Experiences and approaches for supporting Lustre for a large, national science and engineering user environment

Lustre User Group (LUG) Meeting
May 2-4, 2023

Mahidhar Tatineni, Christopher Irving, Subhashini Sivagnanam, Dmitry Mishin, Chris Cox, Yuwu Chen, Haisong Cai

San Diego Supercomputer Center (SDSC)

Outline

- Overview
- System architectures
  - Expanse
  - Comet
  - Triton Shared Computing Cluster (TSCC)
- Expanse Lustre IO performance
- Filesystem usage monitoring
- Application Examples
- Summary

Overview

- San Diego Supercomputer Center provides computational and data resources, services and expertise with three clusters
  - Expanse has AMD Rome processors (>100K cores), NVIDIA V100 GPUs, HDR InfiniBand(IB). NSF funded resource that supports a national research community with a broad range of applications
  - Comet has Intel Haswell processors, FDR IB. Originally NSF funded, has been repurposed to support researchers with the Center for Western Weather and Water Extremes (CW3E)
  - TSCC cluster provides advanced computing (CPU, GPU nodes) and data resources to serve the needs of the UCSD researchers.
- Lustre filesystems are an important component of the systems providing a scalable, high-performance, and open-source IO solution.

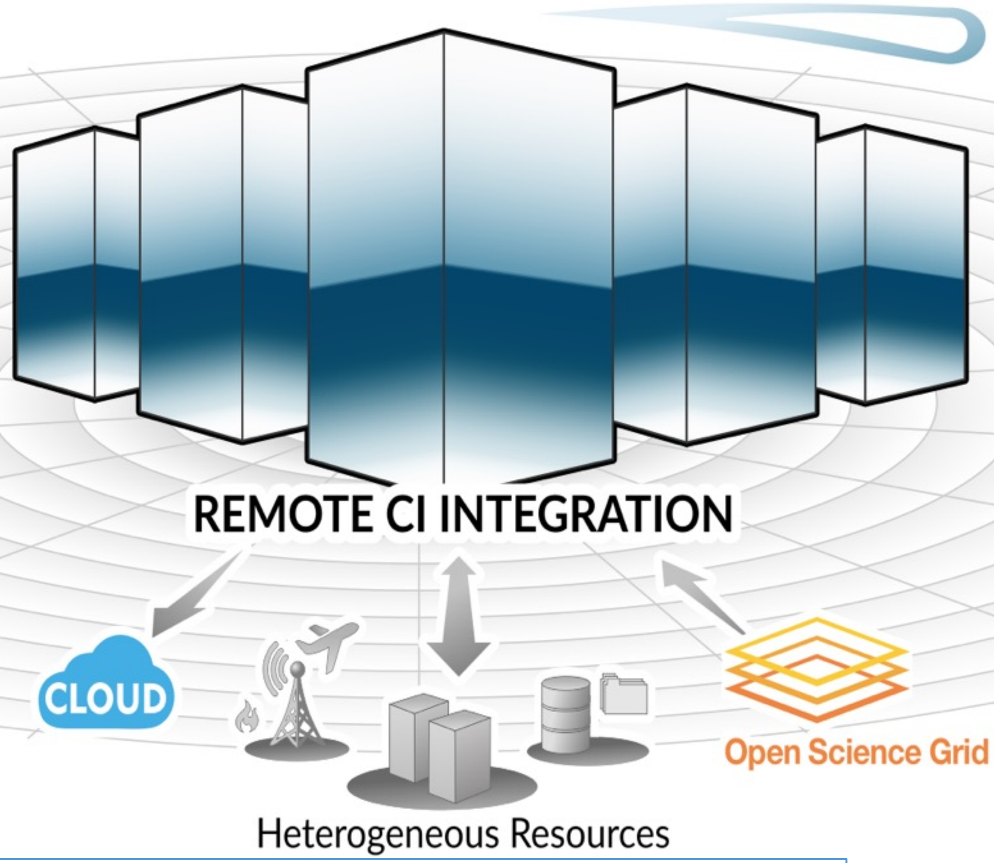
Outline

- Overview
- System architectures
    - Expanse
    - Comet
    - Triton Shared Computing Cluster (TSCC)
- Expanse Lustre IO performance
- Filesystem usage monitoring
- Application Examples
- Summary

EXPANSE
COMPUTING WITHOUT BOUNDARIES
5 PETAFLOP/S HPC and DATA RESOURCE

HPC RESOURCE
14 ~~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

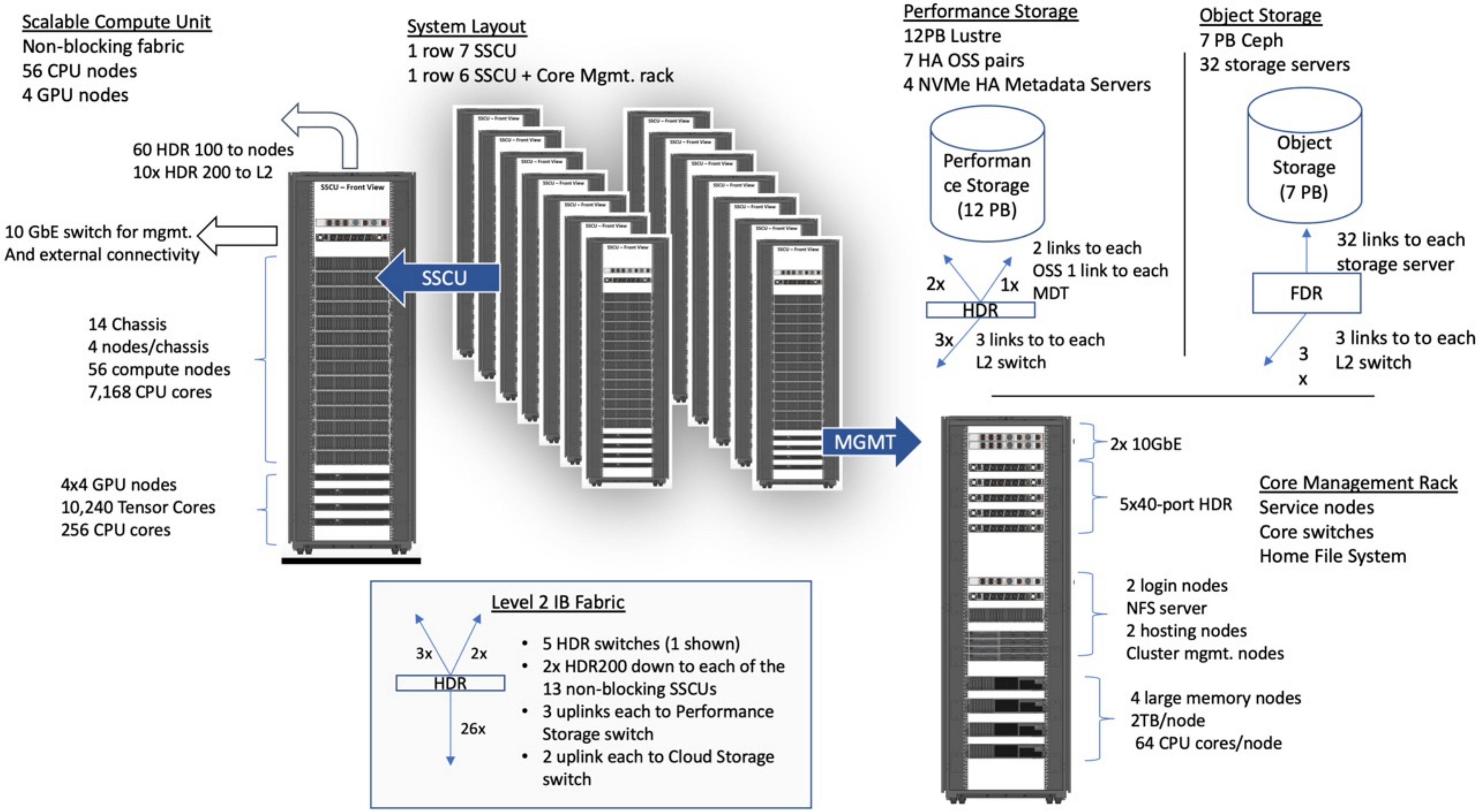
INNOVATIVE OPERATIONS
Composable Systems
High-Throughput Computing
Science Gateways
Interactive Computing
Containerized Computing
Cloud Bursting

