HPC·AI
ADVISORY COUNCIL
NETWORK OF EXPERTISE

OpenFOAM
Performance Benchmarking and Profiling

July 2020

Note

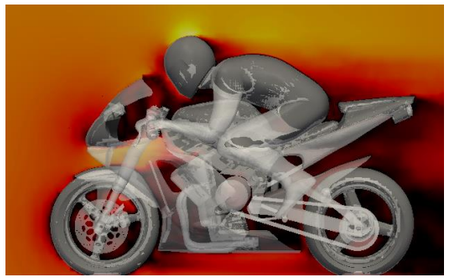
- **The following research was performed under the HPC Advisory Council activities**
  - Compute resource - HPC Advisory Council Cluster Center

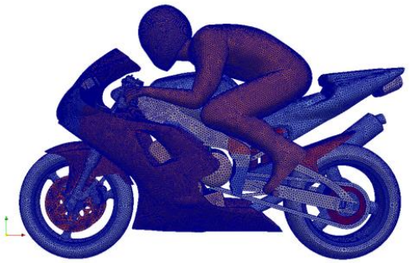
- **The following was done to provide best practices**
  - OpenFOAM performance overview over Intel based platforms
  - Understanding OpenFOAM communication patterns

- **More info on OpenFOAM**
  - https://www.openfoam.com/

OpenFOAM

- **Toolbox in an open source CFD applications that can simulate**
  - Complex fluid flows involving
  - Chemical reactions
  - Turbulence
  - Heat transfer
  - Solid dynamics
  - Electromagnetics
  - The pricing of financial options





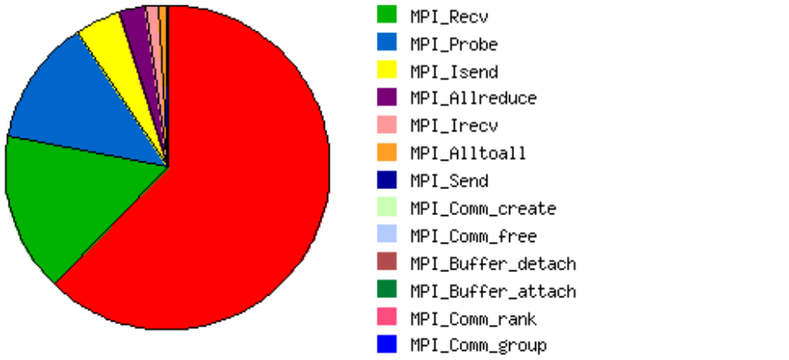
Cluster Configuration

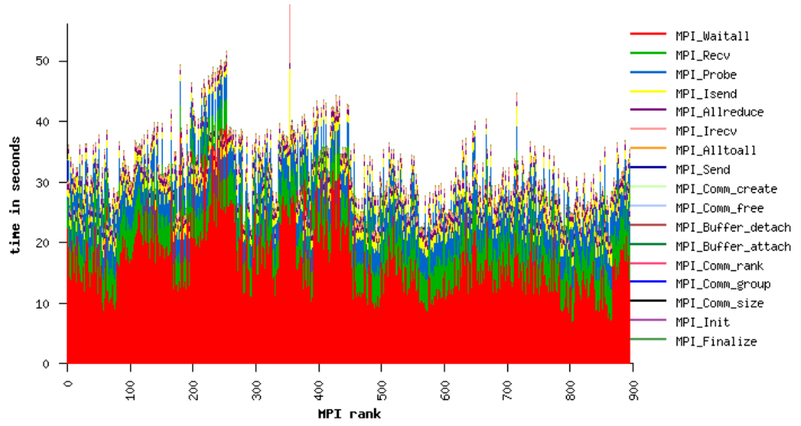
- **Helios cluster**
  - Supermicro SYS-6029U-TR4 / Foxconn Groot 1A42USF00-600-G 32-node cluster
  - Dual Socket Intel Xeon Gold 6138 CPU @ 2.00GHz
  - Mellanox ConnectX-6 HDR100 InfiniBand
  - Mellanox Quantum Switch HDR InfiniBand
  - Memory: 192GB DDR4 2677MHz RDIMMs per node
  - Lustre Storage

