aixcelerate

Does I/O Matter to Me?

Tuning Workshop 2025

December 2nd, 2025 | Aachen

Philipp Martin

NHR4CES NHR for Computational Engineering Science | HPC.NRW | EDIH Rheinland | WEST AI KI-Servicezentrum | IT Center | RWTH AACHEN UNIVERSITY

Agenda for Today (Dec 2nd, 2025)

Time	Topic	Speaker
9:00 – 9:05	Welcome	Sandra Wienke (RWTH)
9:05 – 10:00	Does I/O Matter to Me?	Philipp Martin (RWTH)
10:00 – 10:30	Storage and I/O Technologies on CLAIX	Philipp Martin (RWTH)
10:30 – 11:00	Coffee Break	
11:00 – 12:00	Tips & Tricks & Tools with respect to I/O	Philipp Martin (RWTH)
12:00 – 12:30	From Thousands of Small Files to Few Very Big Files	Jannis Klinkenberg (RWTH)

Contents

Contents

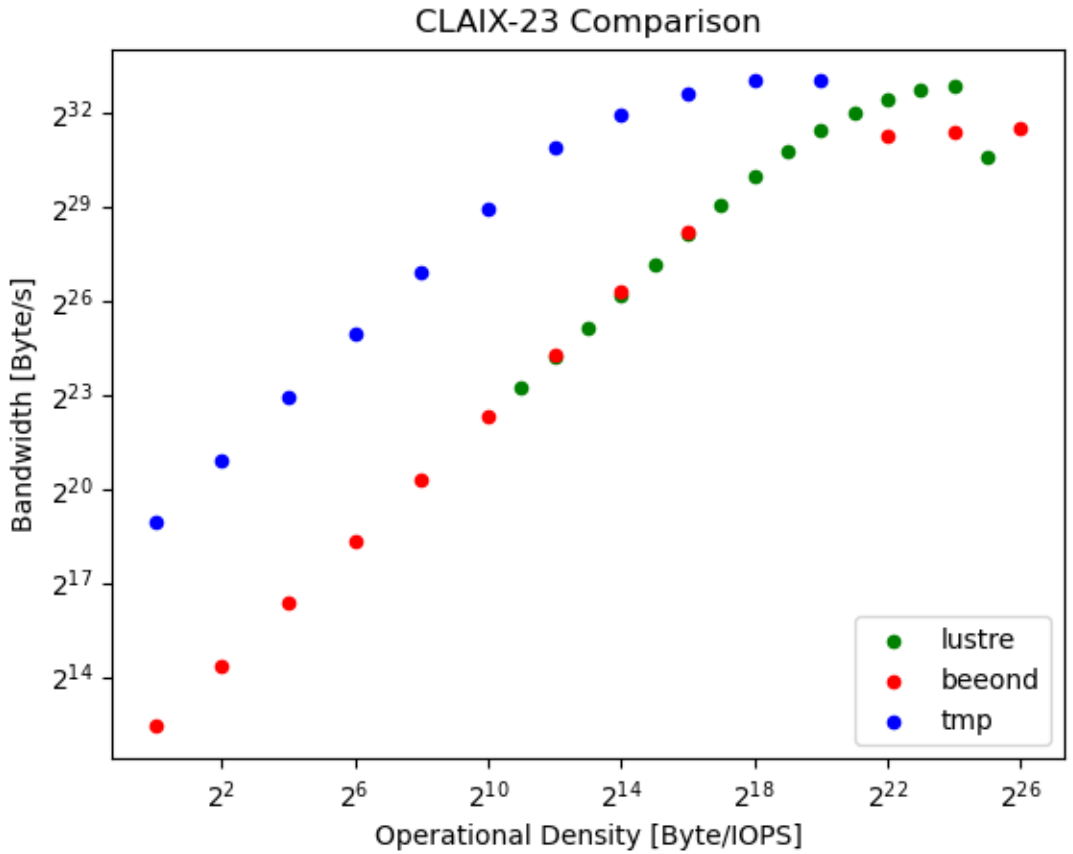
What is I/O? – Why do we need Files?