Category 1: Capacity System, NSF Award # 1928224
PIs: Mike Norman (PI), Ilkay Altintas, Amit Majumdar, Mahidhar Tatineni, Shawn Strande

Expanse is a heterogeneous architecture designed for high performance, reliability, flexibility, and productivity

System Summary

- *14 SDSC Scalable Compute Units (SSCU)*
- *784 x 2s Standard Compute Nodes*
- *100,352 Compute Cores*
- *200 TB DDR4 Memory*
- *56x 4-way GPU Nodes w/NVLINK*
- *224 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*



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

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*

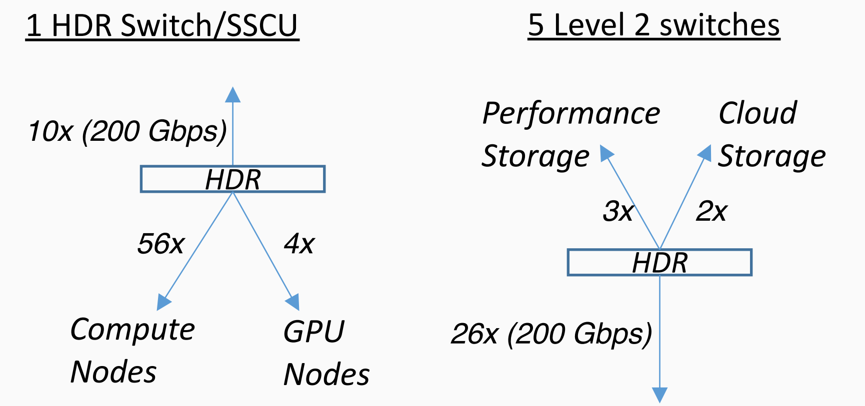
Standard Compute Nodes

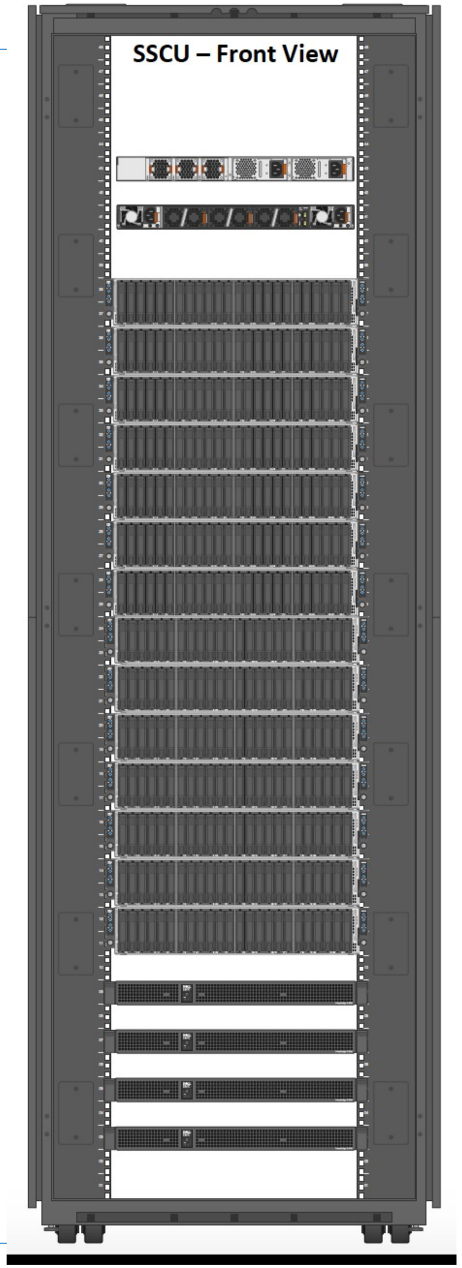
- *2x AMD EPYC 7742 @2.25 GHz*
- *128 Zen2 CPU cores*
- *PCIe Gen4*
- *256 GB DDR4*
- *1 TB NVME*

GPU Nodes

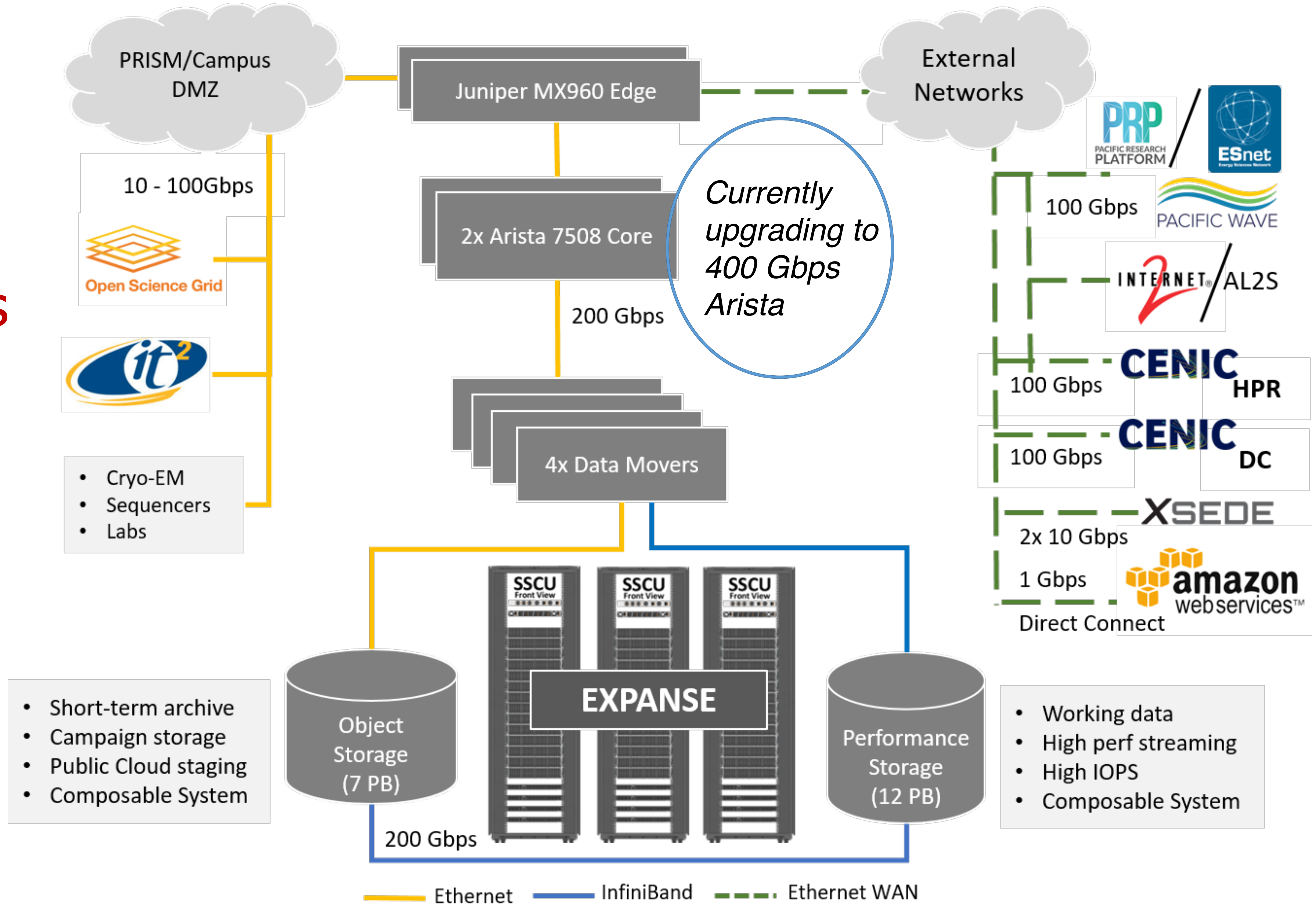
- *4x NVIDIA V100/follow-on*
- *10,240 Tensor Cores*
- *32 GB GDDR*
- *1.6 TB NVMe*
- *Intel CPUs*

