Texas A&M Team Aggie

Sheldon Wei[1,2], Michael Lau[2,3], Dylan Rodriguez[1,2,3], Brian Kelley[2,5], Sean Dormiani[4,5], and Nathan Mandell[2]
Advisor Dr. Jian Tao[1,6,7],

[1]High Performance Research Computing, [2]Department of Computer Science and Engineering, [3]Mays School of Business, [4]Department of Electrical and Computer Engineering, [5]Department of Mathematics, [6]Texas A&M Institute of Data Science, and [7]Texas Engineering Experiment Station

Texas A&M University, College Station, TX

SC18 Dallas, TX | hpc inspires.
Student Cluster Competition

TEXAS A&M UNIVERSITY

Introduction

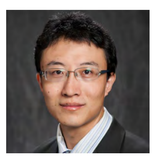
The Team

The Texas A&M Student Cluster Competition team is comprised of students from Texas A&M University's College of Engineering with backgrounds in high performance computing, computer science, electrical engineering, business, and mathematics. The students come from all levels: first year to fourth year. Our team members are pursuing a variety of majors and minors in different colleges and outside work experience. The team was formed via AggiE-Challenge, a special topic course for engineering undergraduates. In an attempt to create a diverse team, membership was open to any student regardless of their field of study which resulted in a team with a collection of varied majors and focuses.

Sheldon Wei
Computer Science
OpenMC / SeisSol

Michael Lau
Computer Science
HPL / HPCG

Dylan Rodriguez
Computer Science
Horovod / OpenMC



Jian Tao
Team Advisor

Brian Kelley
Applied Mathematics
SeisSol / Mystery Application

Sean Dormiani
Computer Engineering
OpenMC / Horovod

Nathan Mandell
General Engineering
Mystery Application

Enhancing Diversity

- Members of the team possess cultures from four continents: South America, North America, Europe, and Asia
- Advisors of all genders helped the team
- Members have a healthy mix of career interests post-graduation
- Two team members were instructors and counselors for Texas A&M's Summer Computing Academy for K-12 students

System Configuration

Hardware

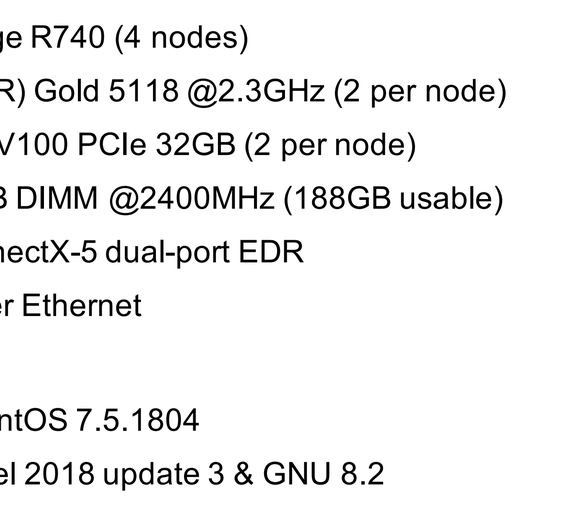
System	Dell PowerEdge R740 (4 nodes)
CPU	Intel(R) Xeon(R) Gold 5118 @2.3GHz (2 per node)
GPU	NVIDIA Tesla V100 PCIe 32GB (2 per node)
Memory	Dell 12 x 16GB DIMM @2400MHz (188GB usable)
Interconnect	Mellanox ConnectX-5 dual-port EDR
I/O	Linux NFS over Ethernet

Software

Operating System	CentOS 7.5.1804
Compiler	Intel 2018 update 3 & GNU 8.2
Resource Manager	N/A
Power Monitor & Control	Intel Data Center Manager (DCM)
GPU Runtime	CUDA 9.0 & 9.2
MPI	OpenMPI 3.1 & Intel MPI 2018

Driving Factors - Power and Performance

- The hardware was chosen in consultation with our vendor sponsors and with regards to feasibility of power management in addition to hardware accessibility.
- The Dell PowerEdge R740 server provided an optimal platform. The versatility of the platform created a suitable environment for testing multiple configurations using 250W cards.



