



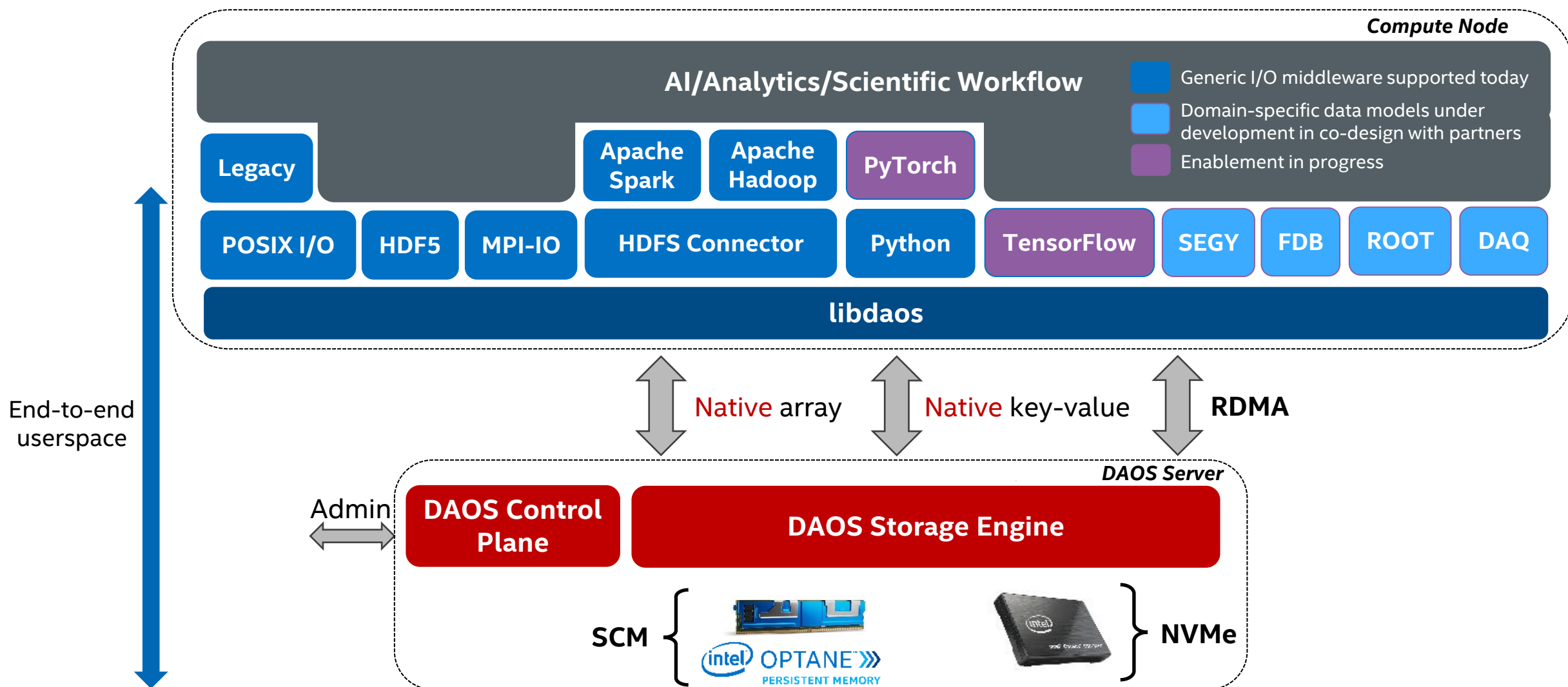
Distributed Asynchronous Object Storage (DAOS)

Evaluating DAOS Storage on ARM64 Clients

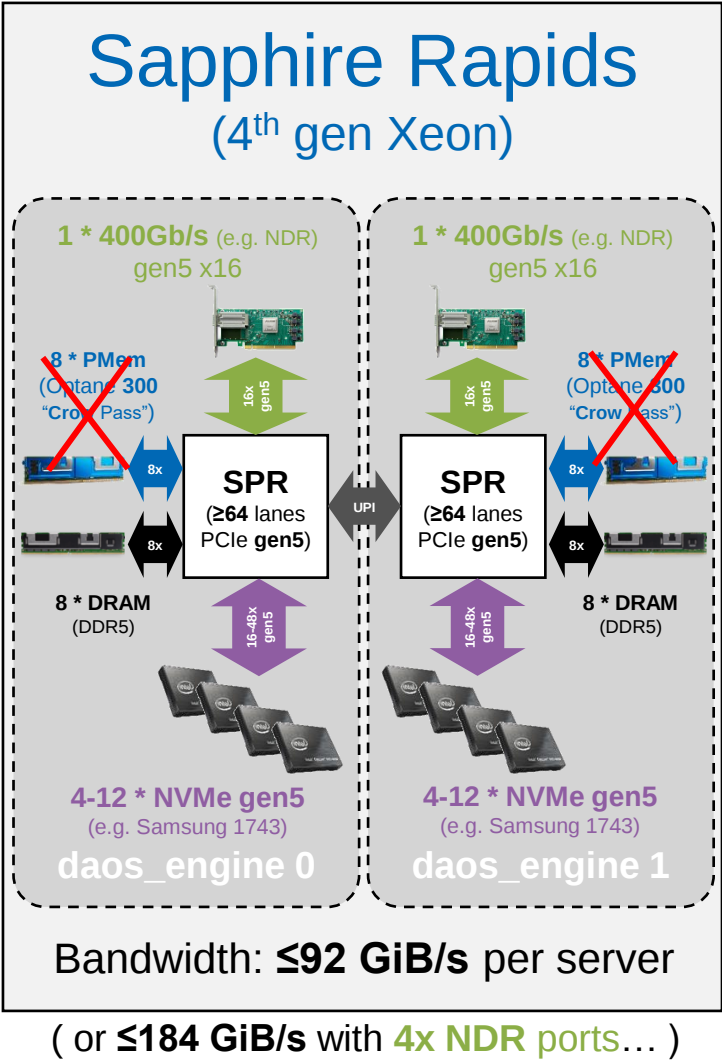
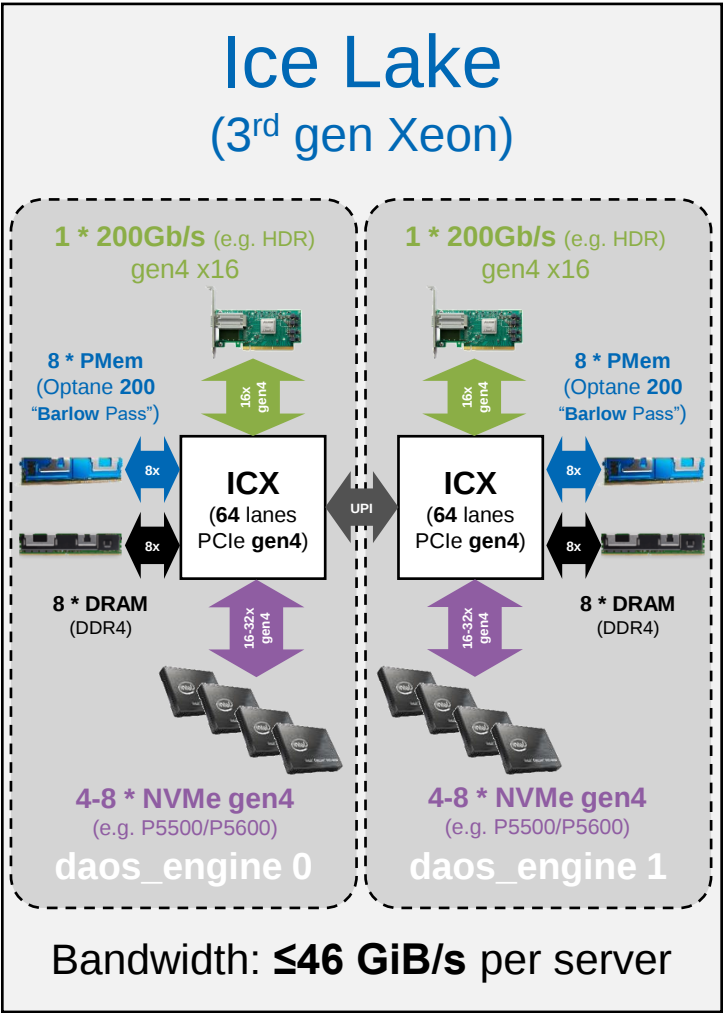
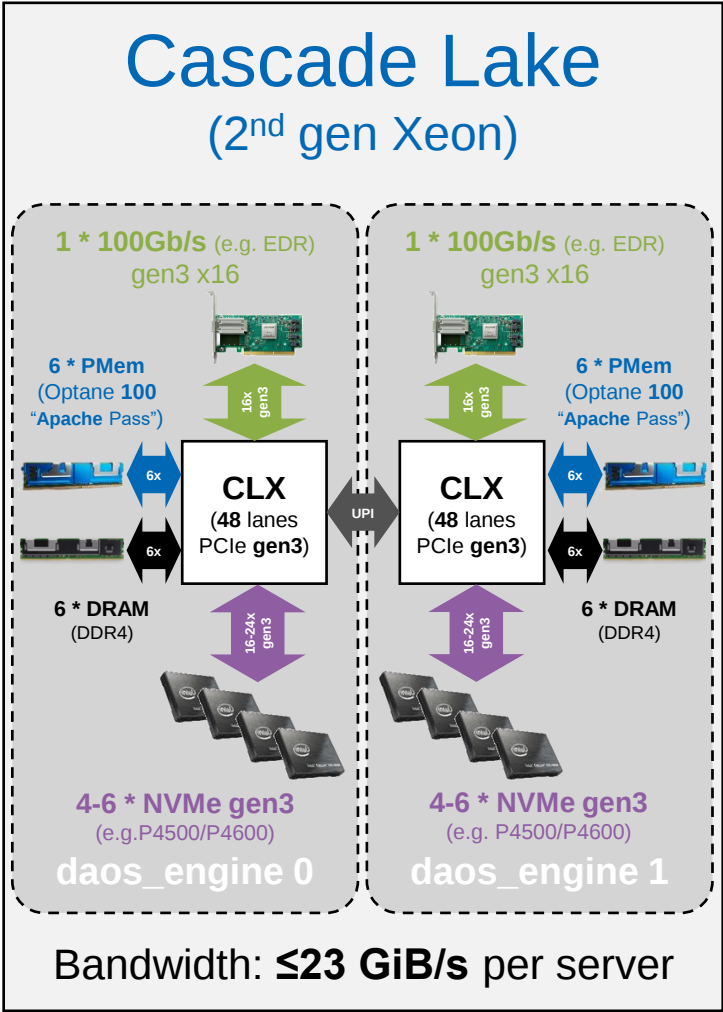
27-Feb-2023 | Michael Hennecke (Intel), Motohiko Matsuda and Masahiro Nakao (Riken)