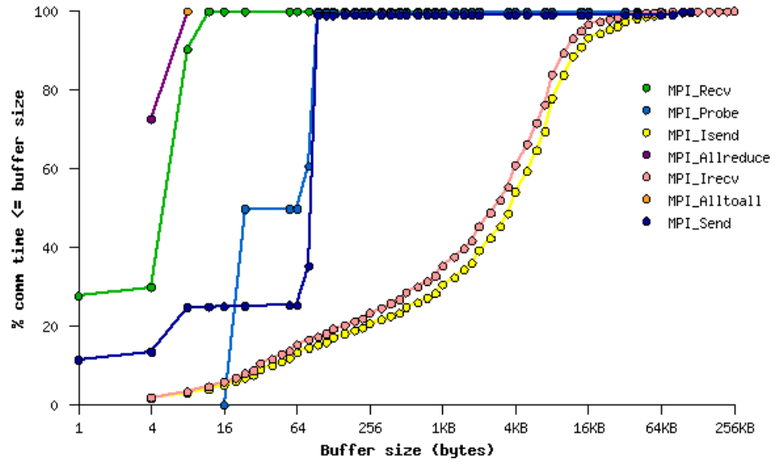
- **Software**
  - OS: CentOS 7.7
  - Driver: MLNX_OFED 4.7
  - OpenFOAM Version: v1912
  - Input: MotorBike_160
  - IO: Lustre/Local Disk
  - MPI: HPC-X 2.6.0/Intel MPI 2019 u7

OpenFOAM Profiling – MPI Time

- **MPI profiler shows the type of underlying MPI network communications**
  - Majority of communications occurred are non-blocking communications
- **Majority of the MPI time is spent on non-blocking communications at 32 nodes**
  - MPI_Waitall (11% wall), 8-byte MPI_Recv (1.4% wall), 1-byte MPI_Recv (0.7% wall)
  - Only 14% of the overall runtime is spent on MPI communications at 32-nodes