- **File use cases**
  - Persistent Storage
  - Input Data
  - Output Data
  - Intermediate Data
- **Files vs Memory**
  - Size constraints
  - Persistence

What is I/O? – Metadata vs Bandwidth

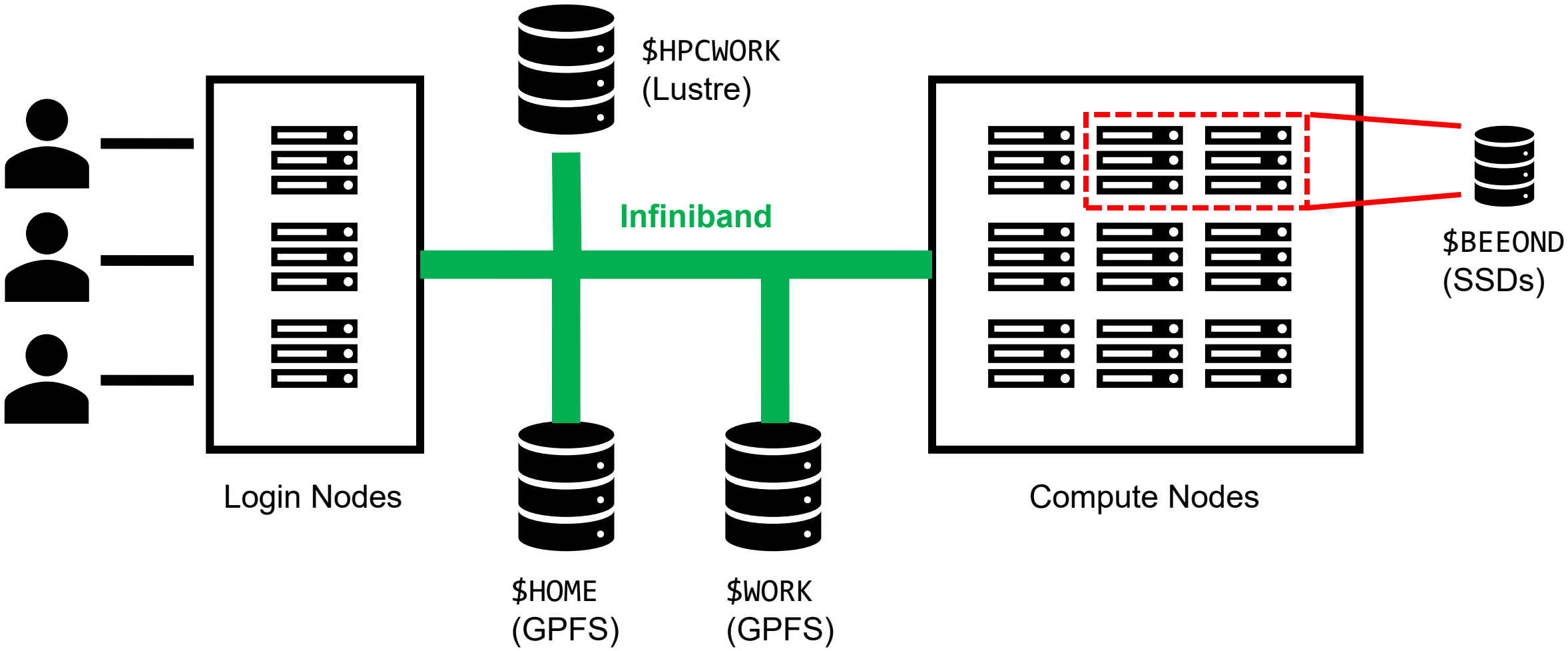
- **Metadata**
  - Anything that isn't the actual contents of the file
  - Filename, Permissions, Access Times, Size, Location…
- **Performance Metrics**
  - Metadata Operations (IOPS)
  - Bandwidth (Bytes/s)
  - Many small files vs. few large files