DAOS Software Ecosystem



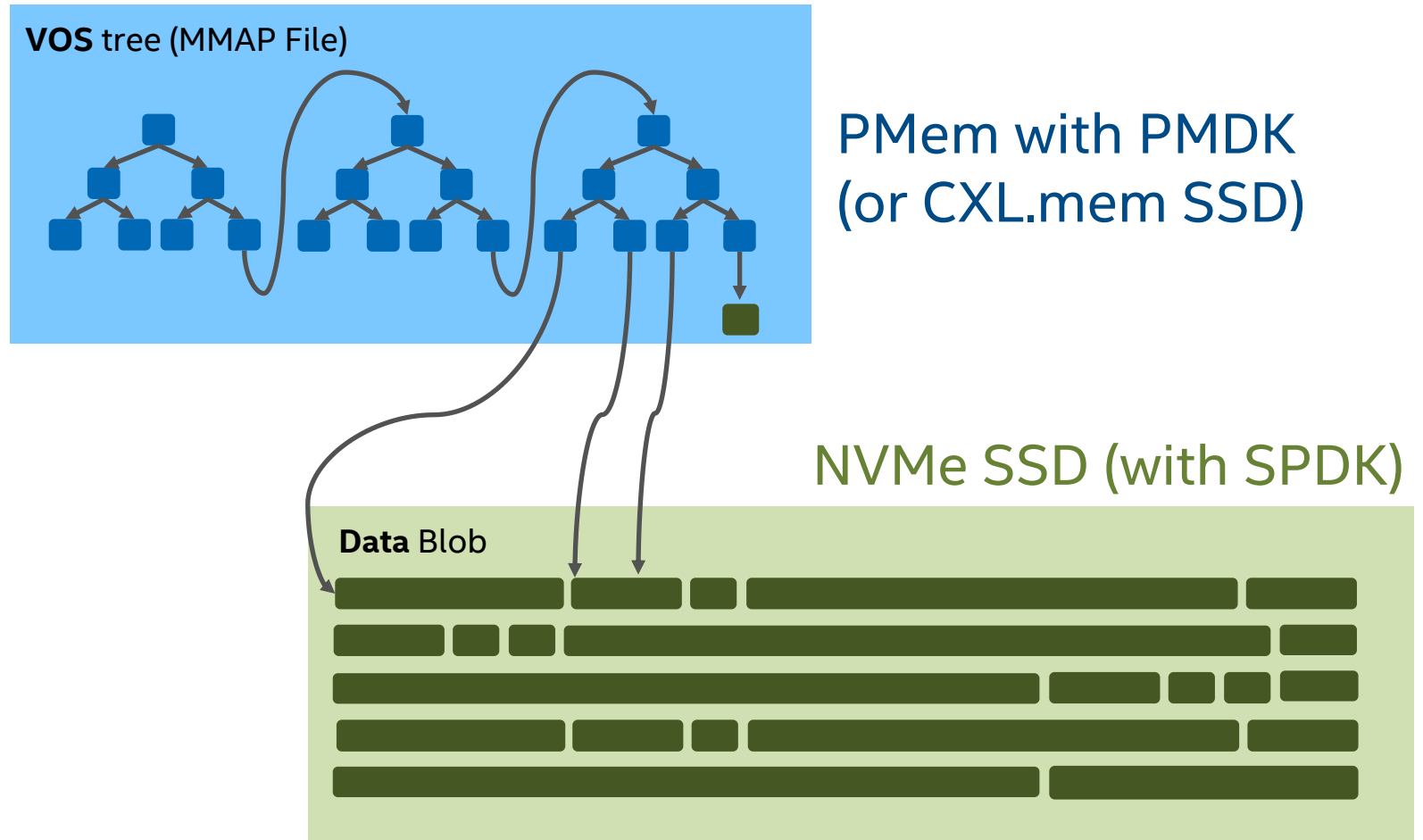
DAOS Servers on Intel Xeon-SP



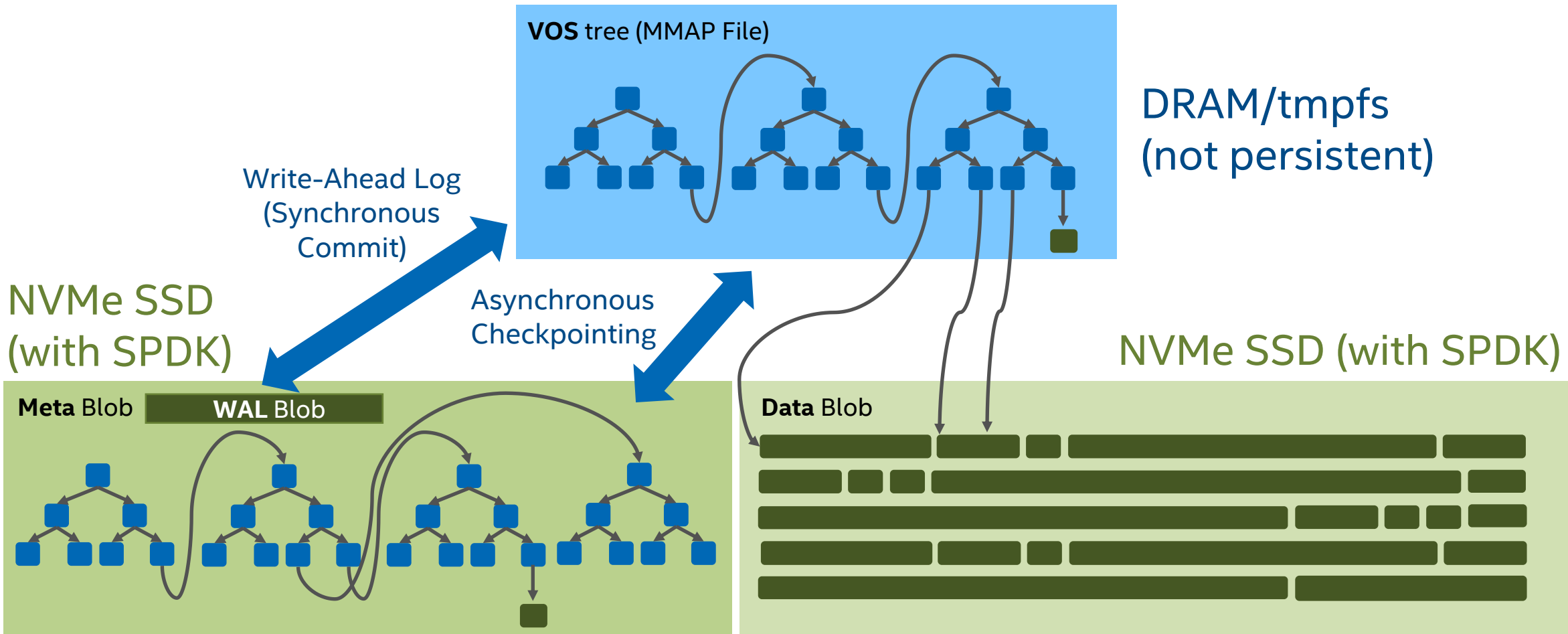
DAOS Metadata Path Evolution

- Today, DAOS utilizes SCM (Optane PMem) over PMDK for **VOS**
 - Byte addressable; hardware “IOPS” performance; PMDK transactions
 - Can emulate SCM by DRAM (tmpfs) – for development, ephemeral storage (e.g. GCP), ...
- Work on **alternative code path** for VOS tree in DRAM **is being accelerated**
 - Using **Write-Ahead Log (WAL)**, and asynchronous **VOS checkpointing** to NVMe
