Final Architecture Proposal
Team UIUC, SCC 19

Hardware

The University of Illinois cluster consists of four Cray CS-Storm 500NX servers with EDR IB interconnect. Its technical details are provided in Table 1. The team has graciously received a donation from Cray for 4 CS-Storm 500NX servers and a donation from Nvidia for 16 V100 GPUs. The cluster also uses an EDR IB switch and adapters previously donated to the team by Jump Trading.

Component	*Name*	*Count*	*TDP per item*
Chassis	Cray CS-Storm 500nx SXM2 1U	4	
CPU	Intel Xeon 6148 2.4GHz 20 Core	8	160W
Memory	16GB DDR4-2666 (12 per node)	48	<6W
GPU	Nvidia Tesla V100 SXM2	16	300W
Storage	1.6 TB Intel DC P4610 SSD	8	15W
	Crucial MX500 1TB SSD (/home)	1	5W
	HyperX Predator M.2 PCIe G2 x4 256GB SSD	4	2W-8W
Interconnect	ConnectX-5 InfiniBand EDR Adapters	4	10W-20W
	SB7890 InfiniBand EDR Switch	1	122W

Table 1: System Configuration

Our cluster is designed to maximize the utilization of GPU accelerators. We believe this design is suitable for the challenge since it has a very desirable performance-to-watt ratio, which is the primary metric to track in a power-limited environment. We have multiple GPUs to accelerate the applications that can utilize them, and the 20-core Xeon CPUs will serve as a capable compute platform for applications that only use CPUs.

Our measurements show that a single node consumes about 350W when idle, and has an

estimated TDP of 1750W. Therefore our major task during the competition is to manage power via CPU/GPU frequency scaling and other power reduction techniques. We intend on underclocking many of our components, namely the CPUs and GPUs, in order to keep our power consumption under the allotted 3,000 watts.

Software

All nodes in our system are running CentOS 7 Linux kernel version 3.10. CentOS 7 was chosen for its widespread industry support and relative ubiquity as an enterprise system. Both our industry sponsors and our advisors have found it to be a stable and reliable platform to run large scale compute jobs. The OS is installed on M.2 SSDs and /home is exported from 1TB SSD on the head node. */scratch* is built across 8 1.6TB SSDs as GPFS.

Intel Parallel Studio XE 2018 Cluster Edition provides all major tools and libraries (compiler, profiler, MPI, MKL, etc.). To maintain compatibility, the GCC compiler and OpenMPI's older versions are installed on the system, however newer versions of OpenMPI and MVAPICH are also installed for testing. NVIDIA CUDA Toolkit is deployed to support the use of GPUs. Arm Forge and Performance Reports tools are used for debugging, profiling, and optimizing code.

The Slurm scheduler is installed to support optimal resource utilization. Team members can either directly access compute nodes from the head node via ssh, or obtain time on the nodes via sbatch. Slurm was chosen due to the team's familiarity with the software and its widespread use in modern HPC systems.

In conclusion, we believe this software stack is suitable for the challenge because it includes the best possible support for the hardware on a set of system software with a proven track record.