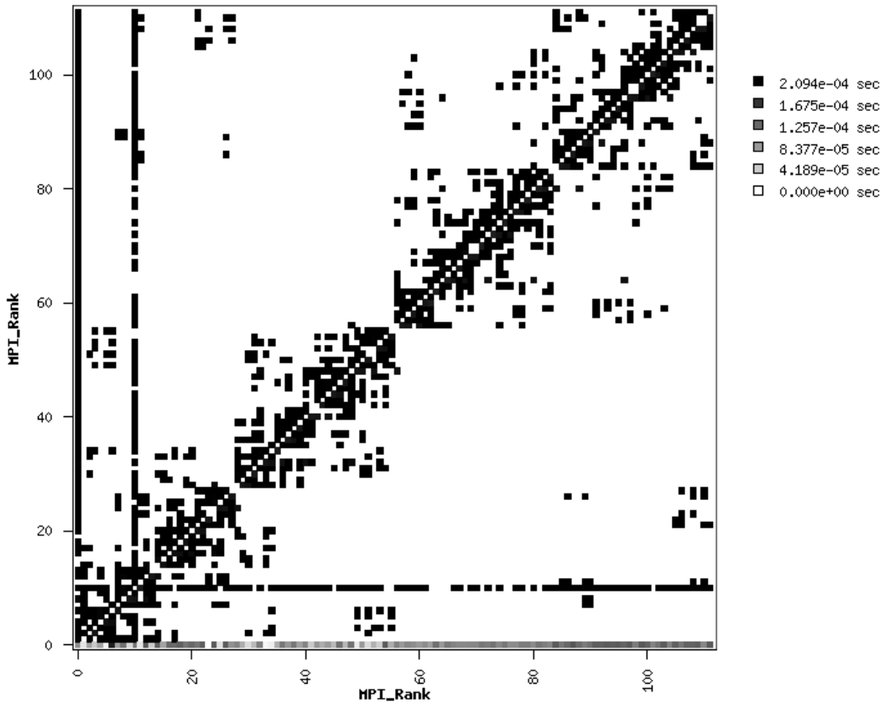




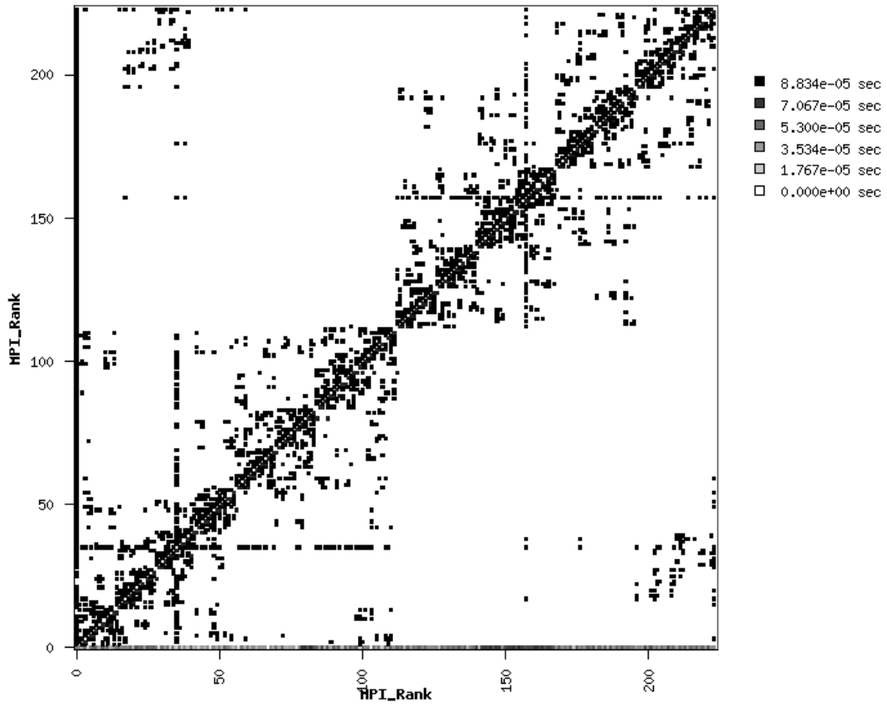


OpenFOAM Profiling – MPI Communication Topology

- Communication topology shows communication patterns among MPI ranks
- MPI processes mainly communicates with neighbors, but also shows some other patterns



