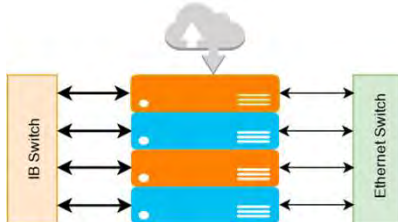


Landon Clipp, Navid Mokhesi, Jaqueline Liang,
Antonio Abinader, Roshan Rajan, Chaitanya Sindagi

Architecture

Overview

- Architecture designed to take advantage of NVLink-connected GPUs.
- Optimized for inter-node data throughput.



Hardware

Component	Description	Count
Chassis	Cray CS-Storm 500nx 1U Server	4
CPU	Intel Xeon E5-2680 v4 @ 2.4GHz	8
Memory	16 GB DDR4 @ 2666MHz (x12/node)	48
GPU	Nvidia Tesla V100 SXM2	8
Storage	960 GB 2.5in SATA SSD	4
Interconnect	Mellanox ConnectX-5 EDR InfiniBand	4

Table 1: Hardware configuration

Software

- CentOS 7.5:** Widely-used OS for cluster computing.
- CUDA 9.2:** Development environment for high performance GPU-accelerated applications.
- SLURM:** Widely-used open-source cluster management and job scheduling system.
- NCCL2:** Performance-optimized multi-GPU and multi-node collective communication primitives.
- Intel 2017 Cluster Suite:** Tools for code development on a distributed memory system.
- TensorFlow-GPU:** Open source GPU-enabled software library for high performance numerical computation.
- Intel-RAPL and Nvidia-SMI:** Tools for measuring and controlling cluster's power usage.
- HDF5-Parallel:** Library for parallel I/O.
- OpenMPI 2.1.5:** Open-source high performance Message Passing Interface (MPI) Library.
- GCC:** the GNU Compiler Collection for compiling applications.
- Arm Forge and Performance Reports:** Tools for application profiling.

Applications

HPL & HPCG: The competition's benchmarking software. In order to excel in this category we tested and tuned our cluster, including CPUs, GPUs and even DRAM, to maximize the power efficiency and achieve the desired results.

Horovod: A distributed training network developed by UBER for TensorFlow, Keras, and Pytorch. To excel in this category we tuned our cluster to maximize its power efficiency by tuning down the CPU clock. We also tuned the installation for NV-Link enabled GPU-to-GPU communication.

SeisSol: A tool for computational earthquake dynamics. We configured and compiled the tool to take advantage of Intel AVX-512 vector instruction set extensions to maximize utilization of Skylake CPUs.

Application	Intended Runtime Schedule	Run Target
HPL	Mon: 9am-11am	Cluster
HPCG	Mon: 11am-1pm	Cluster
OpenMC	Mon: 9pm-finish	Cloud
Horovod	Tues: 1am-10am	Cluster
SeisSol	Tues: 10am-8pm	Cluster
Mystery App	Tues: 8pm-finish	Cluster

Table 2: Application runtime schedule

OpenMC: Monte-Carlo simulation tool for nuclear physics. Used to simulate nuclear reactors. CPU-only code, good candidate for running on the cloud.

Mystery Application: To excel in the mystery application category, it is critical to understand the underlying architecture as well as how to tune system to maximize performance.

Cloud: Utilizing cloud resources in conjunction with cluster resources is the strategy to maximize competition performance. We tested different CycleCloud configurations to maximize performance and minimize costs.

Strategy

Competition preparation

- Thoroughly study applications, profile them to understand performance bottlenecks.
- Work with sponsors to create best possible system configuration, tune it for best performance and optimal power usage
- Learn how to use various tools to profile and optimize for performance.

Team recruitment

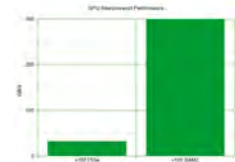
- Create team that represents wide range of expertise and backgrounds.
- Ensure mix of veteran members and newcomers.
- Recruit understudy members to participate in competition preparation to learn and grow into possible future participants.

Vendor sponsorship

- Team solicited sponsorship from industry leaders.
- Worked with NCSA to identify and approach potential sponsors.
- Worked with sponsors to specify system requirements, identify and solve problems.
- Sponsors provided expert support with system installation and performance tuning.

New technology

- Use V100 SXM2 GPUs with NV-Link.
- Use latest IB EDR interconnect technology.



Dynamic power control

- Power control through dynamic power monitoring.
- Clock down CPU and GPU as needed.
- Build power monitoring system in order to visualize power behavior for all applications.

Practice environment

- Use secondary cluster to perform experimental tests.
- Perform competition dry runs to iron out issues.
- Consult with experts to resolve configuration issues.

Diversity and Collaboration

Diverse team

- Diverse levels of experience:** Team includes freshman, sophomore, junior, and senior.
- Mix of newcomers and veterans:** Some members were previous participants of SCC and others are new to the competition.
- Understudies:** Recruited understudy members to learn. Possible participants in future competitions.
- Female participation:** Recruited from Women in ECE (WECE) and Women in CS (WCS)
- Wide range of interests:** Atmospheric Science, System Architecture, Phylogenetic Trees, Cloud Computing, Machine Learning, Economics, Markets, Autosport racing.
- Industry and academic experts:** Provided input to team for building architecture, system tuning, system management, and team management.

Industry sponsorship

- Cray**
 - Provided servers and server rack.
 - Expert support for system deployment.
- Nvidia**
 - Provided V100 GPUs.
 - Expert support for GPU software.
- NCSA**
 - Provided space for hardware.
 - Expert advisors in various domains.
- Jump Labs**
 - Provided hardware for previous competitions, now used for practice.
- Blue Waters Project Office**
 - Provided financial support for the team
 - Expert advisors and guidance.
- Mellanox**
 - Provided InfiniBand cards, switches and cables.

Diverse backgrounds & interests

- Landon Clipp – Senior**
 - Big data, parallel programming, scientific computing, atmospheric science, system administration, AWS
 - OpenMC
- Navid Mokhesi – Junior**
 - CUDA, AWS, Big Data, Systems
 - Passion for Reproducibility
 - SeisSol, Horovod
- Jaqueline Liang – Junior**
 - Phylogenetic trees, HPC, CUDA
 - Horovod, OpenMC
- Antonio Abinader – Senior**
 - NCSA, Cloud computing, Machine learning, system administration
 - HPL, HPCG
- Roshan Rajan – Senior**
 - Economics, Big data, Distributed Systems, AWS
 - Cyclecloud, SeisSol
- Chaitanya Sindagi – Sophomore**
 - System administration, FSAE Electric captain
 - HPL, HPCG