ARGONNE LEADERSHIP COMPUTING FACILITY SITE UPDATE

DAVID MARTIN – DEM@ALCF.ANL.GOV

THANKS TO SUSAN COGHLAN – SMC@ANL.GOV

IXPUG Annual Meeting – October 14, 2020

ALCF RESOURCE

- Mira – IBM Blue Gene/Q – Deployed in 2012, Decommissioned in 2020
- Theta – Cray XC40 - Xeon Phi, Deployed in 2017
- ThetaGPU –NVIDIA DGX A100, Deployed in 2020
- Aurora – Cray Shasta - Intel Xeon with GP-GPUs, Coming Soon!

- Joint Laboratory for System Evaluation – lots of cool hardware

THETAGPU COMPUTE NODES

COMPONENT	PER NODE	AGGREGATE
AMD Rome 64-core CPU	2	48
DDR4 Memory	1 TB	24 TB
NVIDIA A100 GPU	8	192
GPU Memory	320 GB	7,680 GB
HDR200 Compute Ports	8	192
HDR200 Storage Ports	2	48
100GbE Ports	2	48
3.84 TB Gen4 NVME drives	4	96

JLSE HARDWARE AND SOFTWARE

Argonne's Joint Laboratory for System Evaluation provides testbeds for Aurora – available today [under the appropriate NDAs]

- Intel Xeons with Gen 9 Iris Pro Graphics (integrated) (20 nodes) [no NDA required]
- Intel Xeons with Intel development GPU cards (code named DG1) [requires a CNDA with Intel]
- Intel's Aurora SDK that already includes many of the compilers, libraries, tools, and programming models planned for Aurora [requires a CNDA with Intel]

AURORA HIGH-LEVEL CONFIGURATION (PUBLIC)

System Spec	Aurora
Sustained Performance	≥1EF DP
Compute Node	2 Intel Xeon scalable processors (Sapphire Rapids) 6 Intel X^e arch based GP-GPUs (Ponte Vecchio)
GPU Architecture	X^e arch based GPU (Ponte Vecchio) Tile based, chiplets, HBM stack, Foveros 3D integration
CPU-GPU interconnect	PCIe
Aggregate System Memory	>10 PB
System Interconnect	Cray Slingshot Dragonfly topology with adaptive routing
Network Switch	25.6 Tb/s per switch, from 64 - 200 Gbs ports (25GB/s per direction)
High-Performance Storage	≥230 PB, ≥25 TB/s (DAOS)
Programming Models	Intel oneAPI, OpenMP, DPC++/SYCL
Software stack	Cray Shasta software stack + Intel enhancements + Data and Learning
Platform	Cray Shasta
# Cabinets	>100

AURORA SOFTWARE AND TOOLS (PUBLIC)

Area	Aurora
Compilers	Intel, LLVM, GCC
Programming languages and models	Fortran, C, C++ OpenMP 5.x (Intel, Cray, and possibly LLVM compilers), UPC (Cray), Coarray Fortran (Intel), Data Parallel C++ (Intel and LLVM compilers), OpenSHMEM, Python, Numba, MPI, OpenCL
Programming tools	Open|Speedshop, TAU, HPCToolkit, Score-P, Darshan, Intel Trace Analyzer and Collector Intel Vtune, Advisor, and Inspector PAPI, GNU gprof
Debugging and Correctness Tools	Stack Trace Analysis Tool, gdb, Cray Abnormal Termination Processing
Math Libraries	Intel MKL, Intel MKL-DNN, ScaLAPACK
GUI and Viz APIs, I/O Libraries	X11, Motif, QT, NetCDF, Parallel NetCDF, HDF5
Frameworks	TensorFlow, PyTorch, Scikit-learn, Spark Mllib, GraphX, Intel DAAL, Intel MKL-DNN

AURORA I/O

- Storage will include Distributed Asynchronous Object Storage (DAOS) and Lustre components
  - Greater than 230 PB of storage capacity
  - Greater than 25 TB/s of bandwidth
- User see single namespace which is in Lustre
  - "Links" point to DAOS containers within the /project directory
  - DAOS aware software interpret these links and access the DAOS containers
- Data resides in a single place (Lustre or DAOS)
  - Explicit data movement, no auto-migration
- Users keep:
  - Source and binaries in Lustre
  - Bulk data in DAOS
- Applications codes will not need changes in most cases