Non-blocking Interconnect





Connectivity to R&E Networks Facilitates Compute and Data Workflows



Expanse Lustre Designed to Address Performance Bottlenecks

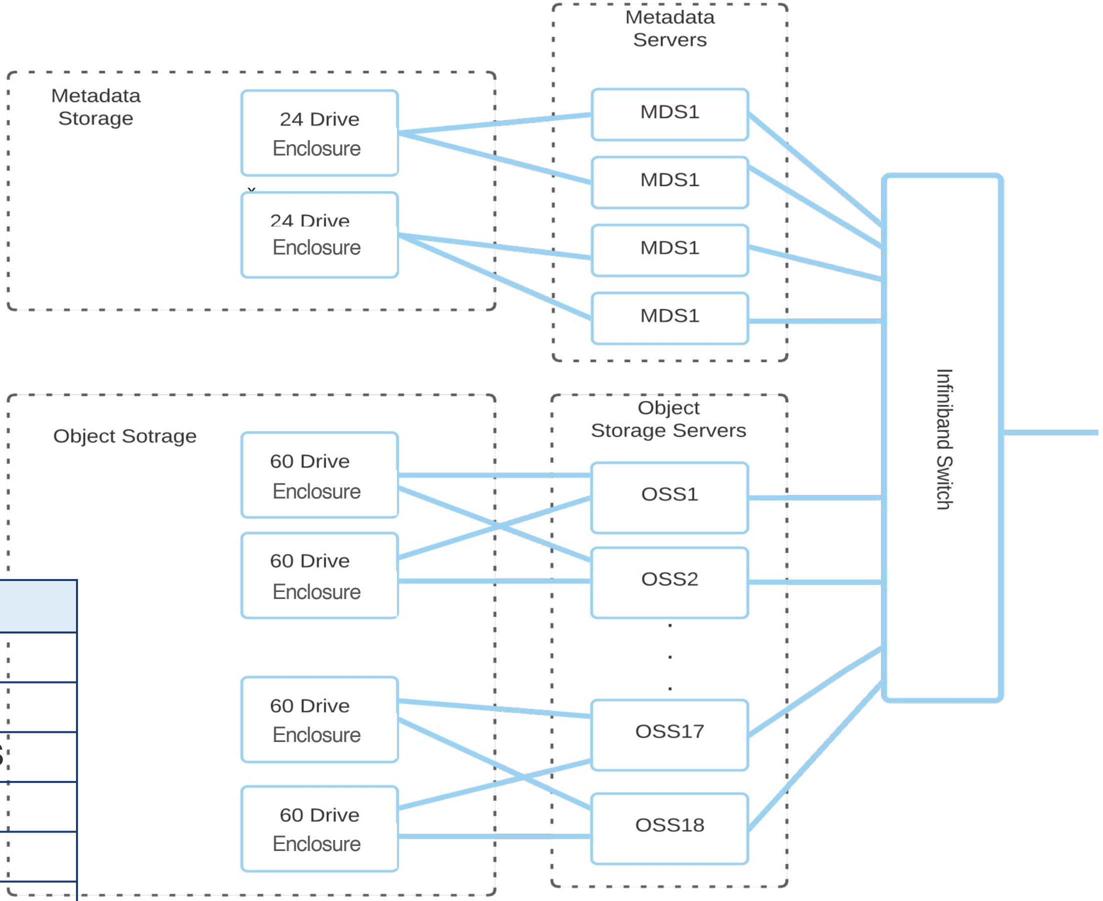
Small IO and Metadata operations are typical bottlenecks on HPC systems.

- Four All NVMe MDS
- Distributed Name Spaces : Improves metadata performance by load balancing operations across all four MDS
- Data on Metadata (DOM). DOM stores smaller files directly on MDS
  - Reduces Latency
  - Takes advantage of the low Latency NVMe drives on the MDSs
- Single Filesystem with two mount points instead two separate filesystems for scratch and projects.
  - Better parallel utilization of spindles
  - Low overhead to moving files when it's within the same filesystem.

Expanse Lustre Architecture

4 Lustre MDS	
Processor	2 X AMD Epyc 7302 (16 Cores)
Memory	512 GB (16 X 32GB DDR4 3200)
MDT Drives	24 X 3.8 TB NVMe per pair
Interconnect	InfiniBand HDR 200
System Drives	2 X 240 GB Intel SSDs

- Lustre Hardware Provided by Aeon Computing
- 12 Peta Bytes of RAW capacity, approx. 11 PB formatted
- File Capacity of approx. 3 billion files.
- 140 GB/s Filesystem Bandwidth
- 200K IOPS
- DNE Phase III, Striped across the four MDTs, NVMe
- 18 OSS (72 OSTs)
- Data on MDT (DoM) for small file performance
- Lustre v2.13.0 w/ zfs 0.7.13



18 Lustre OSS	
Processor	1 AMD Epyc 7402 (24 Cores)
Memory	512 GB (16 X 32 GB DDR4 3200)
JBODS	2 Cross Connected 60 Bay JBODS
OSS Drives	120 X 14 TB 7200 SAS Drives
Interconnect	InfiniBand HDR 200
System Drives	2 X 240 GB Intel SSDs