 - No PMem or PMDK dependency – broadens the DAOS server ecosystem
 - <https://daosio.atlassian.net/wiki/spaces/DC/pages/11196923911/Metadata+on+SSDs>
- **CLX.mem** in CXL 2.0 will provide byte addressability; performance TBD
 - First devices shown by storage vendors at Flash Memory Summit Aug/2022

DAOS Backend using Persistent Memory

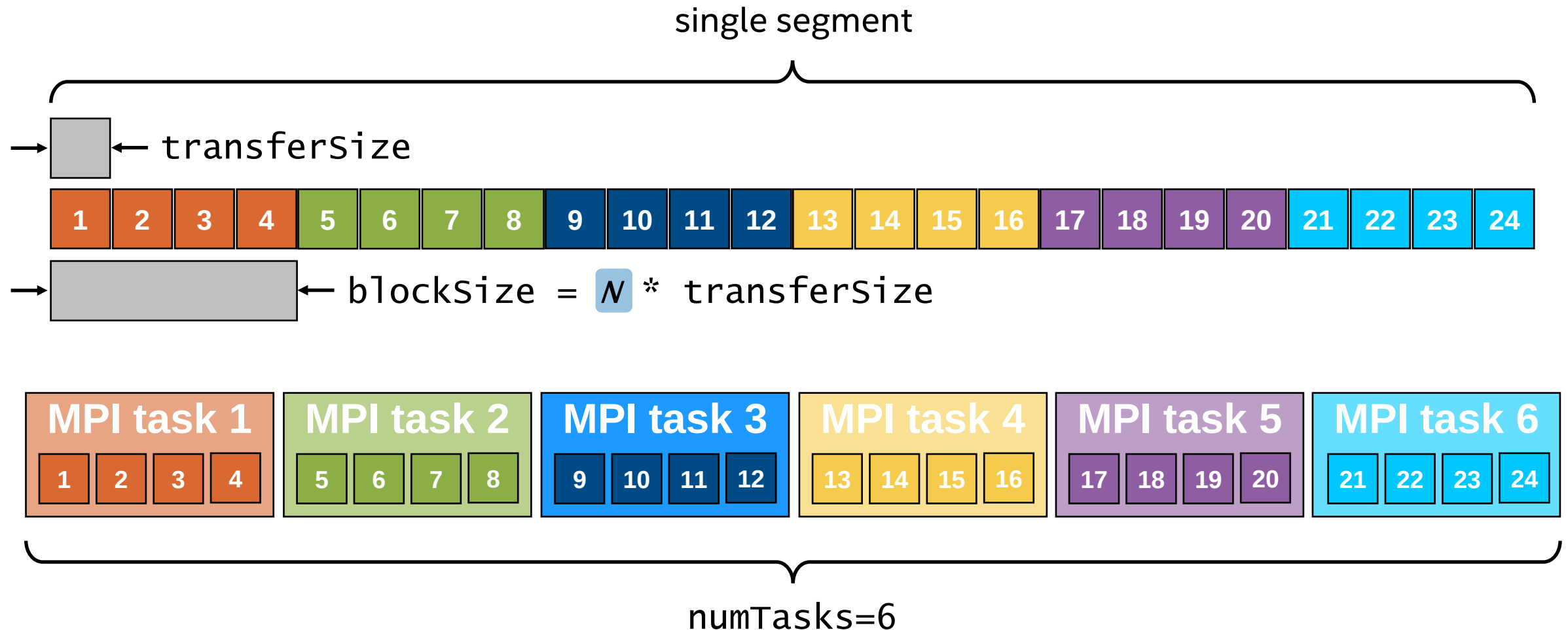


DAOS Backend using Volatile Memory



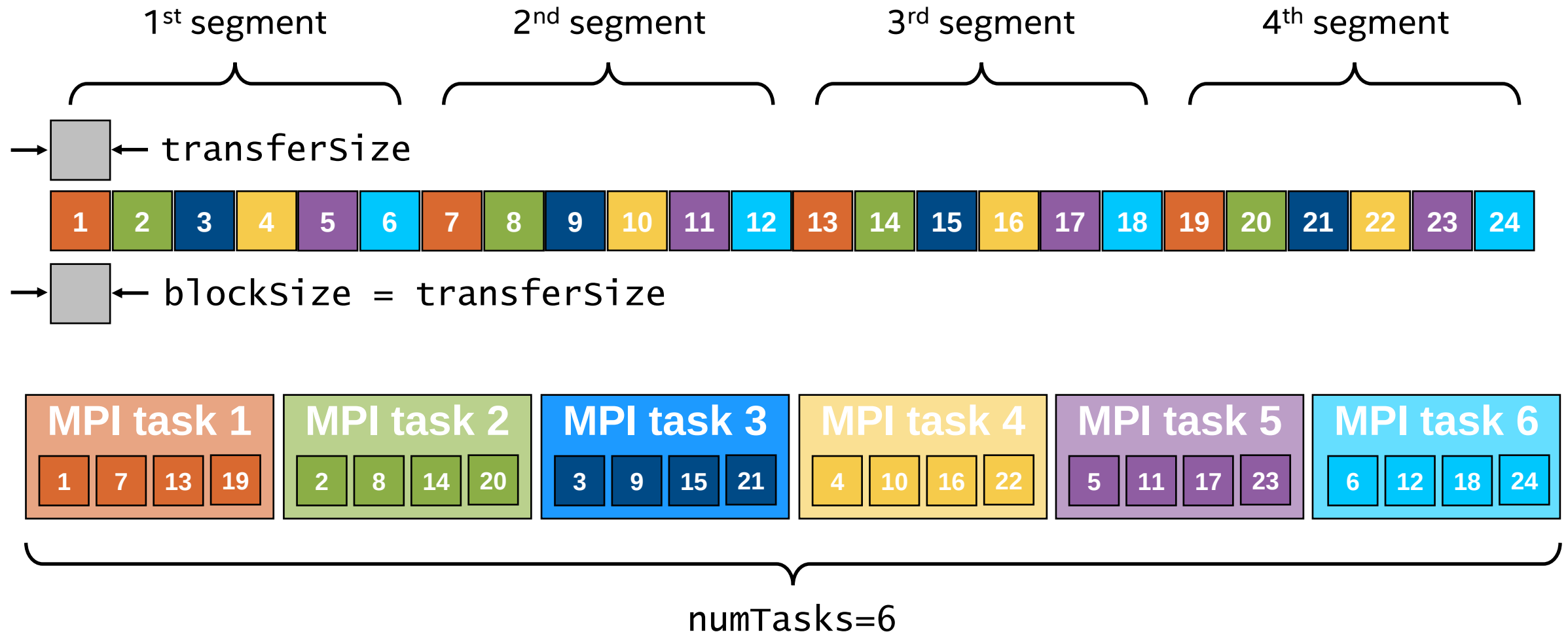
File layout for IOR “easy” (sequential) data distribution

Single shared file (`filePerProc=0`) with *one* segment (`segmentCount=1`)

