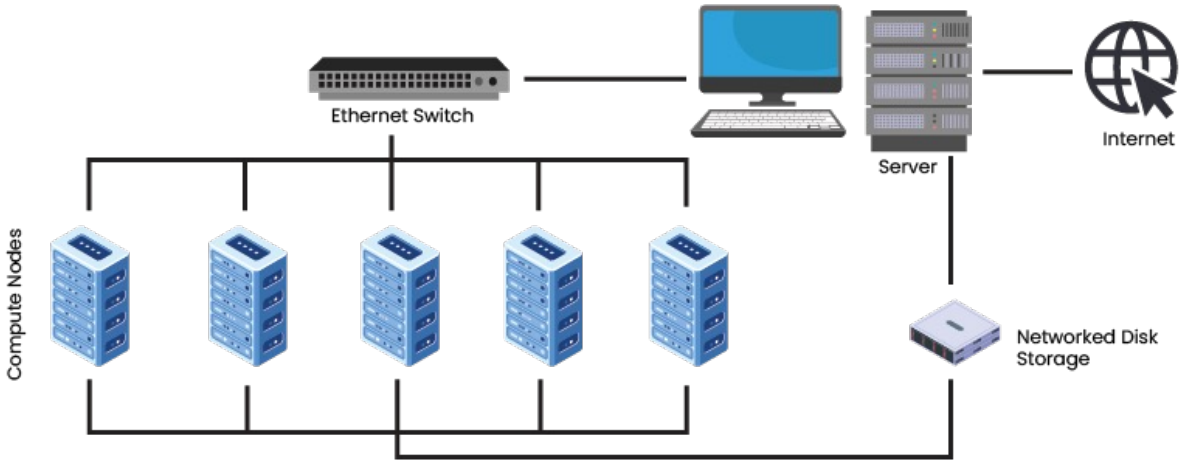


Outline

- Basic HPC/CI Concepts
- HPC/CI Hardware and Architecture: Expanse
- Final Comments



What is a Compute Cluster?



A set of computers (~10^3) that work together so that they can be viewed as a single system. Compute clusters typically have each node set to perform the same task, and are controlled and scheduled by special software.

Source: https://en.wikipedia.org/wiki/Computer_cluste

What is a Supercomputer Cluster?

A very large set of compute nodes (10^5 – 10^{6+}) that work together…

SDSC Expanse: 13+ racks; 728 AMD EPYC nodes, ~ 10^5 cores; 52 GPU nodes; > 12 PB

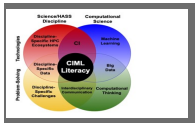
→ 5.2 petaflops.

(10^{15} floating-point operations per second.

ORNL Frontier: 100+ racks; 9000 AMD EPYC nodes, > 10^6 cores; > 32 PB

→ 1.5 exaflops

(10^{18} floating-point operations per second)



Early parallel computers based on commodity hardware

Stone Soupercomputer (2003):

Cheapest cost/flop=$0, ~20 MFlops

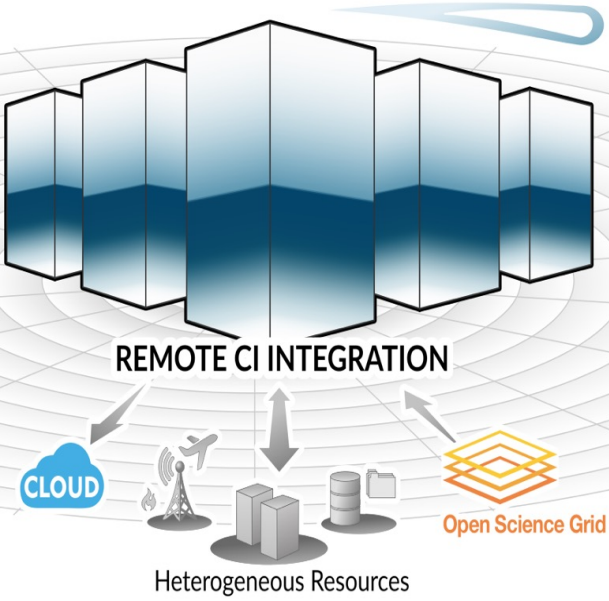
https://web.archive.org/web/20031121211117/http://stonesoup.esd.ornl.gov