What is I/O? – Metadata vs Bandwidth

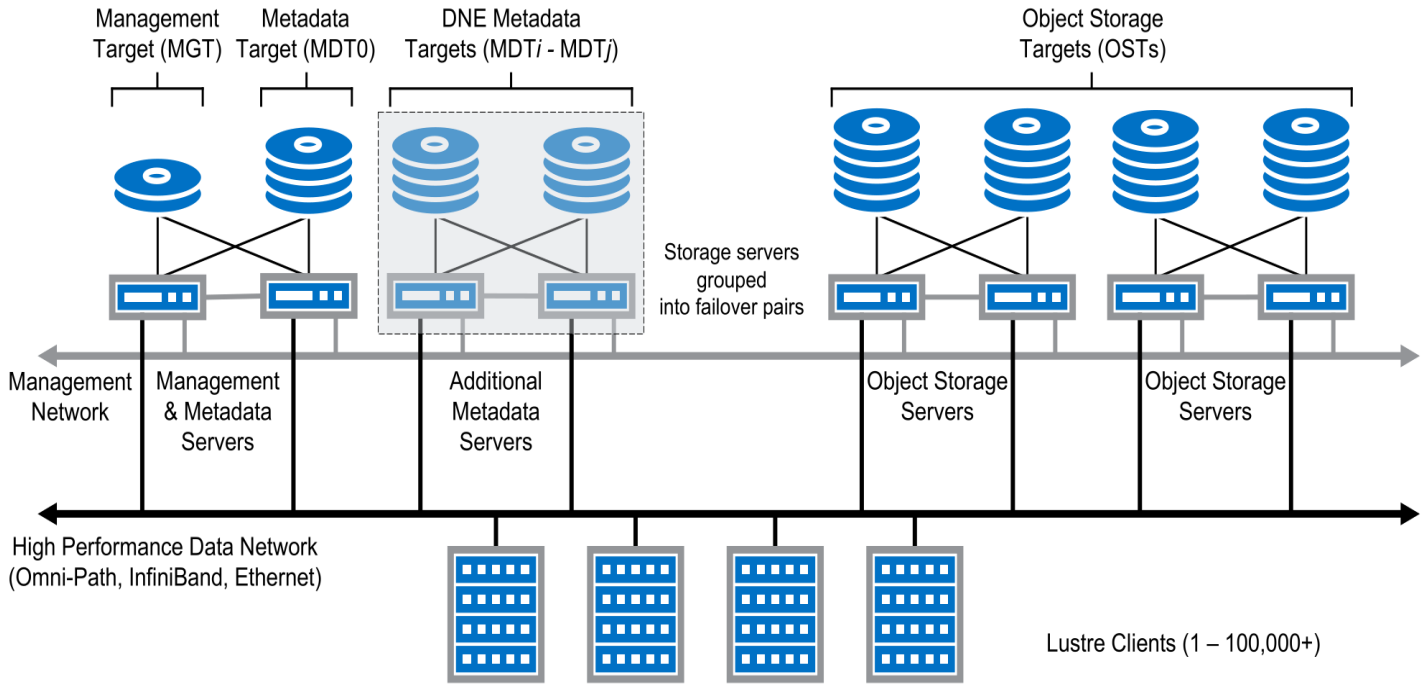


Contents

Datapaths in High Performance