Comet: System Characteristics

- Total peak flops ~2.1 PF
- Dell primary integrator
  - Intel Haswell processors w/ AVX2
  - Mellanox FDR InfiniBand
- 1,944 standard compute nodes (46,656 cores)
  - Dual CPUs, each 12-core, 2.5 GHz
  - 128 GB DDR4 2133 MHz DRAM
  - 2*160GB GB SSDs (local disk)
- 72 GPU nodes
  - 36 nodes same as standard nodes *plus* Two NVIDIA K80 cards, each with dual Kepler3 GPUs
  - 36 nodes with 2 14-core Intel Broadwell CPUs plus 4 NVIDIA P100 GPUs

4 large-memory nodes

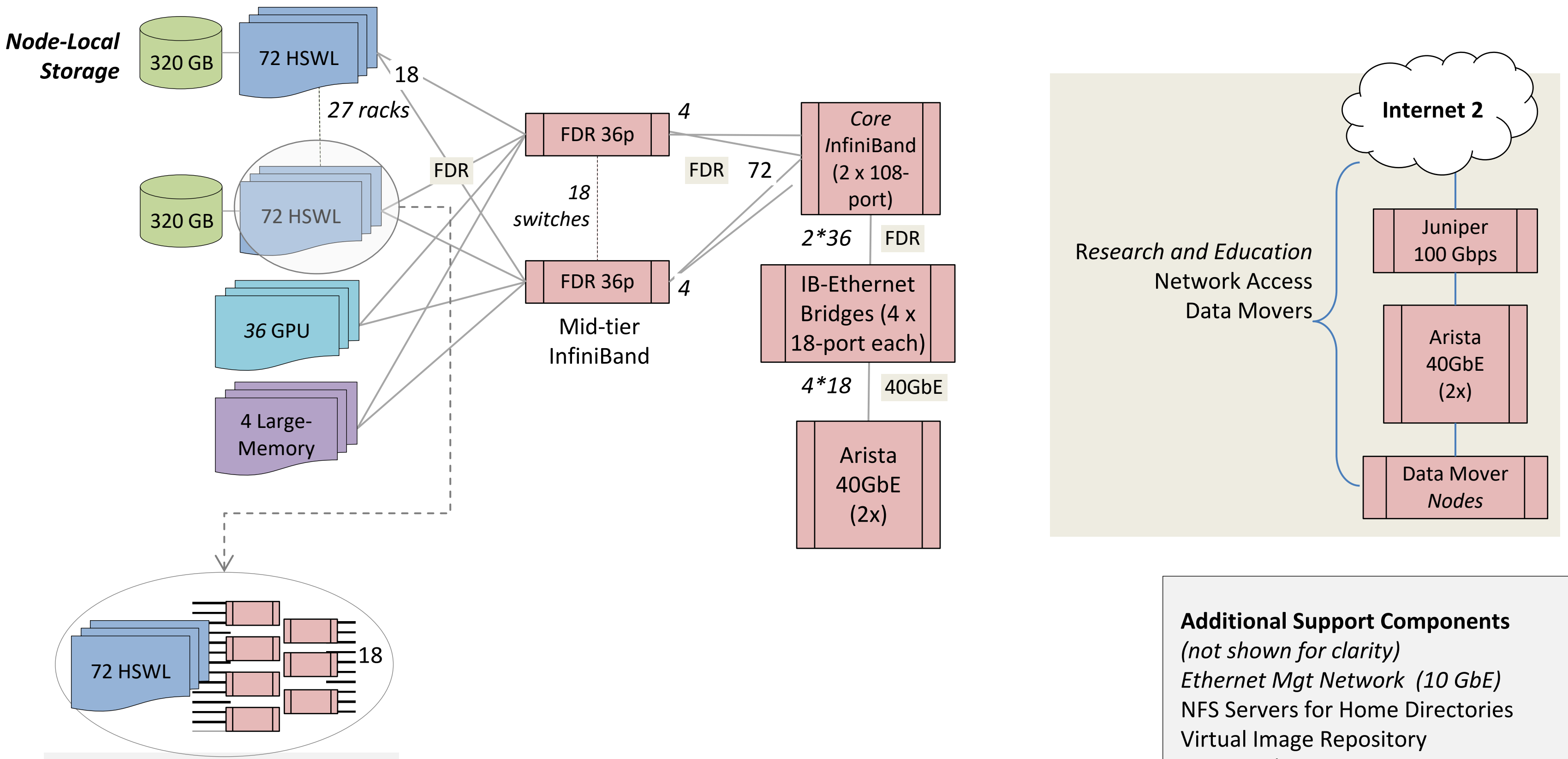
1.5 TB DDR4 1866 MHz DRAM

Four Haswell processors/node

64 cores/node

- *Hybrid fat-tree topology w/ FDR InfiniBand*
  - *Rack-level (72 nodes, 1,728 cores) full bisection bandwidth*
  - *4:1 oversubscription cross-rack*
- *Performance Storage (Aeon)*
  - *7.6 PB, 200 GB/s; Lustre*
  - *Scratch & Persistent Storage segments*
- *Durable Storage (Aeon)*
  - *6 PB, 100 GB/s; Lustre*

Comet Network Architecture



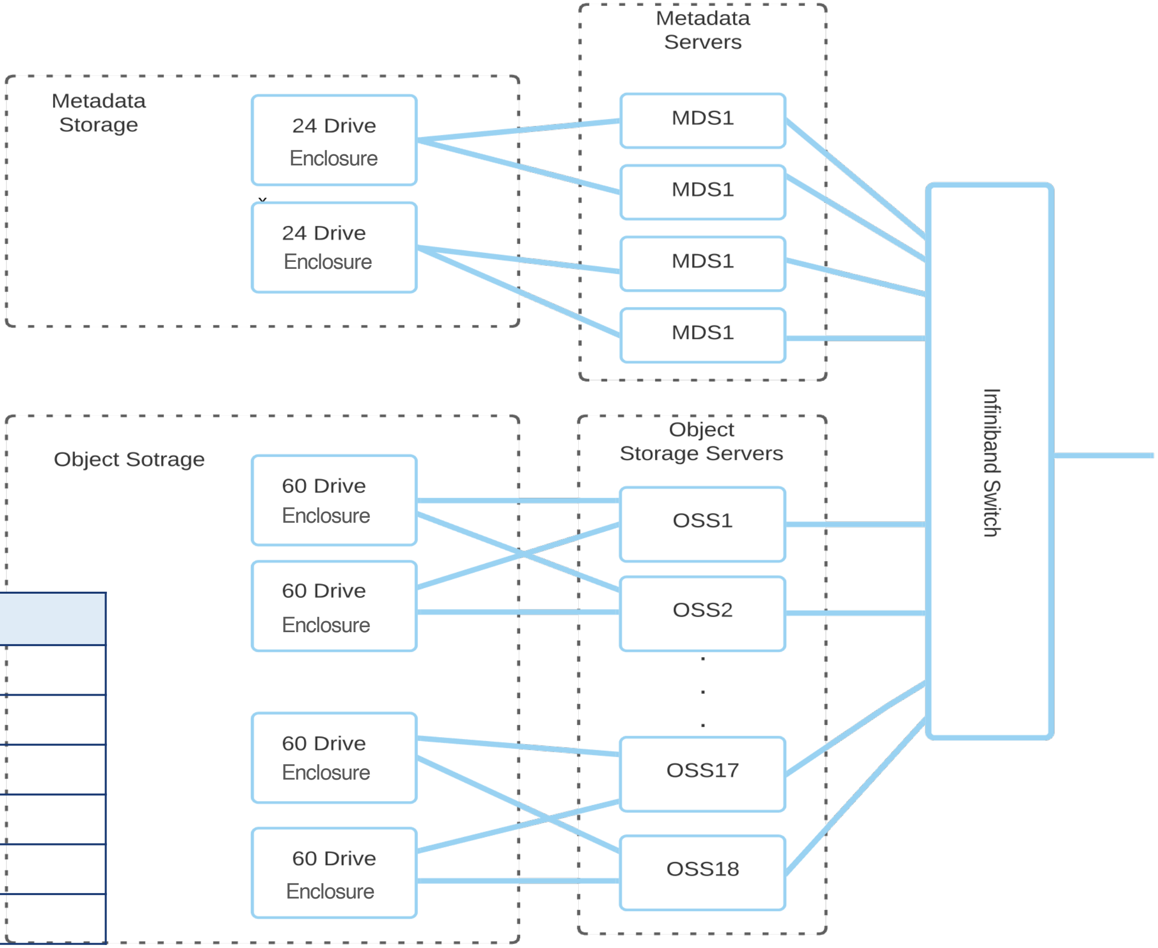
7x 36-port FDR in each rack wired as full fat-tree. 4:1 over subscription between racks.