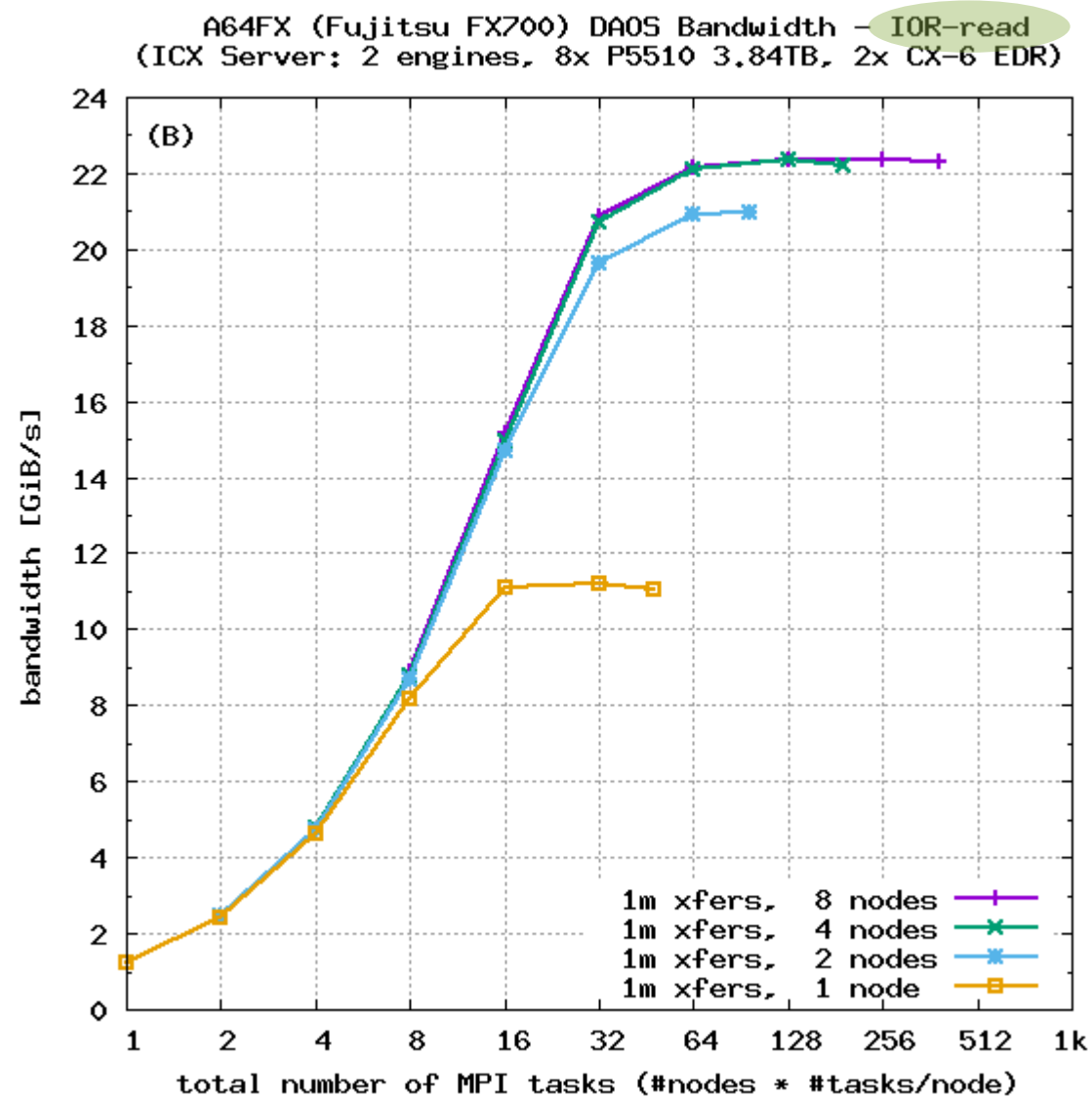
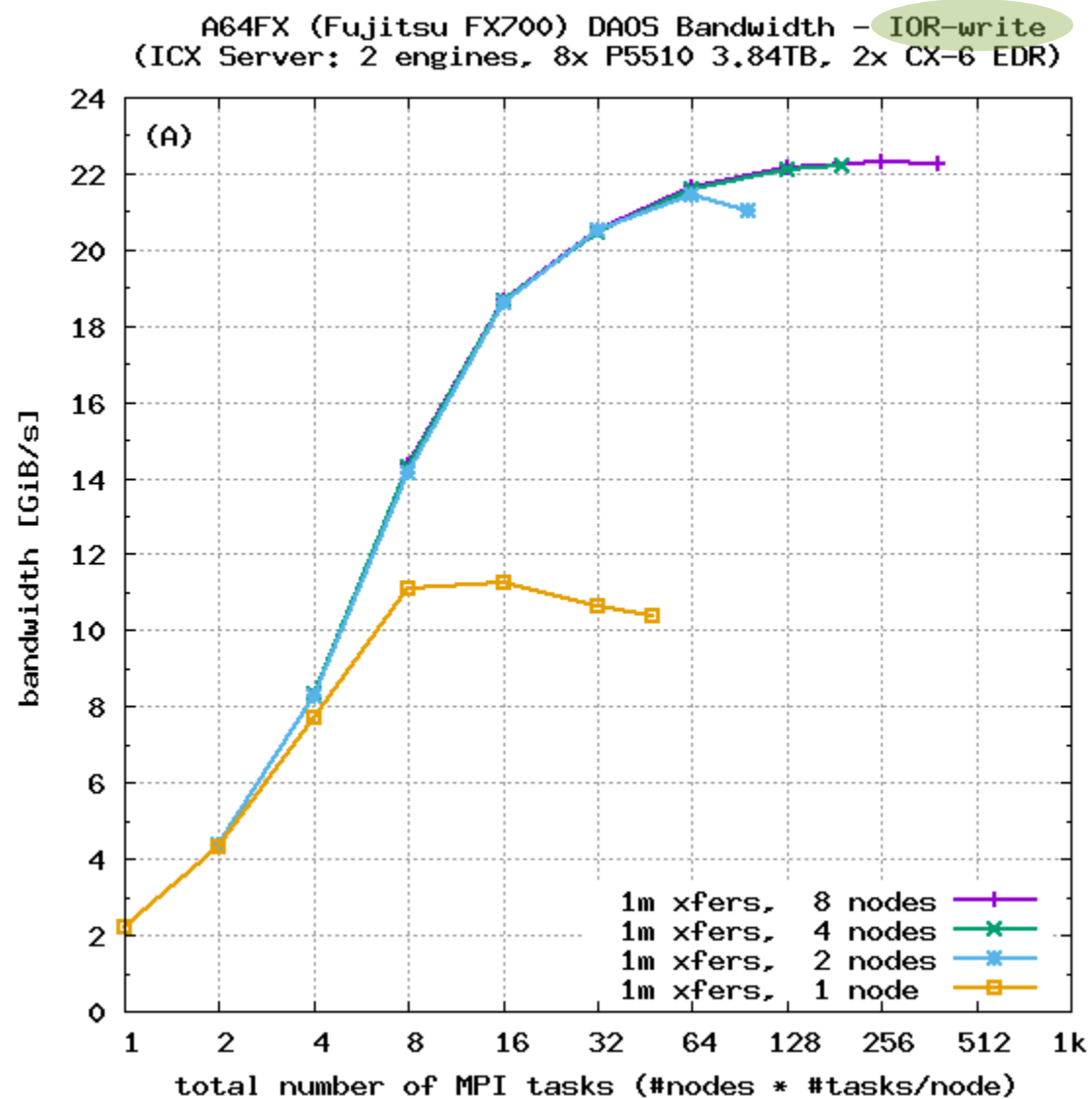