Computers – Cluster Setup

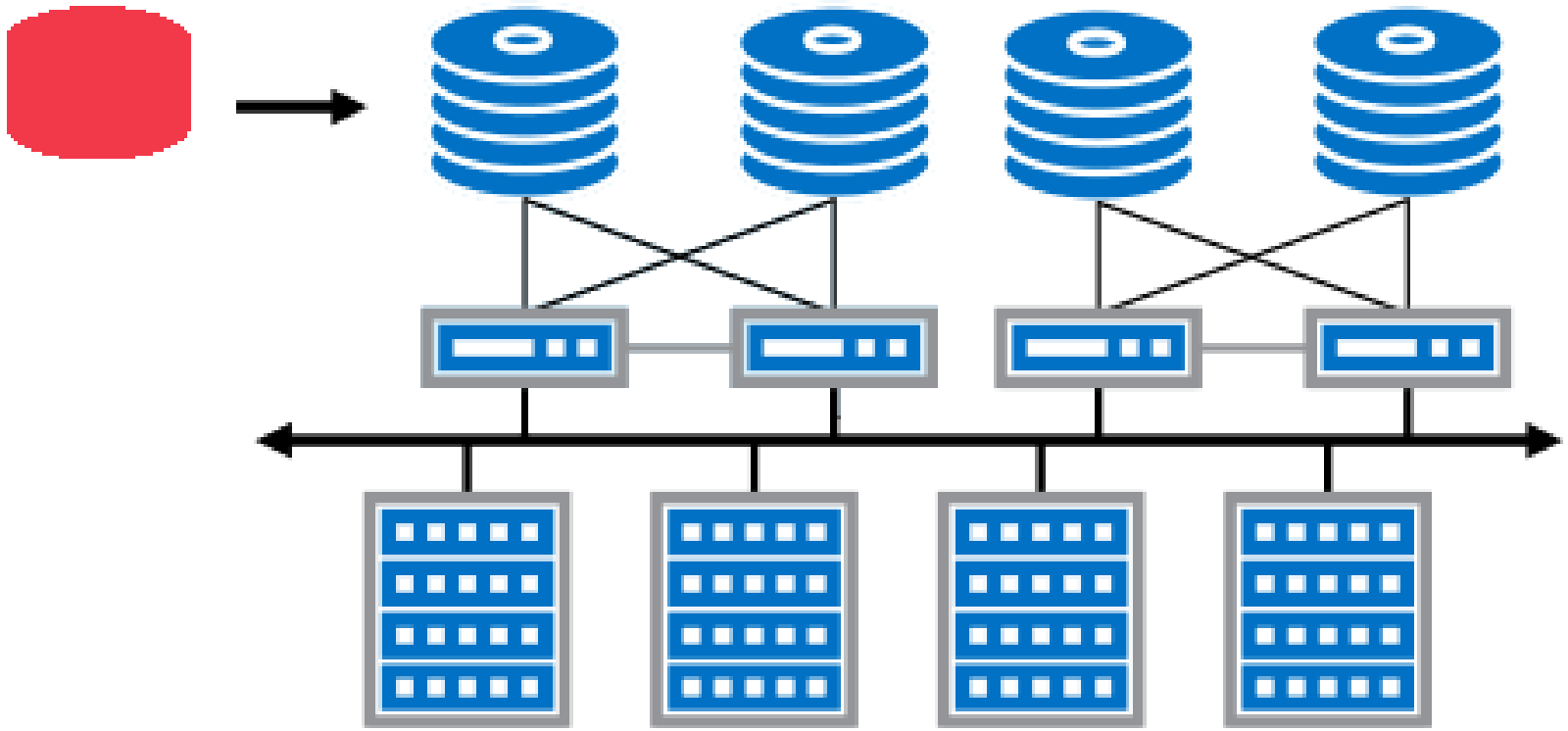


Datapaths in High Performance Computers – Parallel File System