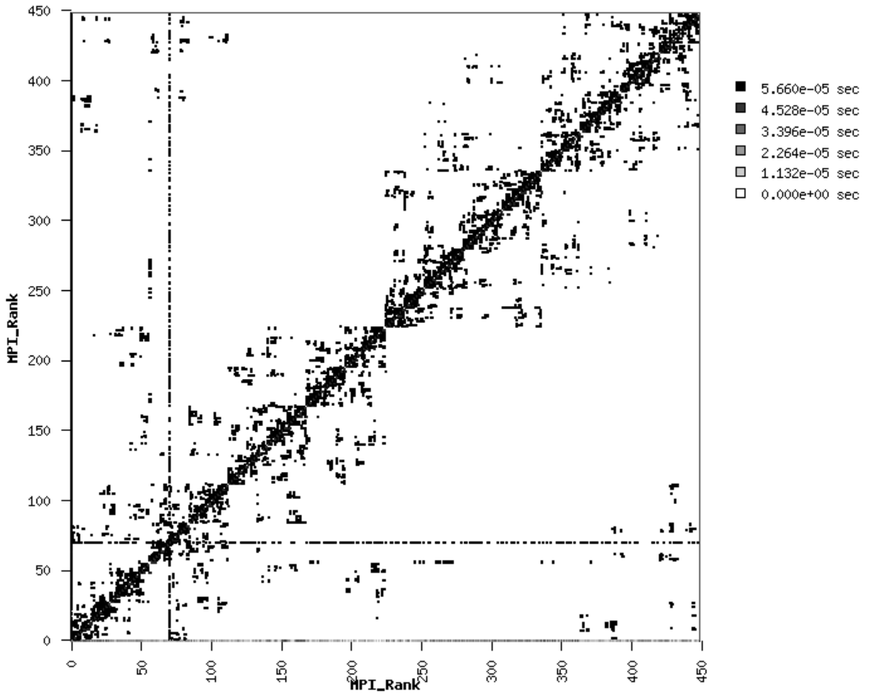
4 Nodes



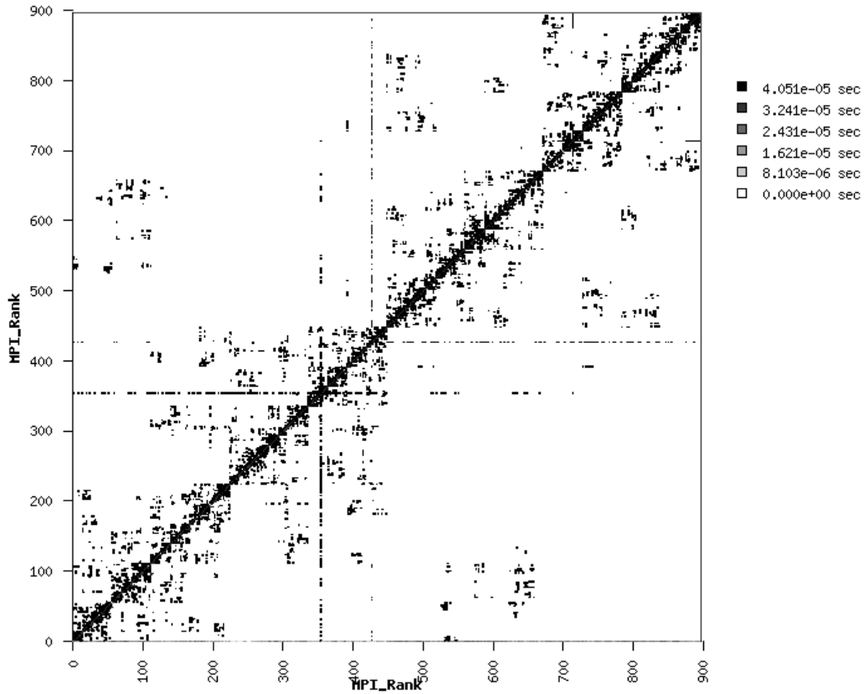
8 Nodes

OpenFOAM Profiling – MPI Communication Topology

- Communication topology shows communication patterns among MPI ranks
- MPI processes mainly communicates with neighbors, but also shows some other patterns