File layout for IOR “hard” (strided) data distribution

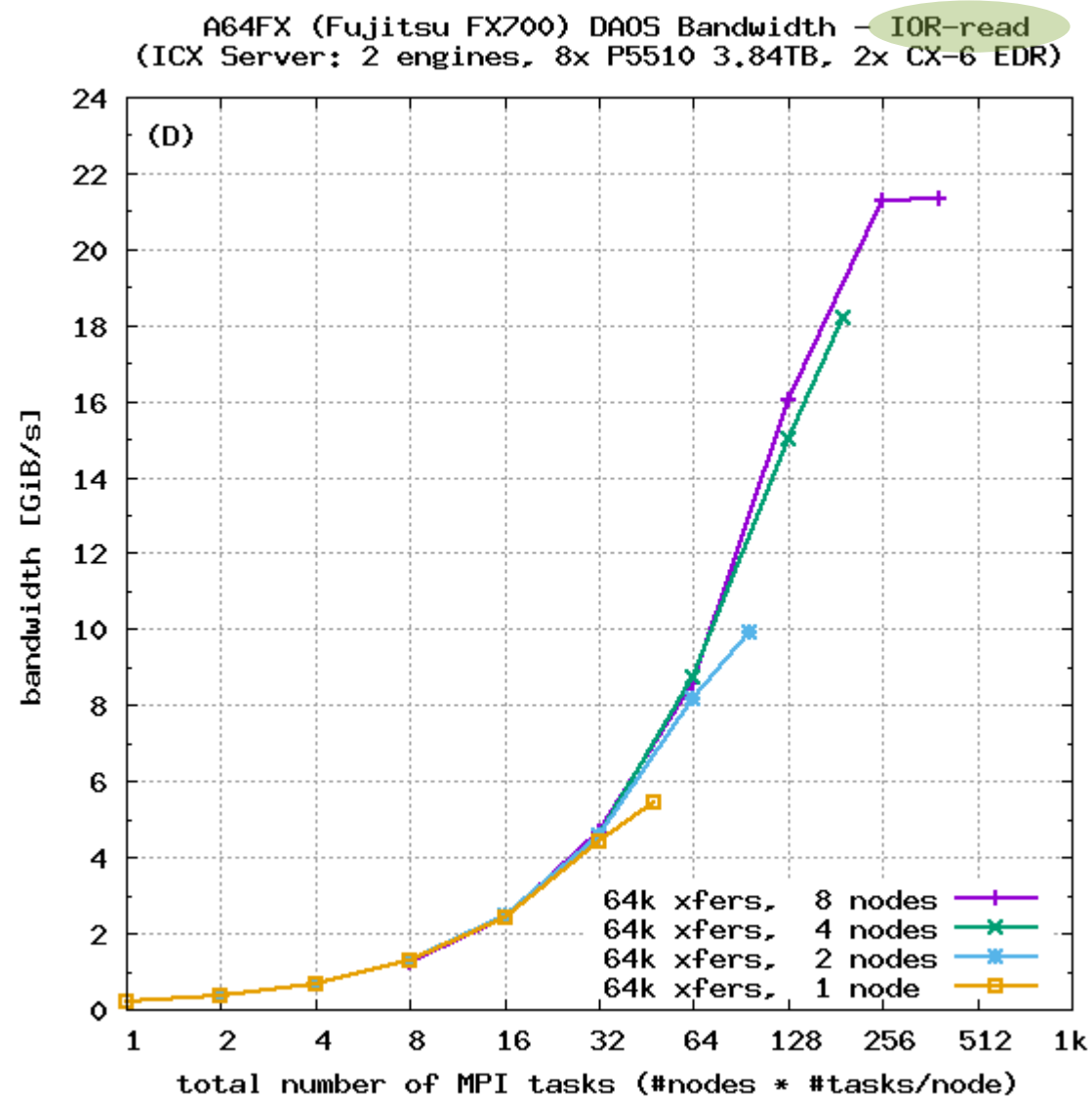
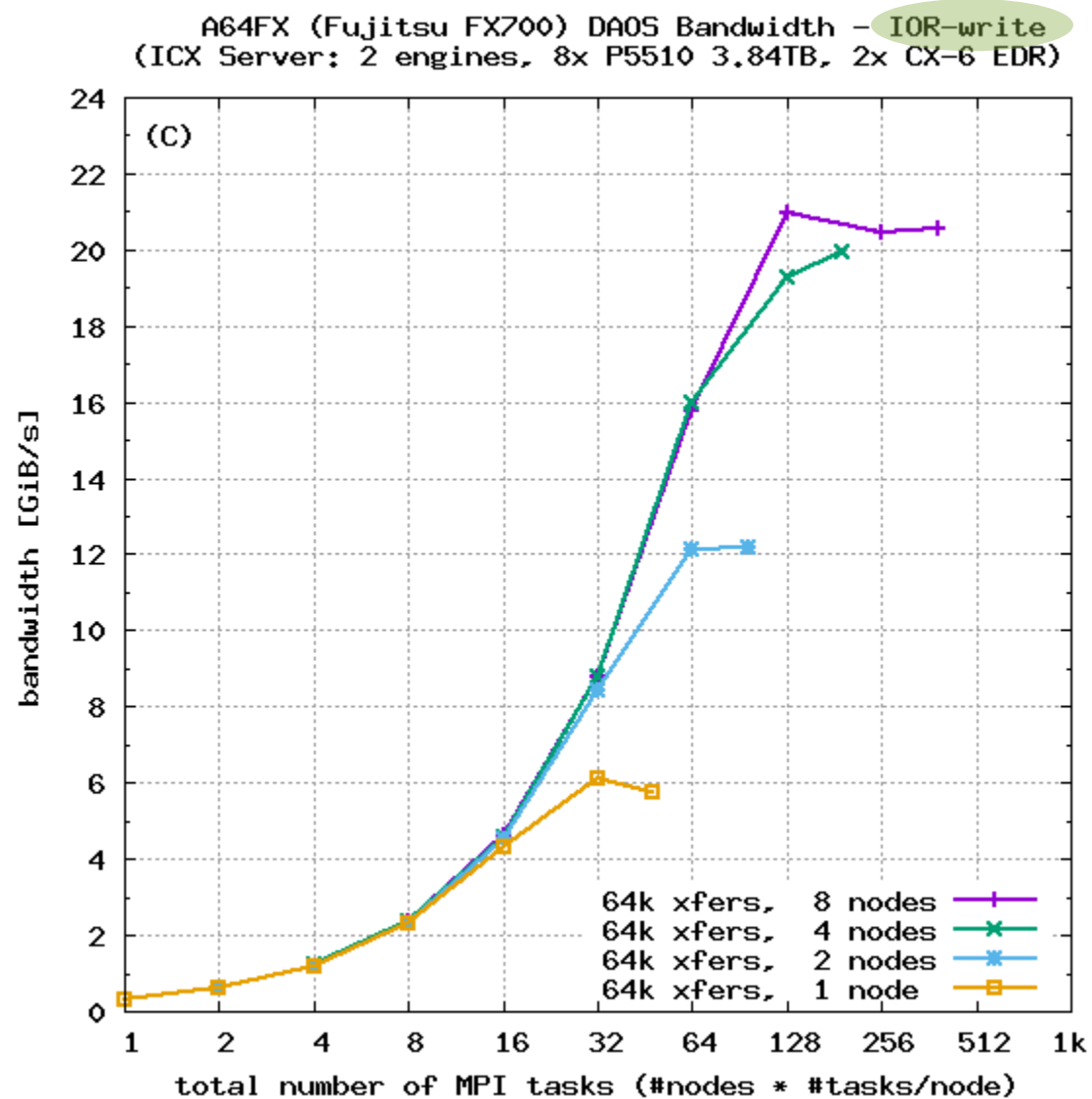
Single shared file (`filePerProc=0`) with N segments (`segmentCount=N`)