EXPANSE

COMPUTING WITHOUT BOUNDARIES
5 PETAFLOP/S HPC and DATA RESOURCE

HPC RESOURCE
13 Scalable Compute Units
728 Standard Compute Nodes
52 GPU Nodes: 208 GPUs
4 Large Memory Nodes

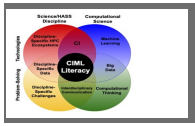
DATA CENTRIC ARCHITECTURE
12PB Perf. Storage: 140GB/s, 200k IOPS
Fast I/O Node-Local NVMe Storage
7PB Ceph Object Storage
High-Performance R&E Networking



LONG-TAIL SCIENCE
Multi-Messenger Astronomy
Genomics
Earth Science
Social Science

INNOVATIVE OPERATIONS
Composable Systems
High-Throughput Computing
Science Gateways
Interactive Computing
Containerized Computing
Cloud Bursting

For more details see the Expanse user guide @ https://www.sdsc.edu/support/user_guides/expanse.html
and the "Introduction to Expanse" webinar @ https://www.sdsc.edu/event_items/202006_Introduction_to_Expanse.html

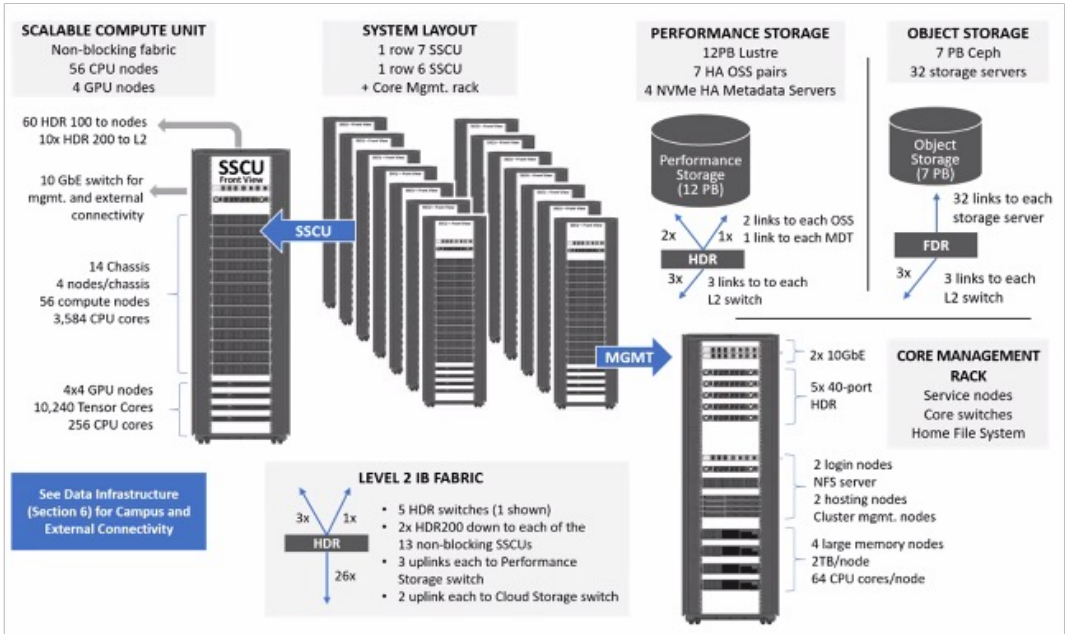


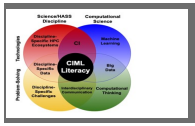
Expanse

Expanse Heterogeneous Architecture

System Summary

- *13 SDSC Scalable Compute Units (SSCU)*
- *728 x 2s Standard Compute Nodes*
- *93,184 Compute Cores*
- *200 TB DDR4 Memory*
- *52x 4-way GPU Nodes w/NVLINK*
- *208 V100s*
- *4x 2TB Large Memory Nodes*
- *HDR 100 non-blocking Fabric*
- *12 PB Lustre High Performance*
- *Storage*
- *7 PB Ceph Object Storage*
- *1.2 PB on-node NVMe*
- *Dell EMC PowerEdge*
- *Direct Liquid Cooled*