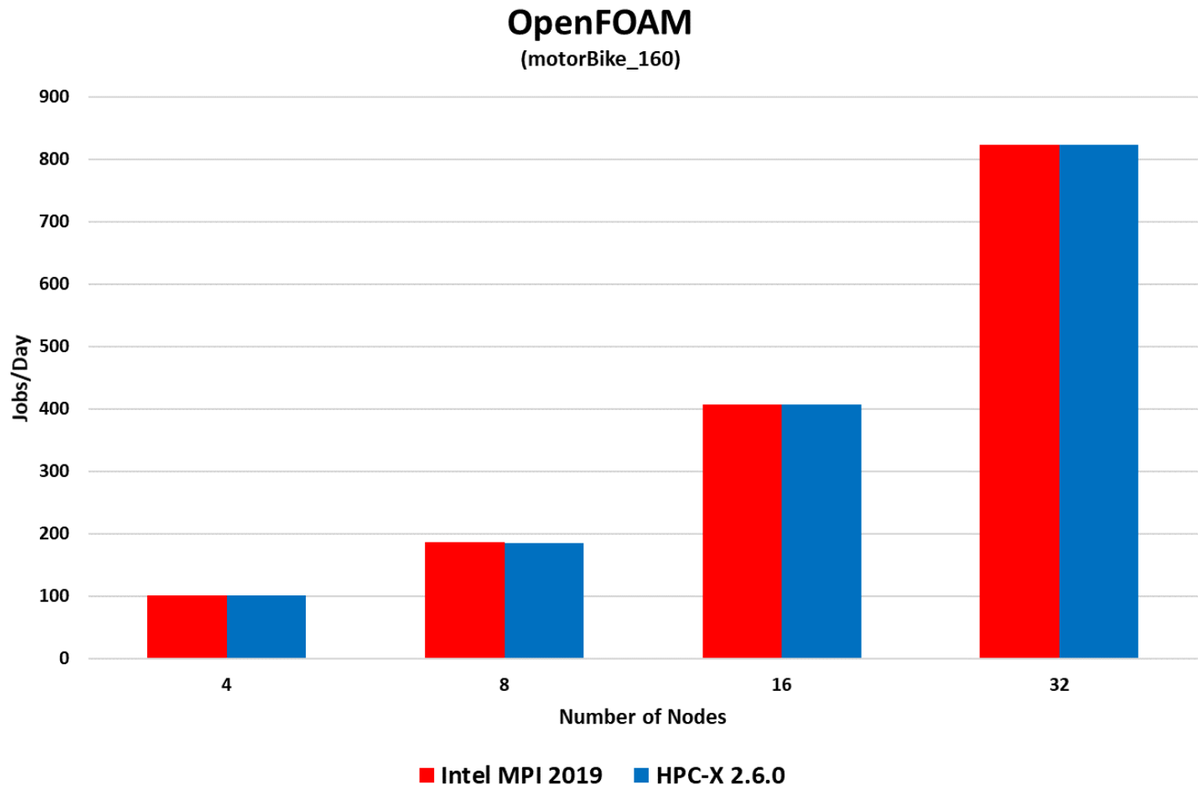
16 Nodes



32 Nodes

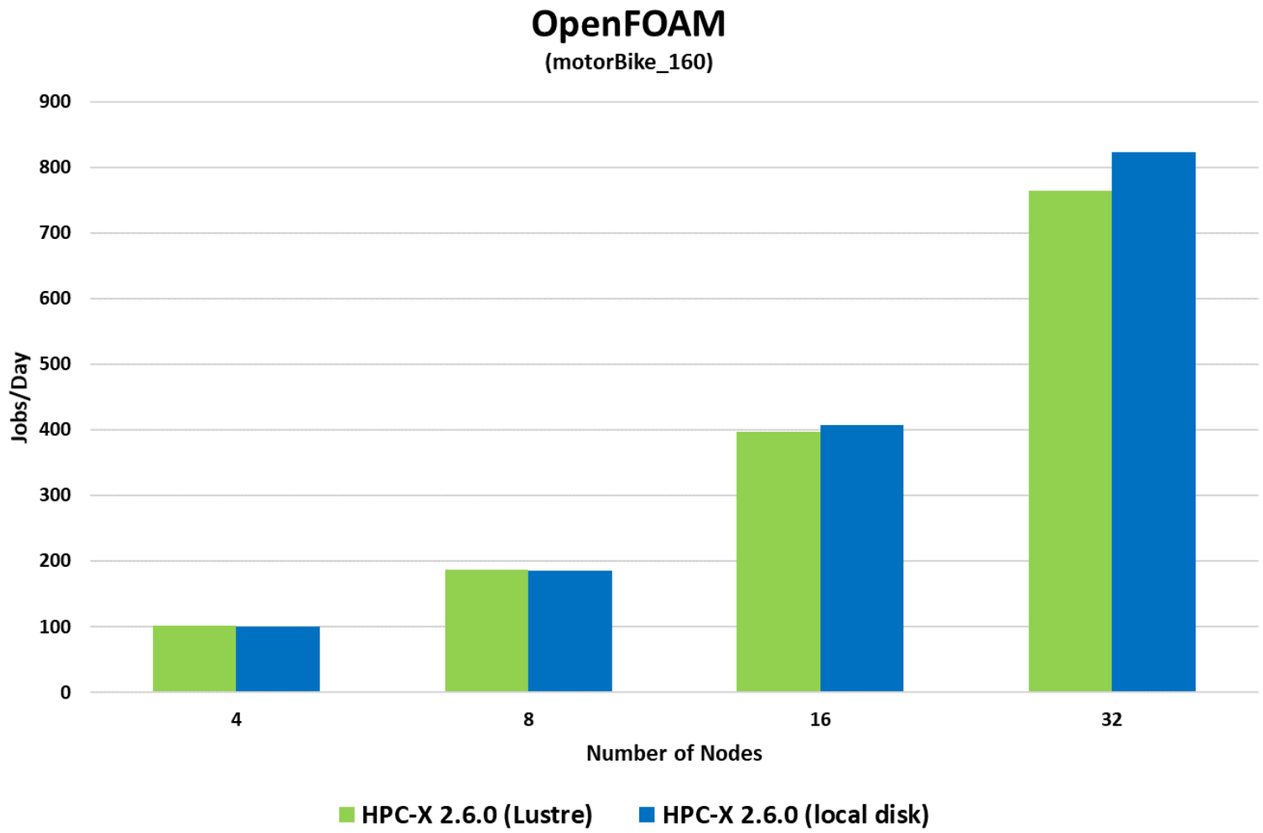
OpenFOAM Performance – MPI Comparison

- Intel MPI and HPC-X MPI demonstrate similar performance
- OpenFOAM demonstrates linear scalability