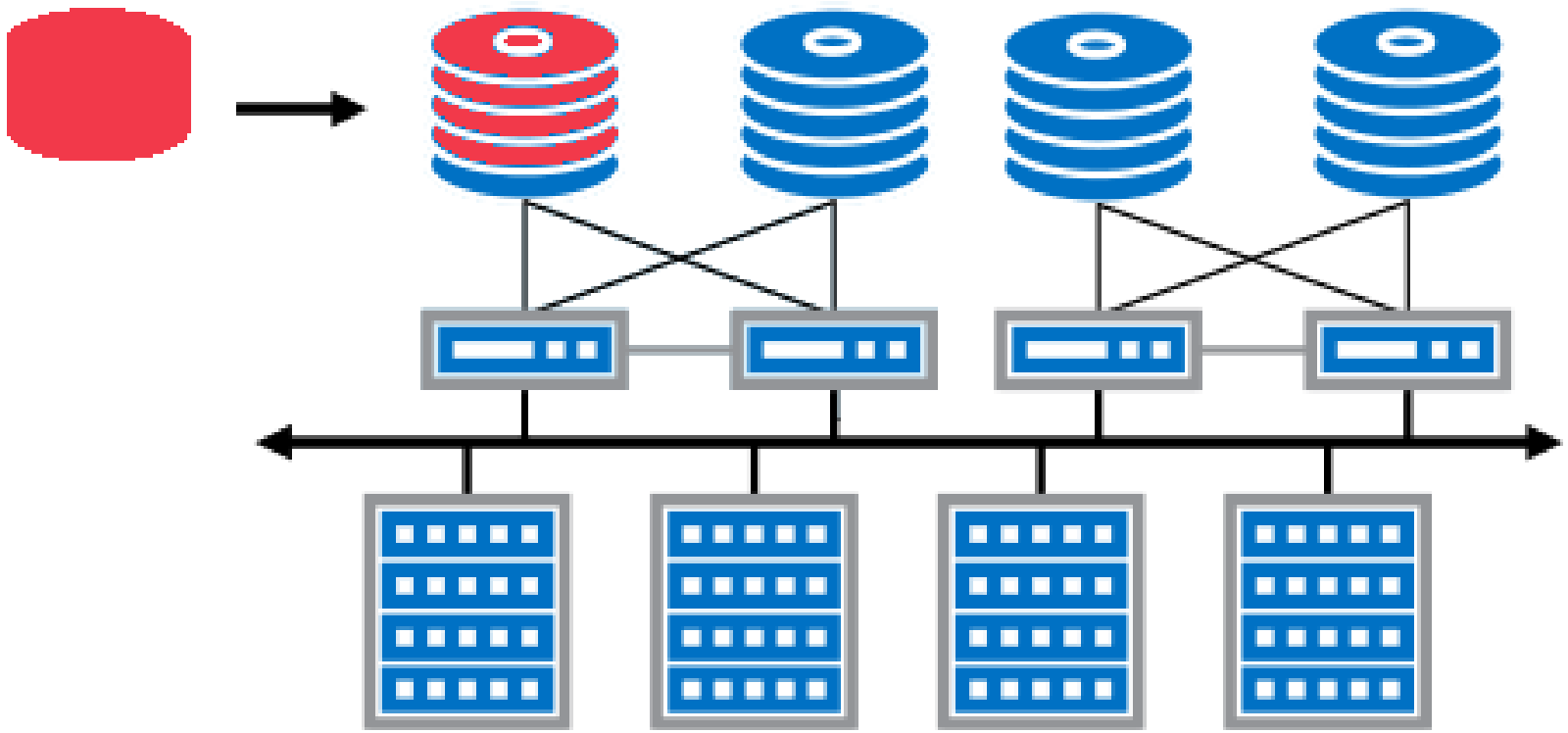
From https://wiki.lustre.org/images/6/64/LustreArchitecture-v4.pdf