SCCU

The SSCU is Designed for the Long Tail Job Mix, Maximum Performance, Efficient Systems Support, and Efficient Power and Cooling

Standard Compute Nodes
- 2x AMD EPYC 7742 @2.25 GHz
- 128 Zen2 CPU cores
- PCIe Gen4
- 256 GB DDR4
- 1.6 TB NVME

GPU Nodes
- 4x NVIDIA V100/follow-on
- 10,240 Tensor Cores
- 32 GB GDDR
- 1.6 TB NVMe
- Intel CPUs

SSCU Components
- 56x CPU nodes
- 7,168 Compute Cores
- 4x GPU nodes
- 1x HDR Switch
- 1x 10GbE Switch
- HDR 100 non-blocking fabric
- Wide rack for serviceability
- Direct Liquid Cooling to CPU nodes

Non-blocking Interconnect

1 HDR Switch/SSCU

10x (200 Gbps)

HDR

56x → Compute Nodes

4x → GPU Nodes

5 Level 2 switches

Performance Storage

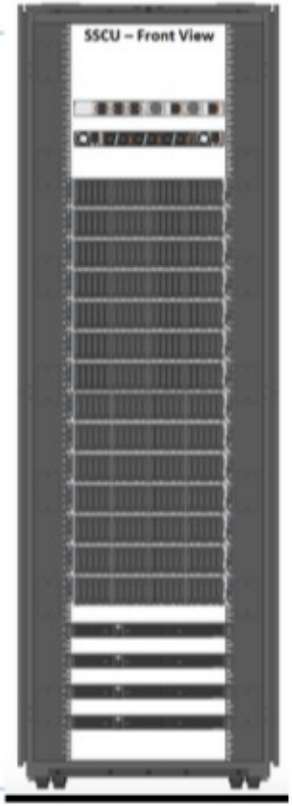
Cloud Storage

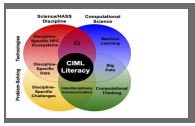
3x

2x

HDR

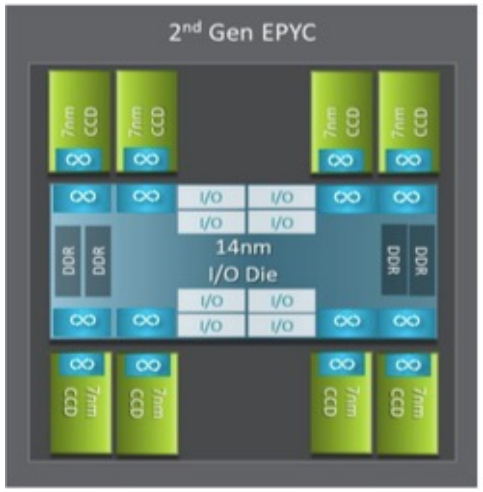
26x (200 Gbps)





