WARSAW TEAM

INTERDISCIPLINARY CENTRE FOR MATHEMATICAL AND COMPUTATIONAL MODELLING,
UNIVERSITY OF WARSAW, POLAND

At Student Cluster Competition,
ISC 2018, Frankfurt, Germany

TEAM INTRO

In our team we have members from three different universities: University of Warsaw, Warsaw University of Technology and University of Wrocław. We all study Computer Science but Maciej and Łukasz are doing a joint degree in Maths and CS.

Dominik is our system admin, deals with all hardware-related problems and manages the basic software stack. Maciej and Łukasz have a strong Maths background. Additionally, Łukasz takes part in National Team Algorithmics Championships and recently he earned the fifth place. Adam is the team captain and manages the team work while also making use of his experience in deep learning. Julia apart from her vast technical skills, has outstanding verbal skills and makes use of her creative abilities. Jakub is new to the team and is the youngest team member but he learns really fast. Resposibilities of each team member were:

- Dominik – hardware and OS administration, HPL and HPCG,
- Łukasz – OpenMC and Azure Cloud,
- Maciej – SeisSol (mainly tasks running),
- Julia – SeisSol (mainly paper analysis) and poster,
- Adam – work management and Horovod,
- Jakub – Horovod.

We know that our hardware will not win the competition for us. We put hard work into competition preparations to ensure that we perform as good as we can. We hope that this effort connected with experience coming from diverse backgrounds – doing different majors and studying at different universities offering different curriculums – will earn us a score that we can be proud of.

HARDWARE AND SOFTWARE STACK

HARDWARE

Our hardware consists of servers supplied by Texas Advanced Computing Center, 8 GPUs sponsored by NVIDIA and a unique networking solution: FabreX provided by GigaIO.

We have four nodes, each equipped with:

- **Dell PowerEdge R740** – supports all the needed hardware,
- **2x Intel Xeon Gold 6126 12 Core CPU** – versatile and performant Intel x86 platform; enough CPU cores per GPU ensuring no data transfer bottlenecks,
- **192 GB DDR4 RAM** – enough to handle the most demanding tasks,
- **2x NVIDIA Volta V100 32GB PCI-E** – state-of-the-art V100 GPU Accelerators,
- **GigaIO FabreX PCI-E Inteconnect** – innovative solution, allowing to reduce latencies and overheads, our hope to gain an edge in the competition,
- **1TB Intel SSD** – lightning fast storage for all the competition tasks.

SOFTWARE

Our software stack configuration is an effect of our competition experiences and requirements of the applications. The most important building blocks are:

- **CentOS 7** – industry-standard operating system,
- **GCC 7.x Compilers** & **Intel Parallel Studio 2018** – compilers and libraries bringing out out the best possible results from the applications and interconnect,
- **Slurm 17.11** – workload manager allowing us to run the computations through the 48-hour competition with minimal supervision,
- **MPICH 3.3** – compiled with additional support for our FabreX Interconnect.

CLOUD RESOURCES

In order to excel in the competition we had to decide where we should run the computations. Using cloud resources is an integral part of the competition. We decided that **we would run OpenMC in Azure Cloud**. Running SeisSol in the cloud is forbidden by the rules and our setup seems perfect for running Horovod (many nodes with multiple GPUs). In addition, GPU nodes in the cloud are much more expensive, so using CPU-only nodes would be a more efficient use of our quota. Another reason for that decision is that we plan to run SeisSol on CPUs, so running OpenMC at the same time would be definitely impossible or inefficient. Of course, there is the Mystery Application to consider, but that is something we cannot prepare for.

MEET THE TEAM

ADAM SOBECKI

ŁUKASZ KONDRACIUK

JAKUB KOWALSKI

MACIEJ SZPINDLER
team advisor

DOMINIK PSUJEK

JULIA BAZIŃSKA

MACIEJ KORPALSKI

MARCIN SEMENIUK
team advisor

COMPETITION PREPARATION AND STRATEGIES

GENERAL PREPARATIONS

Key elements of our preparations for the Student Cluster Challenge were:

- weekly meetings for sharing tasks updates,
- research about optimizations possibilities,
- working in application subgroups,
- support from team members who were not able to go with us to SC18,
- consulting experts (geophysics, MPI),
- cooperation with GigaIO to ensure that their beta-phase interconnect solution works with the competition applications,
- sharing and discussing insights from the mailing list.

Our cluster was not available for us before the competition, so we have to prepare all of the software from scratch during the preparation days. We only had access to a similar setup to run tests.

POWER MANAGEMENT

Before the contest we will run power usage tests, mostly focusing on the GPU applications, because the GPUs are the most power-draining part of our setup. We estimate that we can reach up to 3.5 kW with the all of the GPUs running on 100%. Thus we will power cap the GPUs and manually control the CPU clocks as needed. Disabling Turbo Boost on the CPUs would make their power usage more predictable in case we would be on the verge of exceeding the limit.

SEISSOL

The SeisSol challenge is focused on reproducing the results, not just achieving the maximal performance. Our aim was to fully utilize our cluster resources. Our main activities before the competition were:

- ensuring we have the required knowledge to run SeisSol tasks,
- run examples from the paper and the mailing group,
- test out PUMgen for mesh repartitioning,
- discuss our task with geophysics professor from Polish Academy of Sciences,
- analyze results from the paper and predict whether we would achieve similar results on our setup.

OPENMC

OpenMC uses Monte Carlo randomized algorithms to simulate movement of elemental particles. During the preparations we:

- attempted to port OpenMP Offload Directives to run on GPU – failed due to OpenMC's parallelization on a high level (we would need to rewrite tremendous amounts of code),
- experimenting with various compilation options and different MPI processes to OMP thread ratios (both on our testing cluster and Azure Cloud),
- automated building dependencies and OpenMC with a bash script allowing us to conduct experiments faster.

HOROVOD

The Horovod task is the only one task related with deep learning and neural networks training. Our goal is to make the process of training as fast as possible and maintain the precision and training accuracy.

- we want to utilize all of our 8x NVIDIA Tesla V100 GPUs and the optimal number of CPU cores to maintain the best possible performance and keep the rest of the cores for another application.
- we have done tests on different servers and with the Slurm Workload Manager.
- we are ready to test different hyperparameters depending on the final problem we will get (we are well prepared for the tuning and adjusting it for our cluster),
- we compile TensorFlow and Horovod from sources with GPU support with compute capability matched with our GPUs.

SPONSORED BY

SERVERS PROVIDED BY