Datapaths in High Performance Computers – Striping



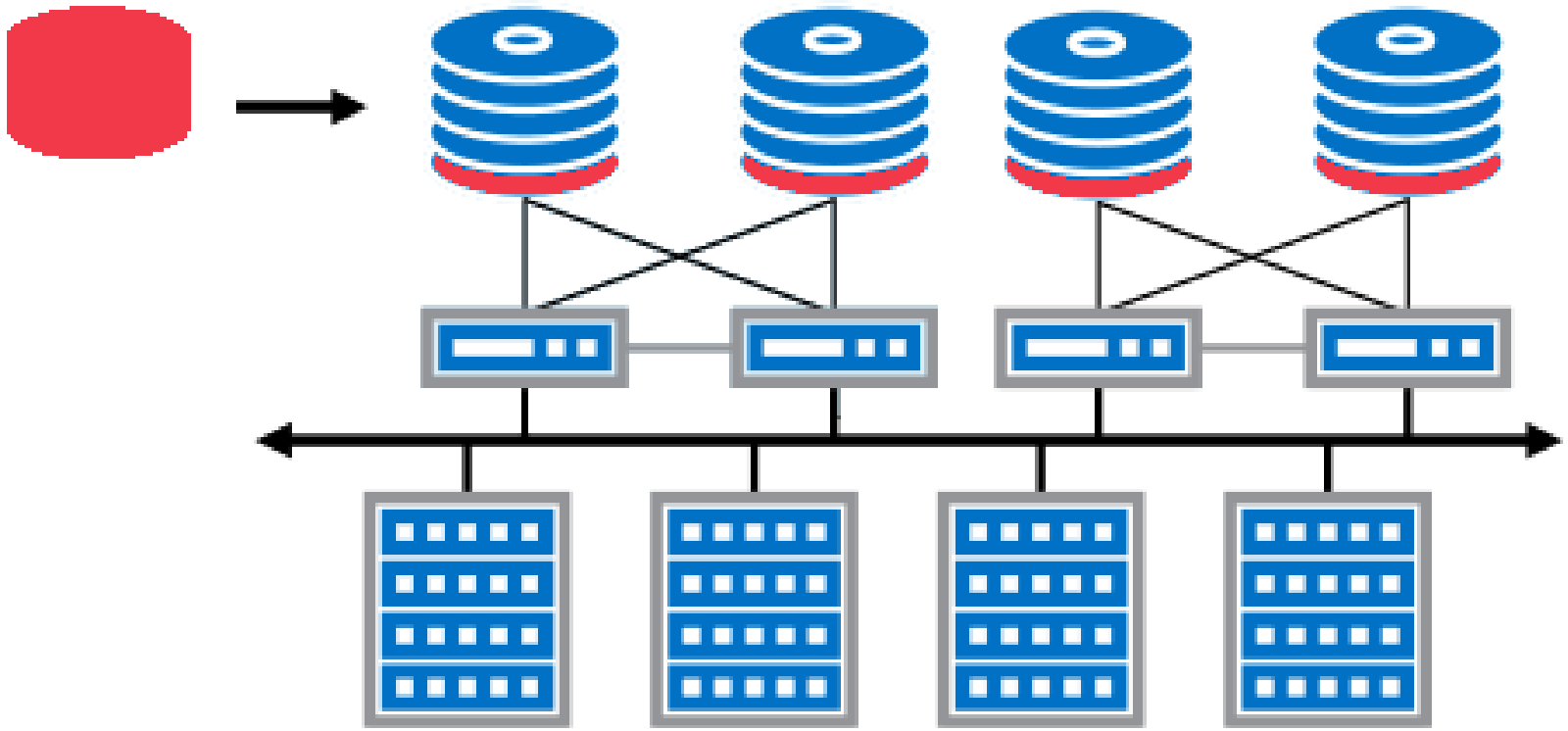
From https://wiki.lustre.org/images/6/64/LustreArchitecture-v4.pdf