Additional Support Components
(not shown for clarity)
Ethernet Mgt Network (10 GbE)
NFS Servers for Home Directories
Virtual Image Repository
Gateway/Portal Hosting Nodes
Login Nodes
Rocks Management Nodes

Comet Lustre Architecture

4 Lustre MDS	
Processor	2 X AMD Epyc 7302 (16 Cores)
Memory	512 GB (16 X 32GB DDR4 3200)
MDT Drives	24 X 3.8 TB NVMe per pair
Interconnect	InfiniBand HDR 200
System Drives	2 X 240 GB Intel SSDs

- Lustre Hardware Provided by Aeon Computing
- 12 Peta Bytes of RAW capacity, approx. 11 PB formatted
- File Capacity of approx. 3 billion files.
- 140 GB/s Filesystem Bandwidth
- 2MDS + SSD based + DNE Phase I
- 18 OSSs (68 OSTs)*
- Data on MDT (DoM) for small file performance
- **Connected into director switch**
- Lustre v2.14.0, ZFS 2.1.0



18 Lustre OSS	
Processor	1 AMD Epyc 7402 (24 Cores)
Memory	512 GB (16 X 32 GB DDR4 3200)
JBODS	2 Cross Connected 60 Bay JBODS
OSS Drives	120 X 14 TB 7200 SAS Drives
Interconnect	InfiniBand HDR 200
System Drives	2 X 240 GB Intel SSDs

** 2 OSS nodes only populated with half chassis drives*

TSCC Cluster

- The Triton Shared Computing Cluster (TSCC) is the primary research High-Performance Computing (HPC) system for UC San Diego faculty and researchers.
- Several types of nodes with range of x86_64 CPUs and NVIDIA GPUs.
- TSCC nodes can have a single port 10GbE network interface. Subset of nodes have an optional EDR InfiniBand/FDR Infiniband (IB) connect for low latency parallel applications.
- 800+ TB Data Oasis Lustre-based high performance parallel file system primarily used for scratch space. There is a 90-day purge policy on this filesystem. Lustre v2.10.7, ZFS 0.7.9, TCP/IP network
- In addition, a dedicated DDN Lustre (2.34PB storage capacity) for processing dbGaP data for cancer research is mounted on a subset of nodes.

Outline

- Overview
- System architectures
  - Expanse
  - Comet
  - Triton Shared Computing Cluster (TSCC)
- Expanse Lustre IO performance
- Filesystem usage monitoring
- Application Examples
- Summary

Expanse Lustre filesystem performance

- Acceptance tests included bandwidth and metadata requirements.
- Bandwidth Benchmarked using IOR with 18 nodes, 8 tasks per client (total of 144 tasks). The aggregate size of the files written was ~27 TB with a transfer size of 1MB.
- Bandwidth Result: **Max Write: 143098.86 MiB/sec (150050.03 MB/sec)**
- The mdtest benchmark used to evaluate the metadata performance on the Performance Storage filesystem
- The Lustre filesystem is configured for storing data on metadata servers for file sizes < 64KB. A total of 1,048,576 files were written and read in each iteration, all of which were written to the metadata server.
- Metadata performance result: **Max IOPs on file reads: 211369.705**

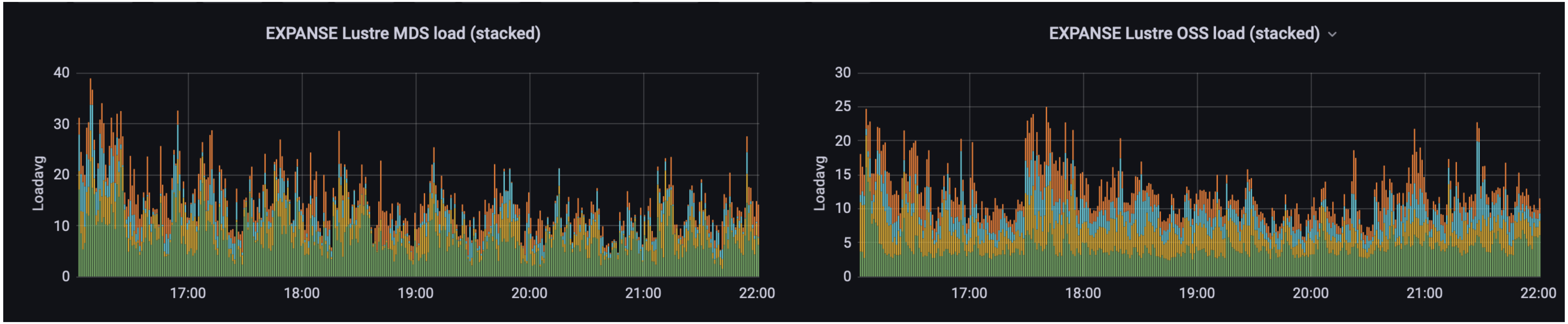
Outline

- Overview
- System architectures
  - Expanse
  - Comet
  - Triton Shared Computing Cluster (TSCC)
- Expanse Lustre IO performance
- Filesystem usage monitoring
- Application Examples
- Summary