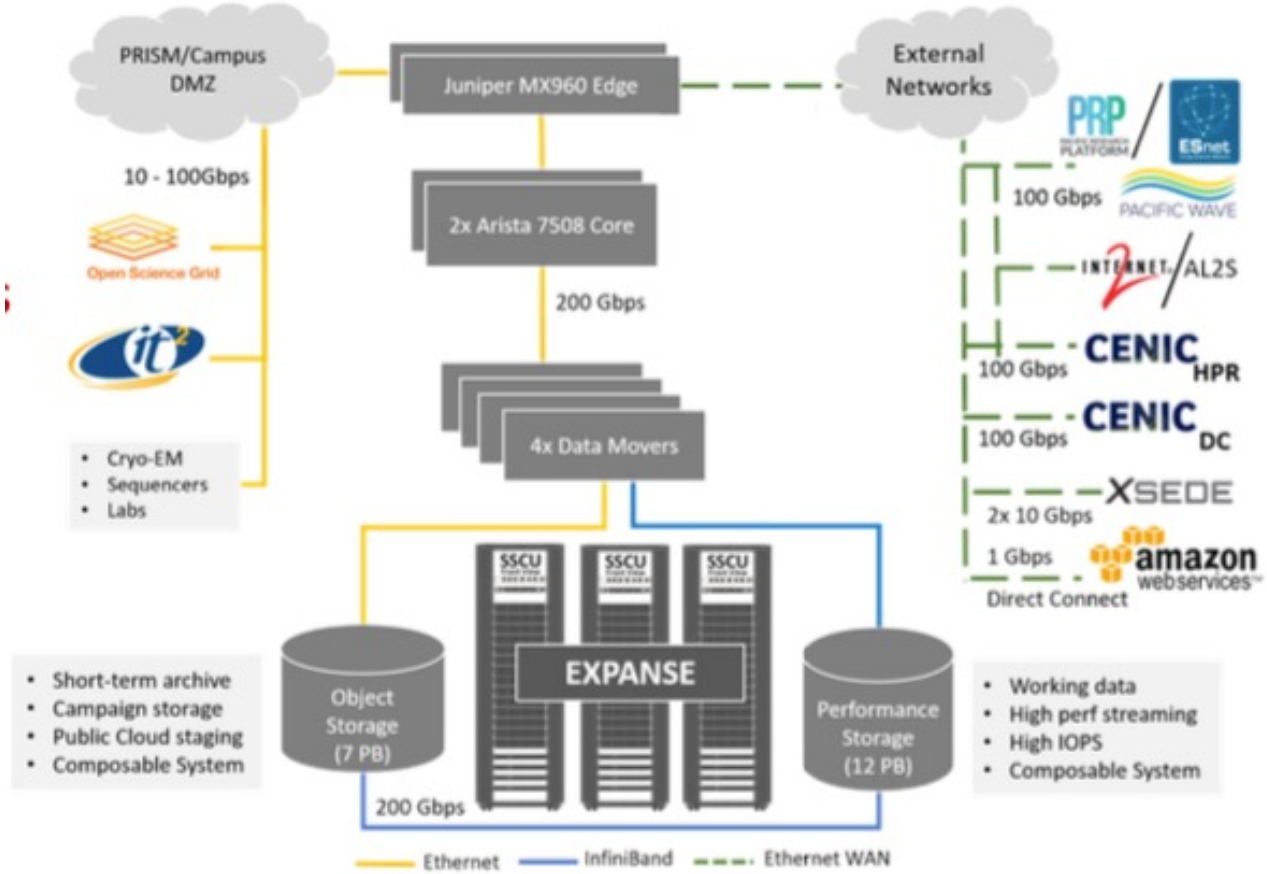
AMD EPYC 7742 Processor Architecture

- 8 Core Complex Dies (CCDs).
- CCDs connect to memory, I/O, and each other through the I/O Die.
- 8 memory channels per socket.
- DDR4 memory at 3200MHz.
- PCI Gen4, up to 128 lanes of high speed I/O.
- Memory and I/O can be abstracted into separate quadrants each with 2 DIMM channels and 32 I/O lanes.
- 2 Core Complexes (CCXs) per CCD
- 4 Zen2 cores in each CCX share a16ML3 cache. Total of 16x16= 256MB L3 cache.
- Each core includes a private 512KB L2 cache.



➔ EPYC Architecture has impact on compiling and batch script configuration

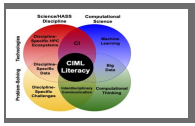
Expanse Connectivity Fabric



Facilitates Compute and Data Workflows

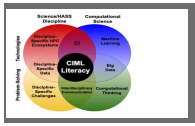
Resources

- Expanse User Guide
  - https://www.sdsc.edu/support/user_guides/expanse.html
- GitHub Repo for this webinar: clone code examples for this tutorial – clone example code:
  - https://github.com/sdsc-hpc-training-org/expanse-101
- SDSC Training Resources
  - https://www.sdsc.edu/education_and_training/training
  - https://github.com/sdsc-hpc-training/webinars
- XSEDE Training Resources



Resources (Optional)

- Expanse :
  - Landing page: _expanse.sdsc.edu
  - User Guide: https://expanse.sdsc.edu/support/user_guides/expanse.html
- Problems? Contact help@xsede.org



Thank You!

If you have problems, please contact help@xsede.org