Datapaths in High Performance Computers – Striping



From https://wiki.lustre.org/images/6/64/LustreArchitecture-v4.pdf

Datapaths in High Performance Computers – Striping



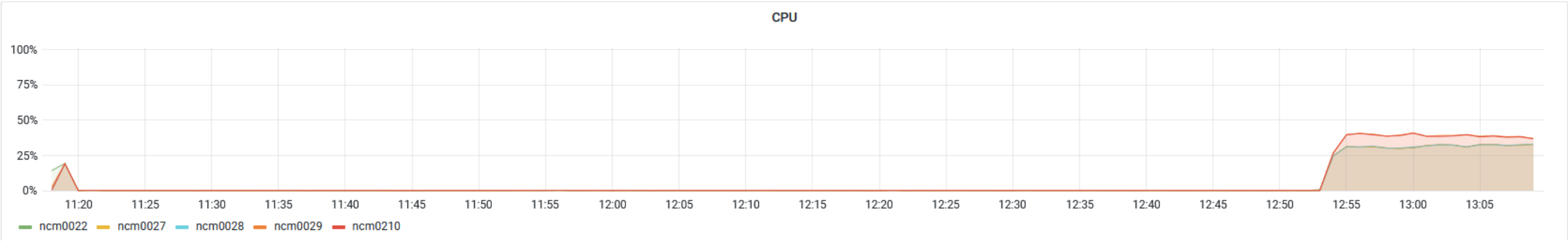
From https://wiki.lustre.org/images/6/64/LustreArchitecture-v4.pdf

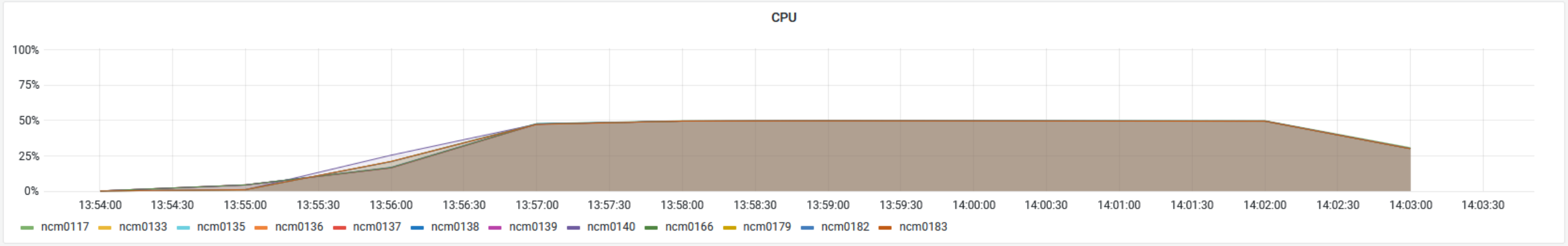
Contents

When Does I/O Matter? – I/O Bottlenecks

- **Application Side Considerations**
  - Am I using a significant number of I/O operations?
  - Am I using a significant amount of data?
  - Simulations: Writing of time-step data
  - Machine Learning: Training Sets & Model Weights
- **Cluster Side Considerations**
  - Is the filesystem suited for my task?
  - Is the filesystem overloaded?
  - Is the filesystem isolated from other users?
  - Does my data (not) fit in the cache(s)?

When Does I/O Matter? – I/O Bottlenecks





When Does I/O Matter? – ML/AI Considerations

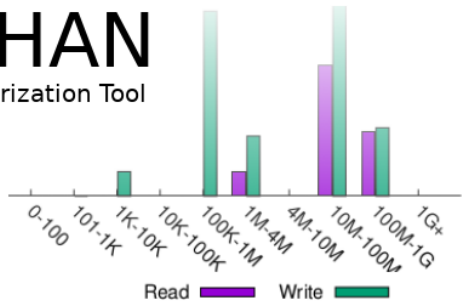
- **Storing training data**
  - Datasets are often very large in size (>1 TiB)
  - Could consist of many single files (>1 million)
- **Loading training data**
  - Entire dataset is required for each epoch
  - Augmentations and reshuffling etc. may be required every epoch
  - Often too large to be cached in the GPU and/or main memory
- **Model weights**
  - Have to be stored during training and accessed during inference
  - One-time load per inference instance
- **Input data**
  - Important for offline inference

Contents

How to Identify and Solve Bottlenecks

- **Profile and analyze**
  - Darshan
  - Score-P
  - Perfmon
- **Improve access patterns**
  - Choose more suitable file formats
- **Choose the correct file system**

DARSHAN
HPC I/O Characterization Tool



Score-P
Scalable performance measurement infrastructure for parallel codes

BeeOND

lustre