Dieut et mon Droid

TACC
TEXAS ADVANCED COMPUTING CENTER

Mathew Abraham, Surendra Anne, Jorge Bolivar, Saiprathik Chalamkuri, Jenna May, Benjamin Nederveld

TEAM INTRODUCTION

The UT Austin SCC team was founded in 2010 and has been mentored by the Texas Advanced Computing Center. We are a competitive team from a wide variety of backgrounds and majors.

Our backgrounds range from physics to electrical and computer engineering. This diverse set of skills allows us to tailor our approach to each application or benchmark and to meet challenges with insight on both the technical aspects of the hardware and the scientific nature of the applications.

Mathew Abraham: Computational Engineering student, worked on HPL

Surendra Anne: Physics and Math, worked on HPCG

Jorge Bolivar: Computational Engineering student, worked on PHASTA

Saiprathik Chalamkuri: Computer Science student, worked on IO500

Jenna May: Electrical and Computer Engineering student, worked on the reproducibility challenge

Benjamin Nederveld: Math and Chinese student, worked on LAMMPS

SYSTEM DESIGN AND MANAGEMENT

Hardware Configuration:

- Head node - x1
  - Dell PowerEdge R750
  - Intel Xeon Platinum 8362 - x2
  - 512 GB RAM
  - Nvidia Ampere A100 - 1x
- Compute nodes - x7
  - Dell PowerEdge R750
  - Intel Xeon Gold 6338 - x2 per node
  - 256 GB RAM per node
  - Nvidia Ampere A100 - 1x per node
- High speed fabric network
  - Rockport LS24T 24 Port Lower Shuffle
  - Rockport NC1225 Network Card (200Gbps) - x8
- Storage
  - Liqid Element LQD4500 PCIe AIC Composable Storage SSD 4TB (/scratch)*
  - 480GB SSD SATA RAID 0 (/home)
- Liqid Composable Infrastructure Platform*
  - LQD400x08P4 Expansion Chassis - 1x
  - Liqid Grid 24 Port Gen 4 Fabric Switch - 1x
  - LQD1416 PCIe Gen 4 Host Bus Adapter - x8

Software configuration:

- CentOS 8 derivative Operating System
- Ansible for change-management
- PXE booting for provisioning
- Intel and GCC compilers
  - Intel MPI and MKL
  - CUDA SDK 11.7
- Rockport Autonomous Network Manager
- Liqid Matrix CDI

Cloud configuration:

- H-series VM's in CycleCloud will allow for fast data processing
  - H16r SKU includes RDMA, which offers high speed fabric
- NCv3-series would provide Nvidia Tesla V100 GPUs if our A100 GPUs are insufficient.

*Based on availability and testing

intel CentOS NVIDIA

APPLICATION STRATEGIES

IO 500:

- Use of NVMe storage hardware to achieve optimal best possible IO performance. Plan to exploit stripping of multiple Liqid Element NVMe devices and therefore expect to achieve around 26 GB/s throughput per drive hosted in our cluster
- Investigation into a full parallel Filesystem (such as Lustre) is unnecessary for our small drive-count

PHASTA:

- Primarily focus on running on CPUs and prepared to allocate partially to the cloud if necessary
- MPI build of application to run across nodes. Additional builds with compressible and/or incompressible solvers paired with optimized supported libraries

LAMMPS:

- Primarily focus on running on GPU hardware and add use of available CPUs for any non-GPU enabled packages
- Research different packages for an optimized build per various scenarios

Reproducibility:

- Run benchmarks and frameworks on both CPU and GPU based nodes
- Use of Intel CPUs, as in the paper, will help with producing similar results
- Runs are relatively short, allowing a shift of some runs to cloud as funds allow.

HPCG:

- Utilize optimal # of GPU nodes parallely for efficiency
- Nvidia A series' high-bandwidth memory allows for faster processing

HPL:

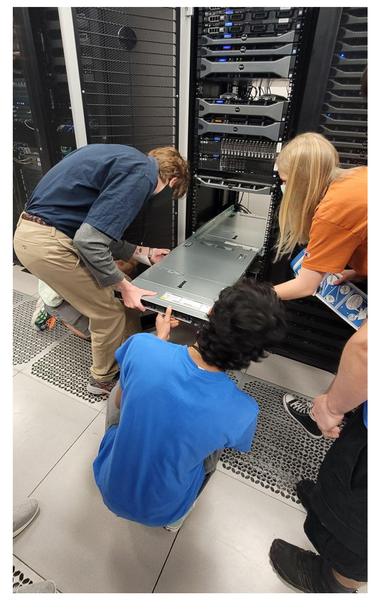
- Utilize Nvidia pre-built binary to run most optimally on GPU architecture
- Run HPL on GPUs with Liqid composable hardware platform
- Tune HPL.dat file parameters and experiment with MPI/affinity

Mystery Application:

- Analyze source code to determine whether to run primarily on CPU nodes or GPU nodes
- Use our team's varied experience to best approach scaling and scheduling

PREPARATION

- Take advantage of TACC resources to gain early understanding of HPC environment
- Practice our software management through TACC machines to emulate our finished nodes and connections
- Active communication with team on weekly basis via virtual and in person meetings
- Team members have primary and secondary app/benchmark to ensure even support among team



- Stress test applications on TACC systems to gain a better understanding of most optimal way to run our codes
- Split cluster and cloud administration among team members to ensure adequate support throughout competition

Why We Will Win

The support from TACC and our sponsors has allowed us to be ambitious with our cluster design. We are prepared with a competitive base cluster along with nontraditional hardware such as Rockport and Liqid that will prove advantageous.

While we are new to HPC we've needed to tackle every step along the way ourselves from scratch. So, we both enjoy and welcome any challenging problem that is presented.