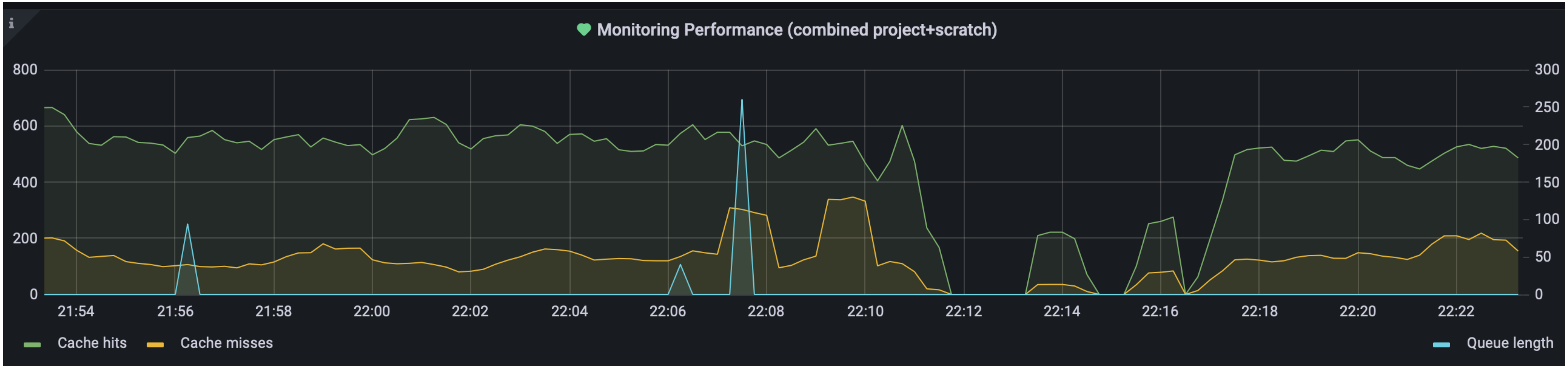
Expanse Usage Monitoring

- Two monitoring components on Expanse Lustre:
  - MDS Changelog: A service node interfaces with the MDS changelog facilities (enabled on MDS nodes). A process called pdm (written in go by SDSC developer) is directly collecting data from MDS change log and is being periodically scraped by prometheus.
  - Telegraf: An agent running on OSS/MDS nodes and forward stats directly to Graphite via proxy.
- The statistics from both monitoring components are forwarded to a Graphite instance with a Grafana frontend
  - Used for monitoring system health
  - User services staff can identify anomalous workloads and work with users to optimize as necessary

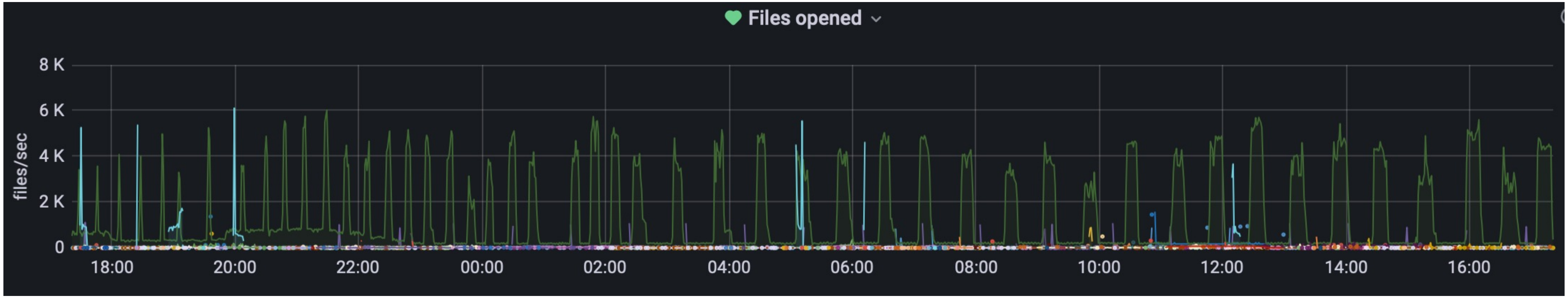
Expanse Lustre Monitoring

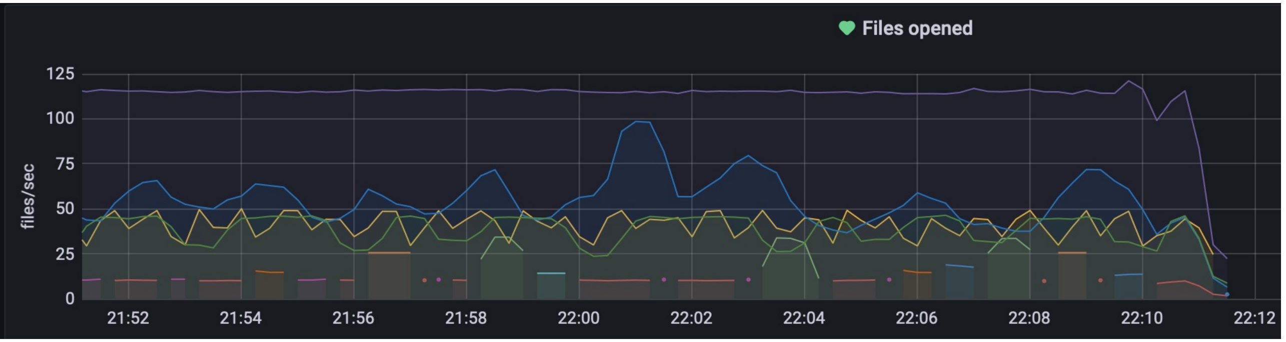


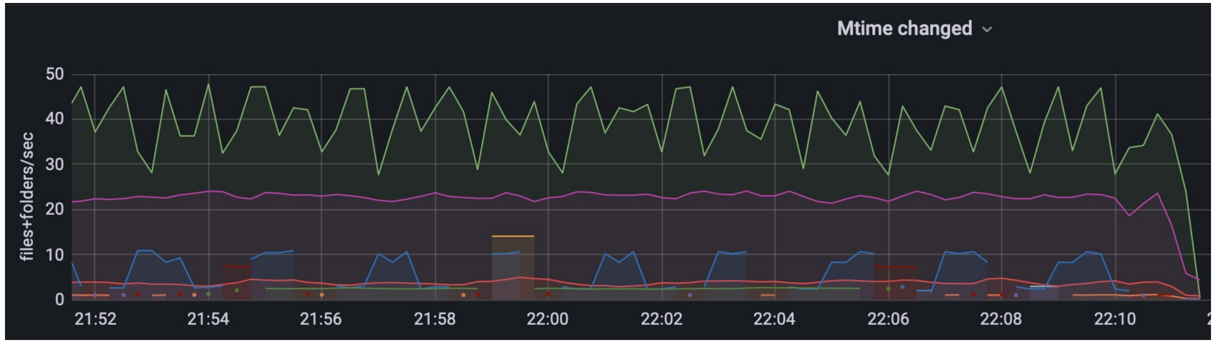
Expanse Lustre Monitoring



Expanse Usage Monitoring - User Info







Typical Issues w/ Lustre

- Sustained metadata intensive activity - some users have workloads that generate a huge number of scratch files (with reads and writes) at rapid rates.
  - Lustre handles them for short durations (a few hours) but sustained loads for days have caused MDS server problems.
  - Typically solved with combination of NVMe local scratch and Lustre. Scratch files redirected to NVMe. If needed aggregated to large tar files in Lustre
- IOPs heavy reads, e.g. Anaconda installs in Lustre, flat file databases.
  - On Comet, system admins have redirected anaconda installs to Data on MDT (DoM) locations and it has sped up such reads
  - Also used Qumulo filesystem mounts for some of these installs
- Some issues specific to Lustre + ZFS - usually upgrading/patching Lustre, ZFS has helped.