- Our partnership with Dell Technologies provided a knowledge base for the R740 platform.
- Configuration of previous championship team's systems was also taken into consideration.
- A heterogeneous architecture that included GPUs was prefered owing to the team's familiarity GPUs and CUDA.
- The Xeon Gold 5118 provides 12 cores and the high-performance AVX-512 instruction set but uses only 105W.
- The Tesla V100, one of the fastest GPUs available provides 7 teraflops (double-precision) at 250W.
- GPUs will support Horovod, HPL, and HPCG.
- CPUs will support SeisSol, OpenMC.

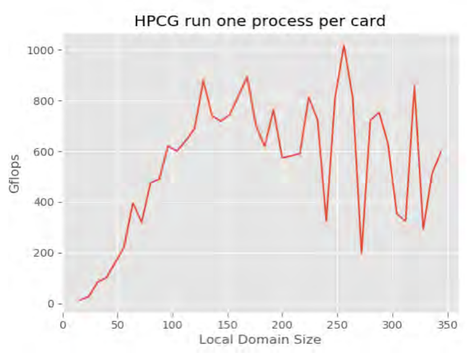
Why We Will Succeed

- System administrators and User Support specialists at TAMU High Performance Research Computing (HPRC) offered expert guidance.
- TAMU HPRC provided access to on-campus computing resources, like the Ada and Terra clusters for preliminary application testing in the early stages of the competition.
- Dell Technologies, NVIDIA and Mellanox provided resources to prepare for the competition.
- Working in teams created redundancy with application knowledge, ensuring everyone was knowledgeable in multiple applications.
- Two team members work at Texas A&M High Performance Computing Research.
- The team receive support from the Texas A&M AggiE-Challenge program.
- Team members gained invaluable experience from Asia Supercomputing Community Student Challenge.

Applications & Optimization Strategies

HPL & HPCG: Most widely used cluster performance benchmark suites.

- Used NVIDIA-provided binaries for HPL and HPCG benchmarks optimized with CUDA 9.2
- Automated HPCG runs to determine optimum input parameters
- Intel's DCM provided accessible power management, particularly during HPL runs
- With 8 Tesla V100 HPCG does not reach the power cap



Horovod: A distributed deep learning framework which improves efficiency of communication between nodes.

- Designed to accommodate growing problem sizes which require multiple GPUs
- Relatively easy to distribute training across GPUs
- NCCL and cuDNN libraries contribute to accelerated computation
- Optimization testing included manipulation of hyperparameters until reaching an ideal speed power consumption ratio
- Easy-to-implement checkpointing ensures minimal data loss in the event of disruption

OpenMC: An open-source tool for running simulations of neutron systems in various geometries and calculating the criticality of these systems using Monte Carlo techniques. After testing many different builds, we decided on using OpenHPC modules with Infiniband support.

- Studied Monte Carlo simulation to better understand the parameters
- Made custom scripts for easy execution (loading modules, activate venv) and preserving the output

Mystery Application: Preparation included referencing applications in this category from previous years. As team members worked on the listed applications, we maintained notes on the build process. These documentation processes, along with previous experience building software on Linux distributions contribute to our confidence with regards to the Mystery Application. The versatile platform provides the highest likelihood of achieving optimal performance during the "mystery application" challenge.

Checkpointing: Checkpointing during runs lessens the risk of data loss.

- Saves time in event of disruptions like loss of power
- Horovod and SeisSol have checkpoining built in
- Application state is periodically written to disk so it can be resumed later

The Reproducibility Challenge

SeisSol: Software package for simulating wave propagation and dynamic rupture on arbitrary high-order derivative discontinuous Galerkin method.

- Used Intel C/C++ compilers (produces faster code than GCC)
- -O2 optimization flag yielded higher performance
- Enabled AVX512 instruction set to fully utilize Skylake microarchitecture
- Demonstrated strong scaling on the sample problem TPV33, across 96 cores
- Expected to produce a report confirming the SC17 paper about SeisSol's performance

Acknowledgements

Special thanks to Francis Dang, Jack Purdue, Ping Luo, Dhruva Chakravorty, Rick McMullen