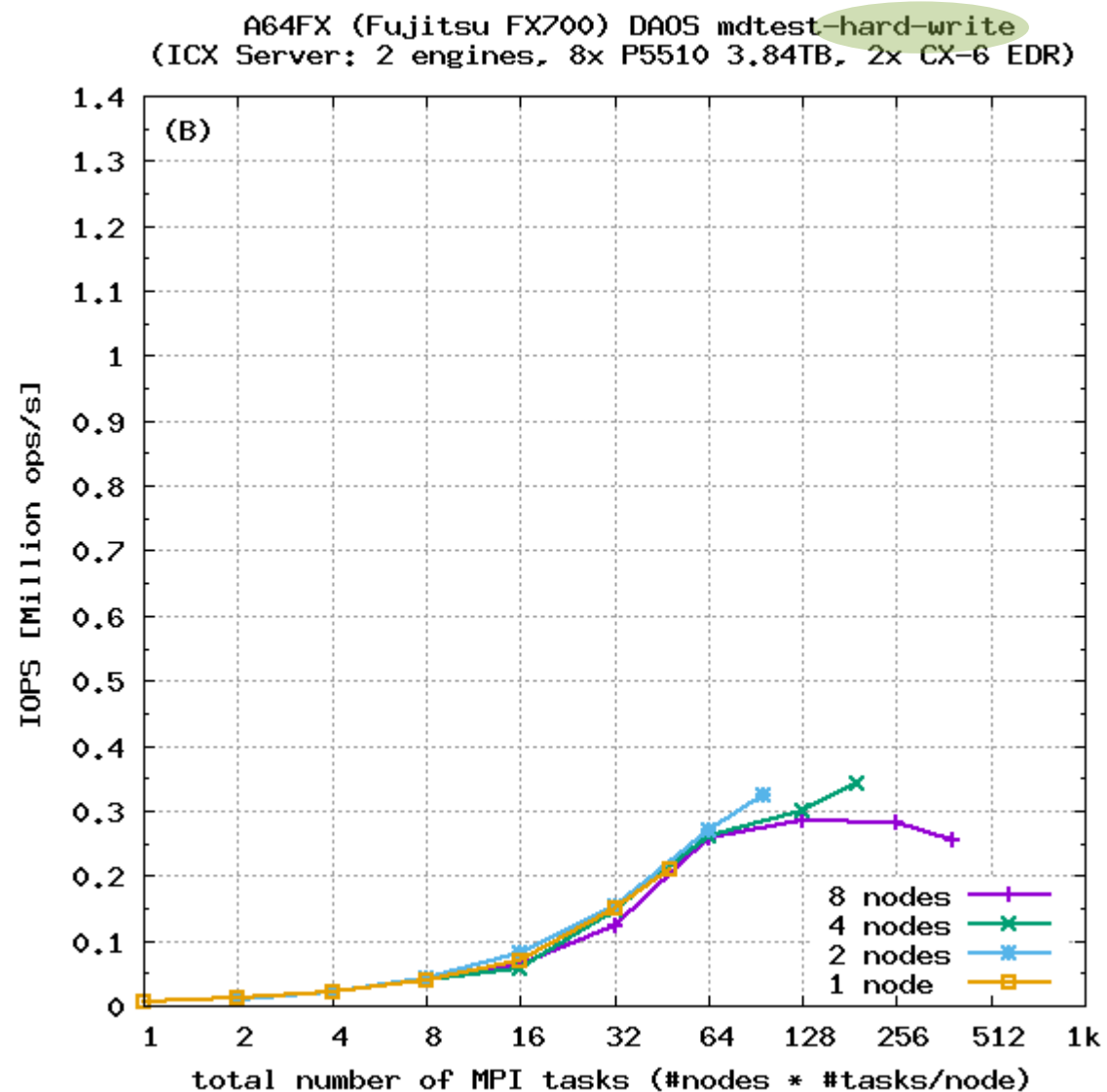
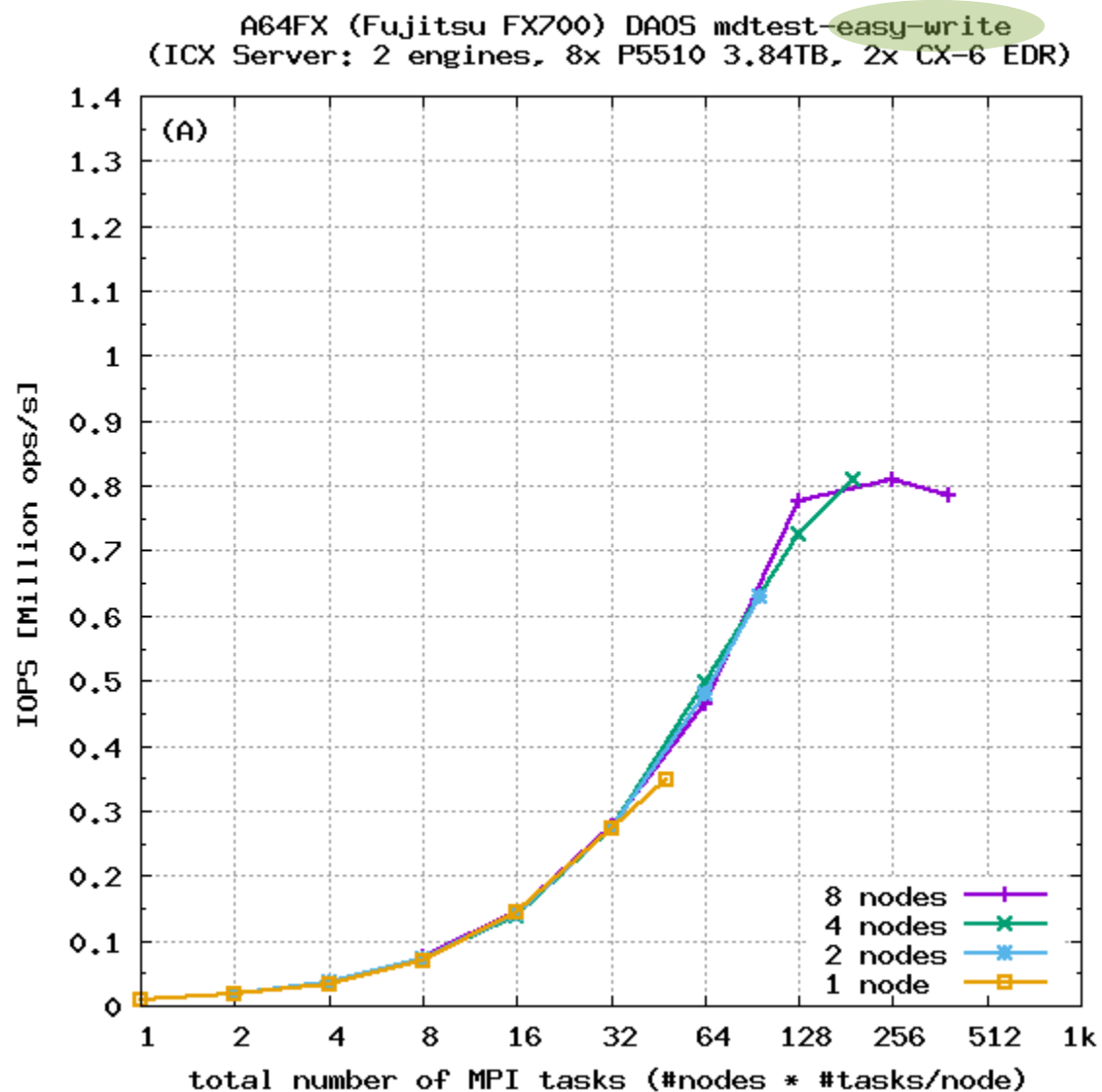
DAOS Bandwidth on A64FX Clients (1MiB xfers)