Outline

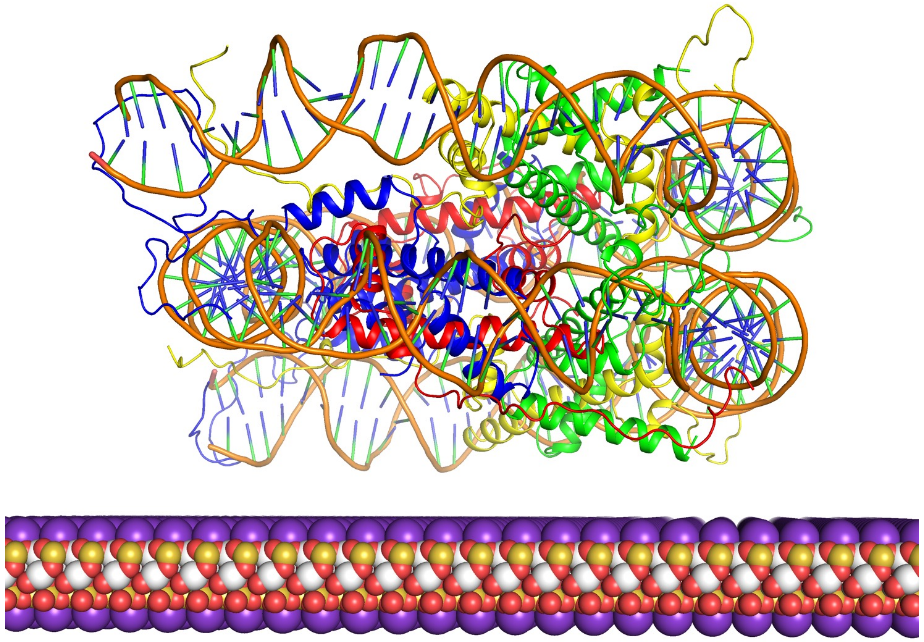
- Overview
- System architectures
  - Expanse
  - Comet
  - Triton Shared Computing Cluster (TSCC)
- Expanse Lustre IO performance
- Filesystem usage monitoring
- Application Examples
- Summary

Expanse Applications

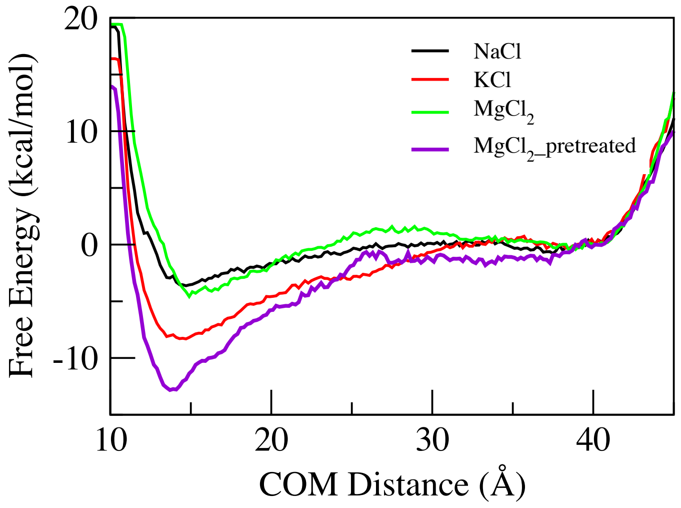
- Expanse supports a wide range of applications in bioinformatics, computational chemistry, cosmology, engineering, weather prediction codes, and many other scientific domains.
- Lustre usage over 3.4PB and 713 million files.
- Varying IO patterns
  - High throughput in jobs with continuous IO from each job
  - Large scale parallel IO
  - File per task type IO. File sizes can be large (e.g cosmology codes) or very small (e.g. OpenFOAM application).
- Nearly 3,000 researchers (not including gateway users) and educators from over 400 institutions and spanning 133 science, engineering and social sciences disciplines have Expanse accounts. A few examples in following slides.

Molecular mechanisms of epigenetics

The Wereszczynski lab at the **Illinois Institute of Technology** uses Expanse to study epigenetic mechanisms at the molecular level. This involves large-scale atomistic MD simulations of protein/DNA complexes to examine the relationship between their structure, dynamics, and function, and how nature tunes these to affect gene expression.

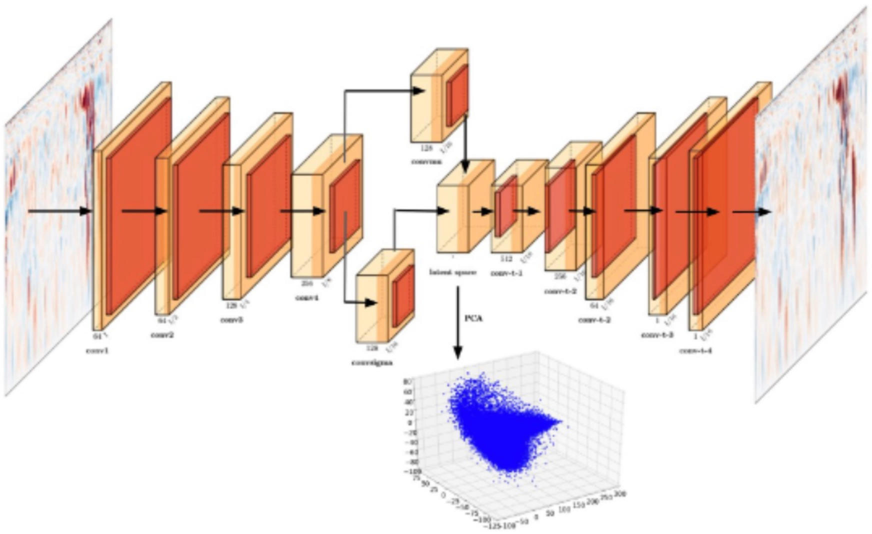


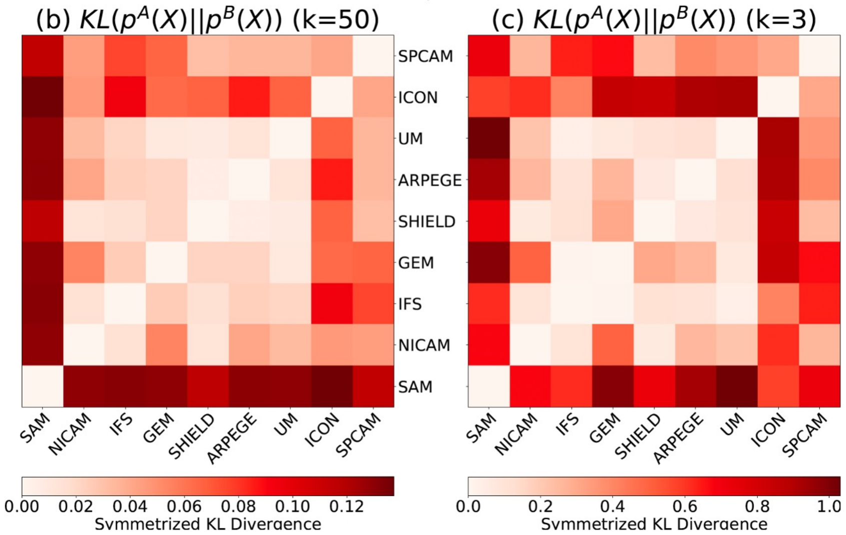
Left: the nucleosome core particle interacting with a mica surface, a commonly used experimental setup.
Bottom: Free energy profiles of DNA binding to mica surfaces in varied ionic environments.