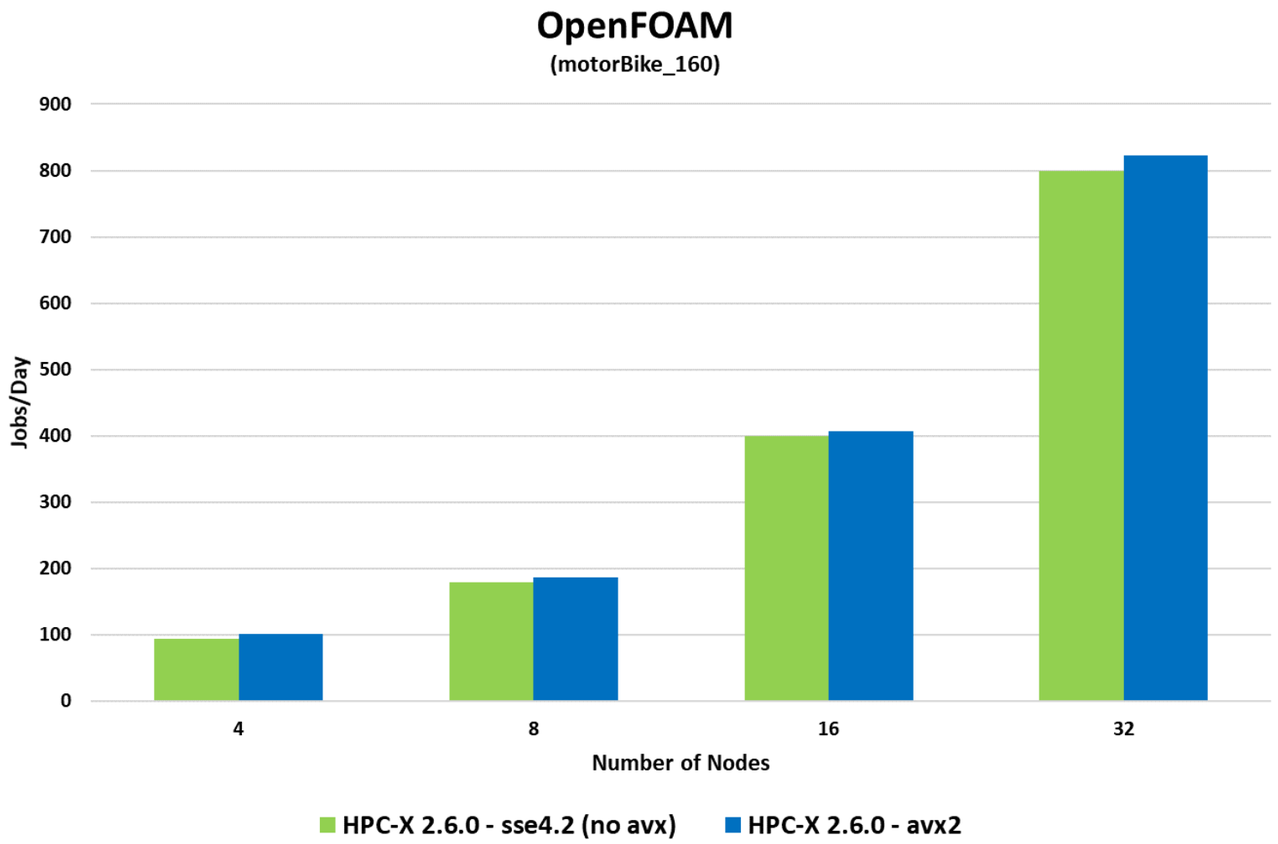
OpenFOAM – IO Comparison

- **With Local storage OpenFOAM demonstrates 8% higher performance at 32 nodes**



OpenFOAM – AVX Comparison

- **OpenFOAM showcase 3% higher performance at 32 nodes with AVX**



Summary

- OpenFOAM imposes high demands on the cluster interconnect
- Intel MPI 2019 u7 and HPC-X 2.6 use the same UCX library from the UCF (Unified Communication Framework) consortium, and demonstrate similar performance
- Enabling AVX2 for OpenFOAM had 3% advantage over SSE4.2 (No AVX)
- OpenFOAM running mounted to local disk demonstrated 8% higher performance versus Lustre

HPC·AI ADVISORY COUNCIL
NETWORK OF EXPERTISE

Thank You