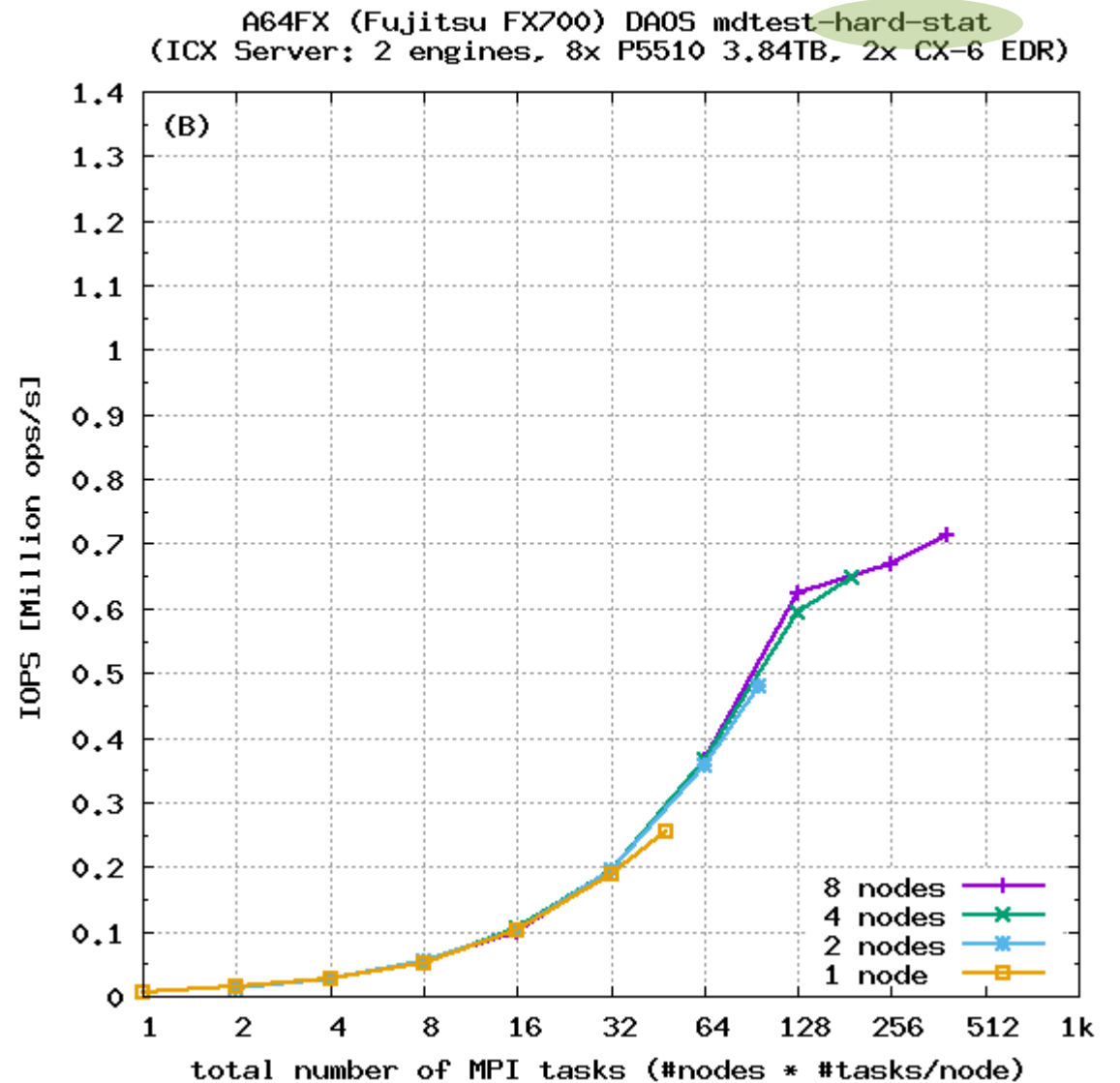
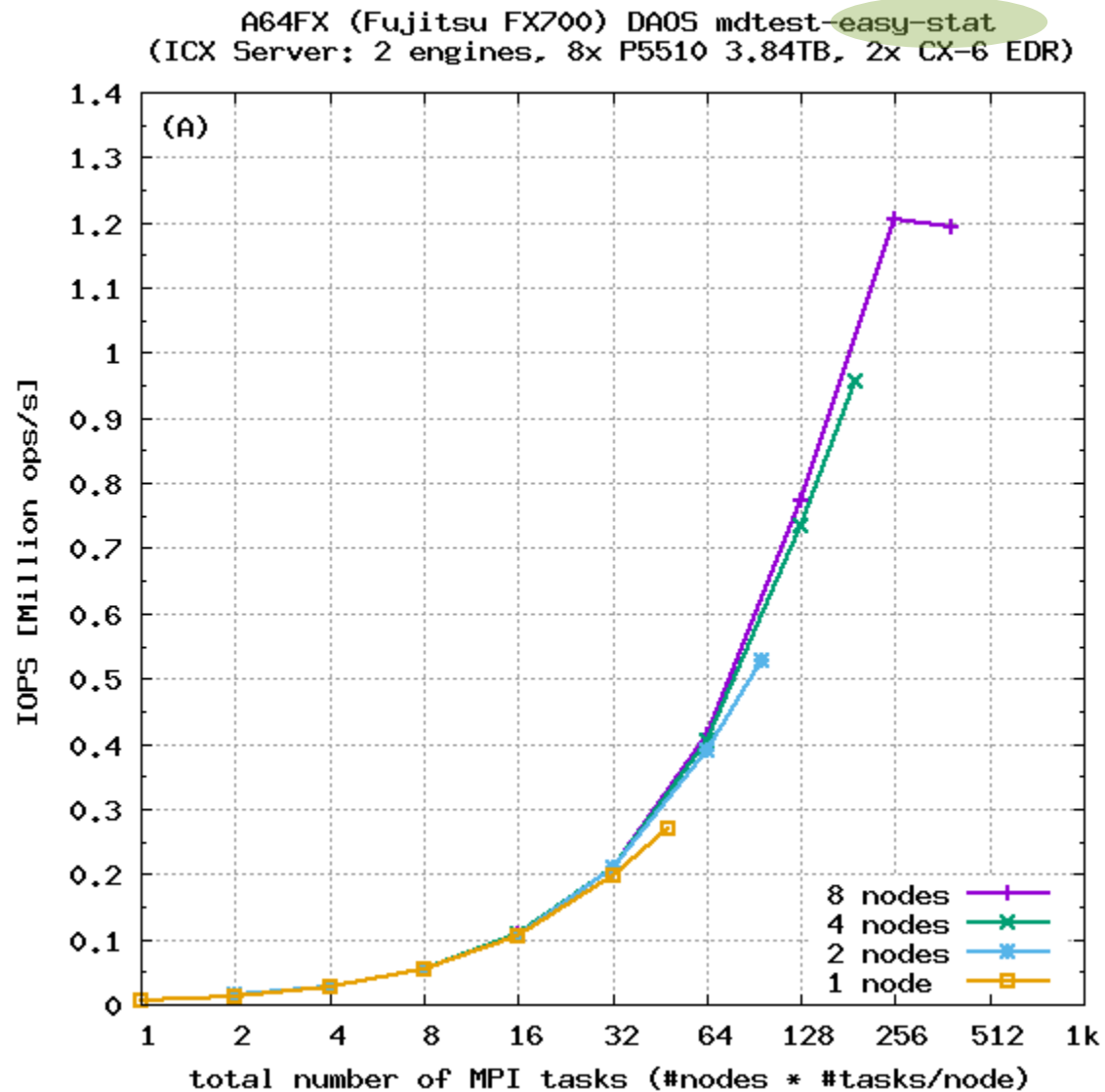
DAOS Bandwidth on A64FX Clients (64kiB xfers)