Data-driven intercomparison of global storm-resolving climate models

The Pritchard Group (**UC Irvine**) uses *Expanse's* GPUs for the novel analysis of high resolution climate simulation data. Large volumes of output from multiple climate models serve as training data for Variational Autoencoders (VAEs), which can then construct low-dimensional latent representations of the data.



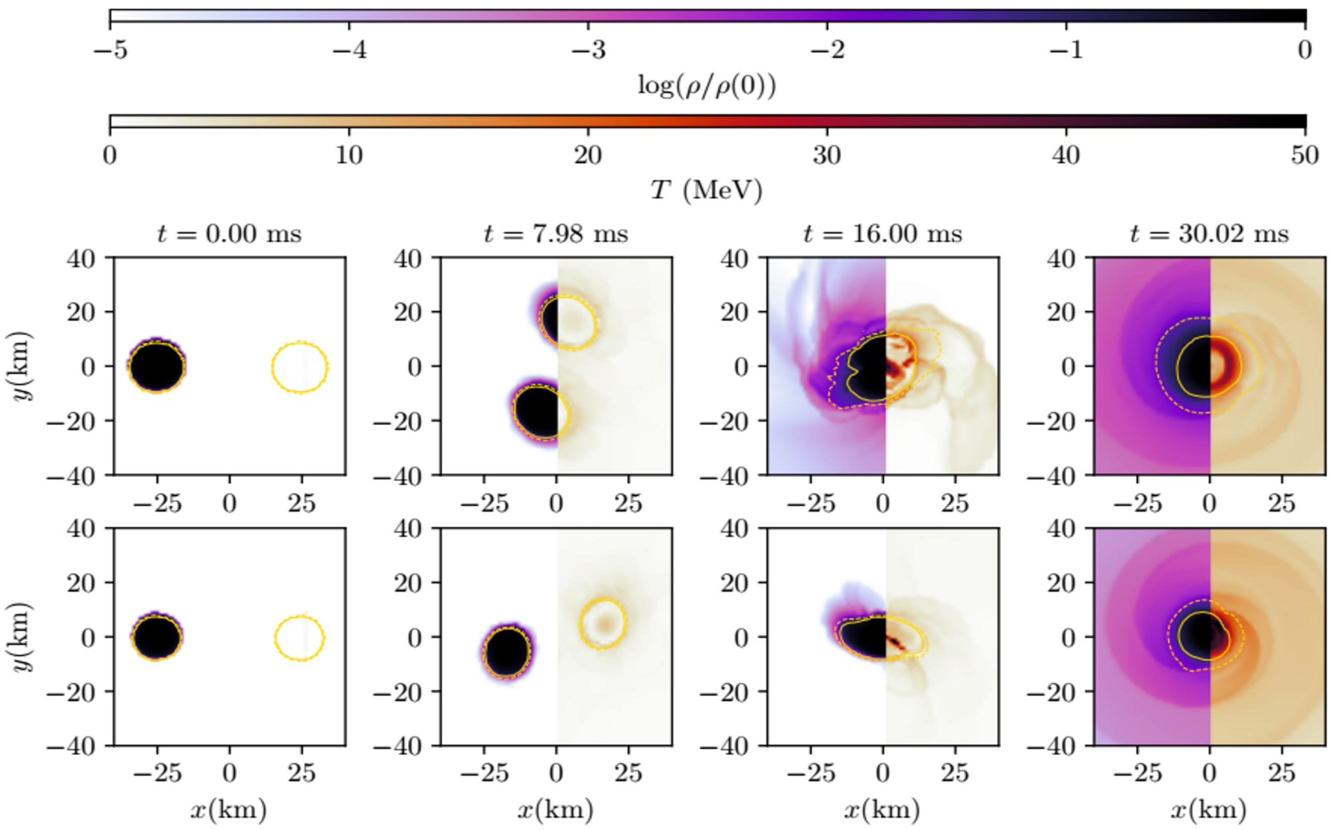


In the lower dimensional latent form, one can apply other machine learning techniques on these model outputs to elucidate the atmospheric characteristics of each simulation and measure "Distribution Shifts" between different climate models. Only six of the nine models investigated are dynamically similar (light reds) in their representation of the atmosphere, while three (SAM, SPCAM, ICON) show disparate (dark reds) atmospheric conditions.

Numerical simulations of neutron star mergers

The team led by David Radice (**Penn State**) is performing numerical relativity simulations of black holes to study their gravitational waves and electromagnetic signatures.

- Discovered that the threshold mass for prompt black hole formation in neutron star mergers is correlated with the compressibility of neutron star matter
- First long-term merger and post-merger simulations with sophisticated neutrino transport
- Developed a new code infrastructure for binary black hole merger simulations optimized for new computing architectures



Equatorial snapshots of binary neutron star merger simulations. We depict the rest mass density (top color bar) and temperature (lower color bar) on the left and right half of each panel, respectively. From Espino et al., in prep.

Comet Applications: CW3E Users

- Petascale Computing to Support Research and Prediction for Center for Western Weather and Water Extremes (CW3E)

  https://cw3e.ucsd.edu

- Computations include Ocean to Atmospheric Rivers, Precipitation, Hydrology, S2S, and Use of Community Models
- Key project involves 200-member Operational Forecast Ensemble
  - Runs every year during wet season
  - A 9-km 200-member West-WRF ensemble to provide near real-time forecast
  - Predict the timing and magnitude of extreme events associated with atmospheric rivers
  - Extremely IO intensive with nearly 1.9 PB of data

Summary

- Lustre filesystems are key components of the IO subsystems on three major clusters - Expanse, Comet, and TSCC.
- Network architectures vary on the 3 systems. Expanse and Comet feature InfiniBand networks and the TSCC cluster is primarily ethernet based.
- Expanse and Comet Lustre filesystems feature data on MDTs. Expanse delivers > 140GB/s on bandwidth and > 200K IOPs.
- Monitoring infrastructure in place with Grafana frontend. Helps system administrators monitor filesystem health, and user services teams monitor IO behavior at user and group levels.
- The machines support a variety of application workloads in bioinformatics, computational chemistry, cosmology, engineering, weather prediction codes, and many other scientific domains.

Thank you to our collaborators, partners, users, and the SDSC team!

NSF
DELL
AMD
NVIDIA
intel
XSEDE
Extreme Science and Engineering Discovery Environment
AEON COMPUTING
Mellanox TECHNOLOGIES

Ilkay Altintas
Haisong Cai
Amit Chourasia
Trevor Cooper
Jerry Greenberg
Eva Hocks
Tom Hutton
Christopher Irving
Marty Kandes
Amit Majumdar
Dima Mishin
Sonia Nayak
Mike Norman
Wayne Pfeiffer
Scott Sakai
Fernando Silva
Bob Sinkovits
Subha Sivagnanam
Michele Strong
Shawn Strande
Mahidhar Tatineni
Mary Thomas
Nicole Wolter
Frank Wuerthwein

EXPANSE
COMPUTING WITHOUT BOUNDARIES