DAOS Metadata Rates on A64FX Clients – create/write

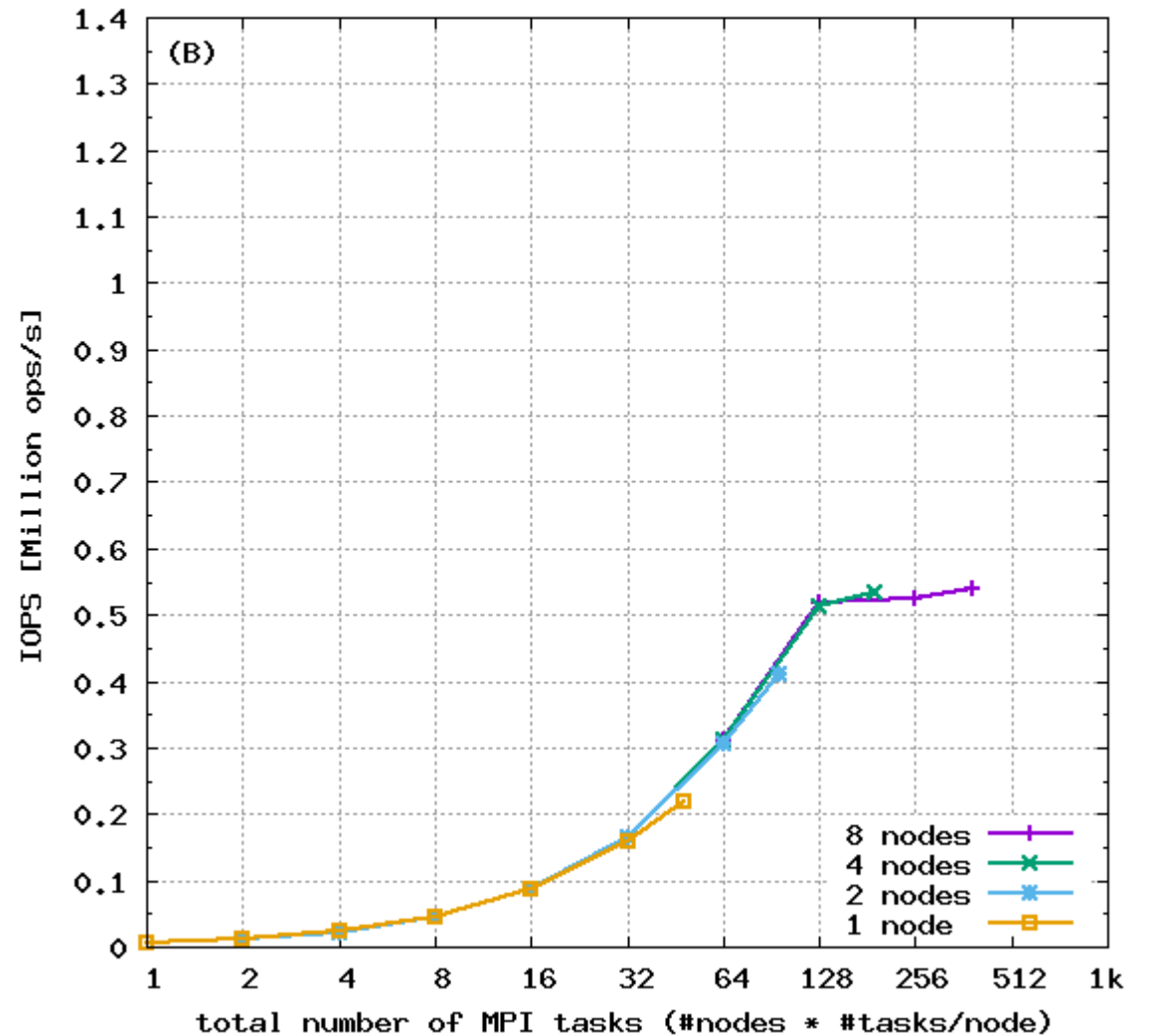


DAOS Metadata Rates on A64FX Clients – stat

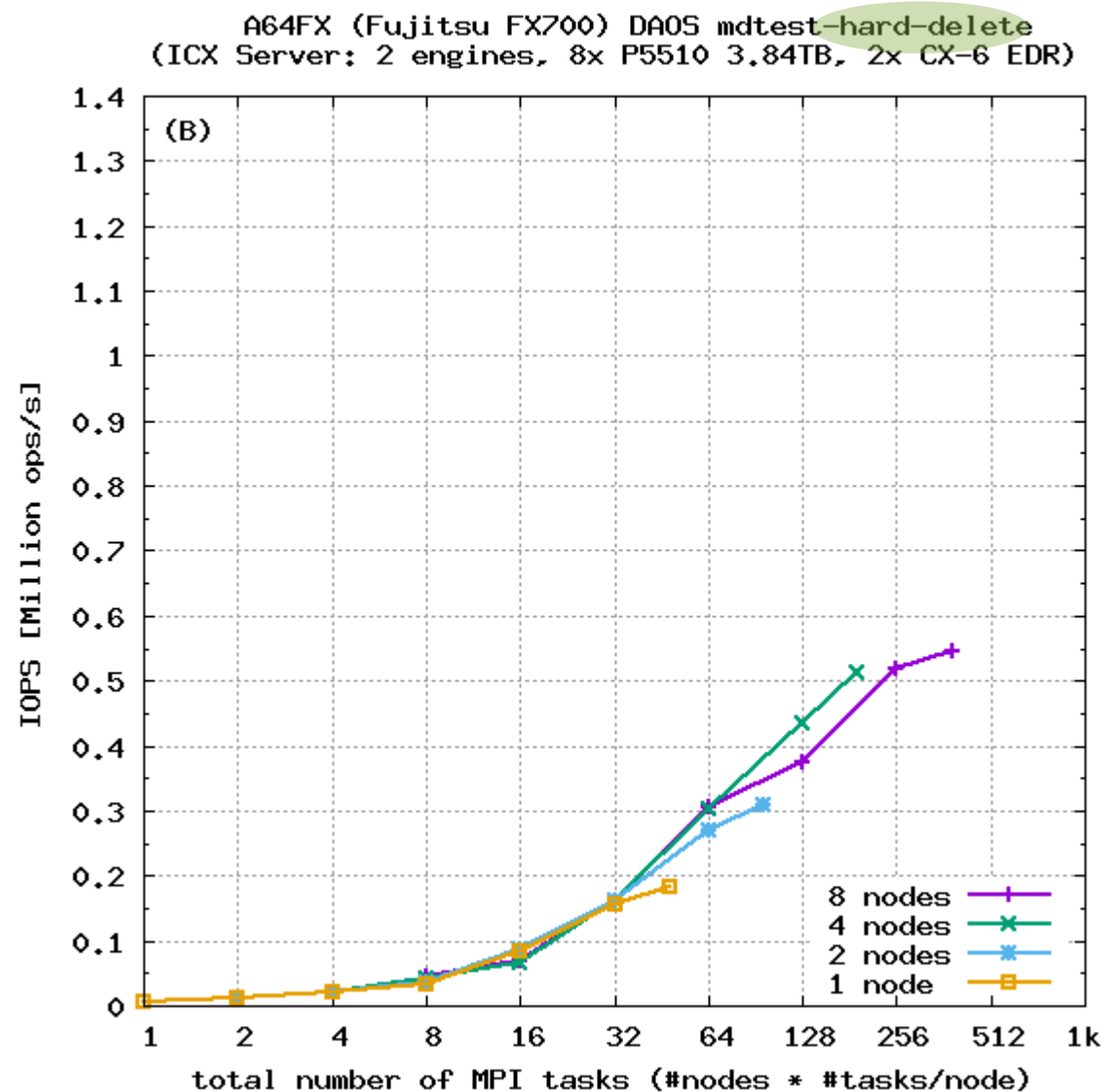
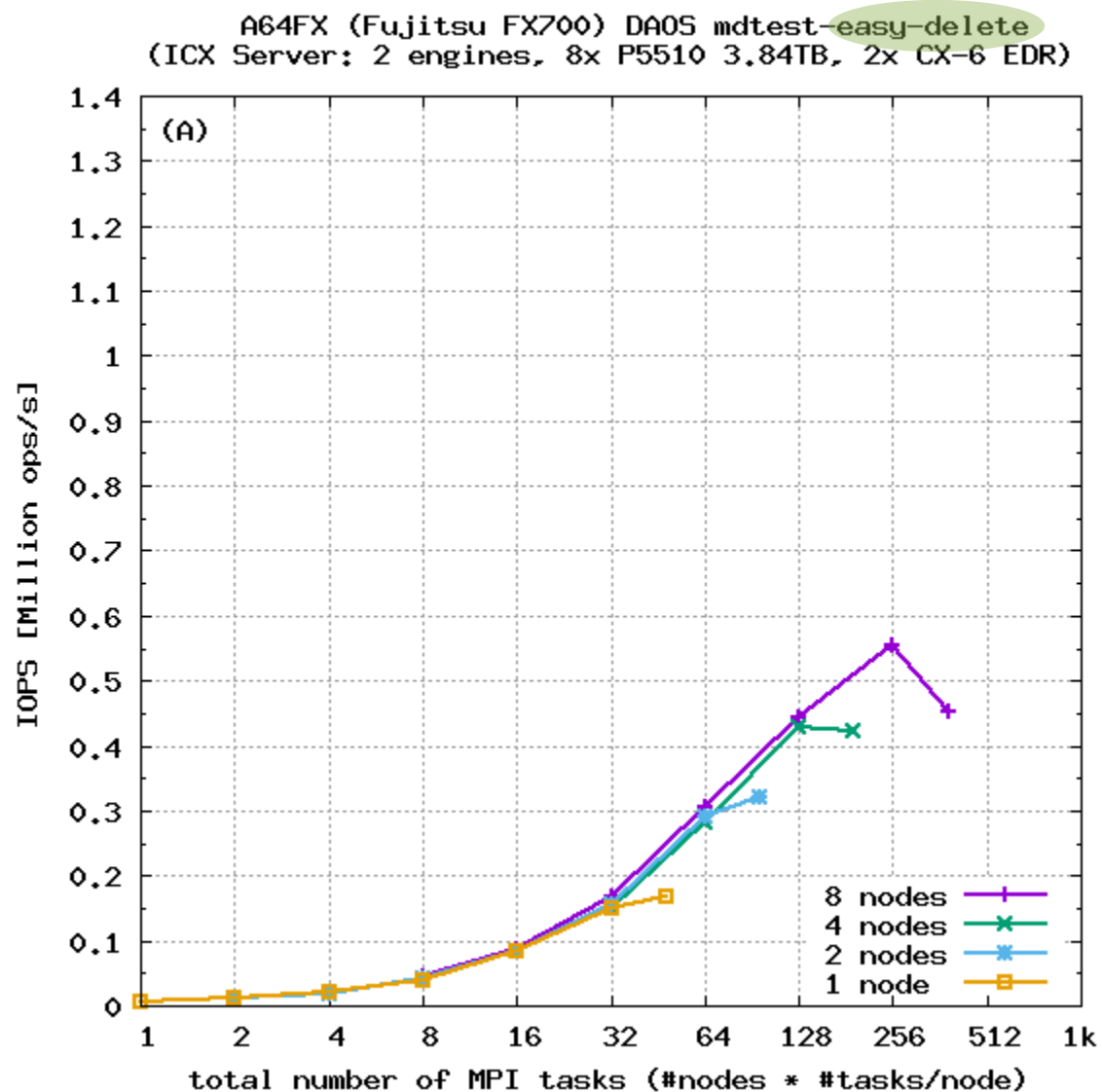


DAOS Metadata Rates on A64FX Clients – read

A64FX (Fujitsu FX700) DAOS mdtest-hard-read
(ICX Server: 2 engines, 8x P5510 3.84TB, 2x CX-6 EDR)



DAOS Metadata Rates on A64FX Clients – delete



Summary and Outlook

- DAOS community has been active to advance ARM64 support
 - Added ARM64 runner to the DAOS github actions
 - Working on a path to build and provide aarch64 RPMs
- DAOS 2.3.101 works fine on ARM64 Clients with InfiniBand or TCP
 - Stretch goal to evaluate DAOS on Fugaku clients (with libfabric/tofu)
- Testing DAOS Server software stack on ARM64 now
 - With MD-on-SSD technology preview (no more PMem required)
 - Minor issues on A64FX: only 32GiB HBM2 (no DRAM), hugepage size 512MiB ... mostly fixed

DAOS Resources

■ Community Resources

- Github: <https://github.com/daos-stack/daos>
- Online doc: <https://docs.daos.io/>
- Mailing list & slack: <https://daos.groups.io/>
- YouTube channel: <https://video.daos.io/>

■ 6th DAOS User Group (DUG'22)

- Recordings available at <https://dug.daos.io/>

■ Intel landing page

- <https://www.intel.com/content/www/us/en/high-performance-computing/daos.html>

