TUMany segFAUlts

HIGH PERFORMANCE COMPUTING@RRZE

GCS Gauss Centre for Supercomputing

SPPEXA

Our Team and Universities

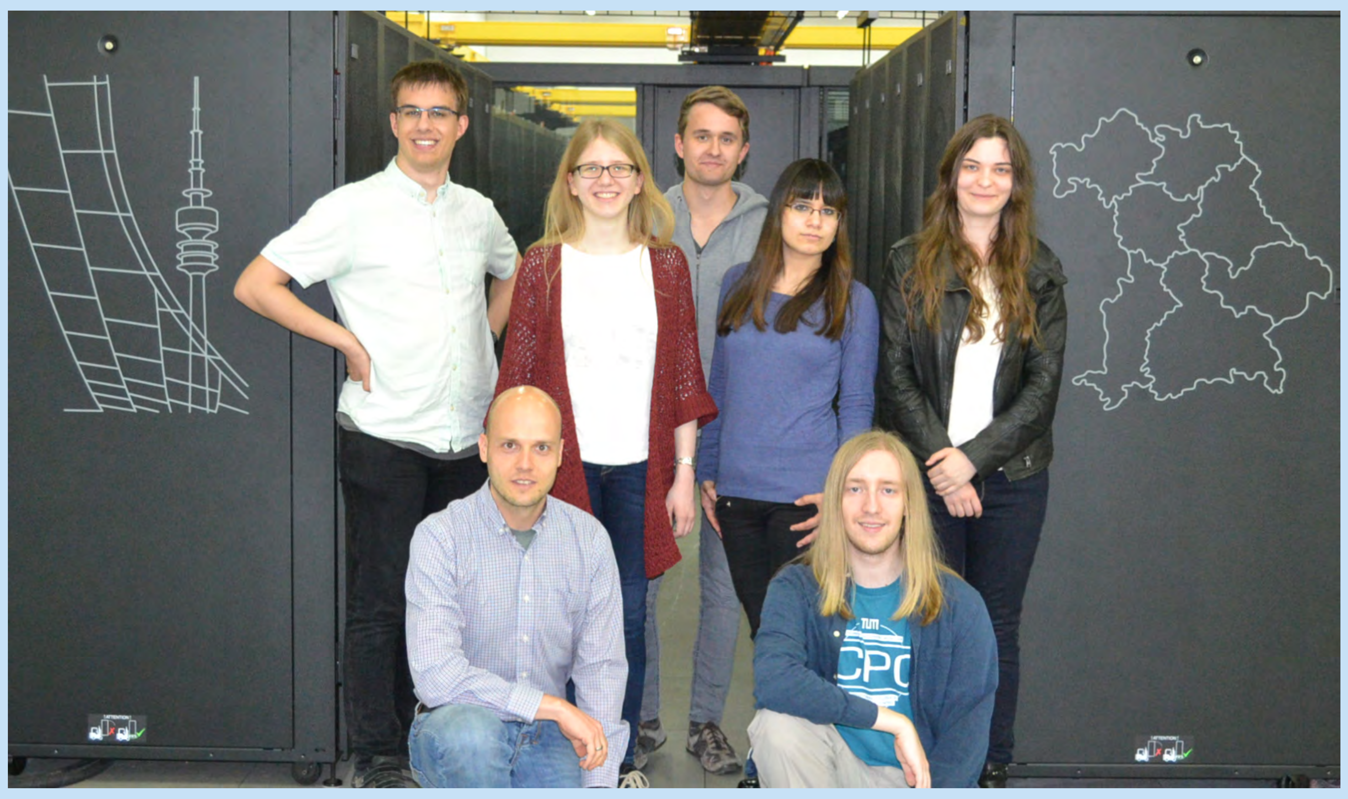
Our team is composed of students of the two greatest Bavarian universities, four from the Friedrich-Alexander-University Erlangen-Nuremberg (FAU) and two from the Technical University of Munich (TUM). We are very keen on integrating different cultures and minorities into our team. All our team members were handpicked and asked to participate in this competition with diversity being one of our criteria.

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Technische Universität München

Three of our team members are female, one of whom is Latina. Furthermore, we got one student from Bulgaria who studies abroad in Bavaria. Three of our team members are the first in their family to go to university, despite the fact that only 23% of people with non-academic parents in Germany do so. Since the majors IT, Computational Engineering and even the minor archaeology are represented in our team, we bring a very diverse set of skills with us.

Both universities offer great courses to their students, the TUM in particular has a strong focus on technical subjects and its computer science program is regularly ranked as one of the best in Europe. The TUM offers access to the LRZ (Leibniz Supercomputing Center) housing some of the fastest supercomputers in Europe. The FAU routinely sends its students to international competitions like the SCC, the ASC and the NWERC. Also of international importance are some of the inventions made by our/both universities e.g. MP3 format, B tree and refrigerator.



From left to right, David Schneller, Lisa Dreier, Benedikt Schröder, Ingrid Münch, Ivonne Rausch, Alexander Ditter, Sharru Möller

Preparation

All of the team members hailing from the FAU were new to HPC. In order to get them caught up we started preparations by having everyone choose basic software topics (e.g. SLURM), read up on them and demonstrate to the rest of the team how this software is used in HPC. After these presentations everyone felt ready to tackle the applications. So the TUM students were invited to the FAU campus, for a week long practice session during the summer holidays, where we all met in person and set up our cluster Adam.
Now, work on the applications could begin for real. All in all we had three applications, two benchmarks and the cloud computing component to contend with. To be more flexible when dealing with problems at the competition we chose to assign three students to each of these topics, making sure that each group includes participants from both universities.

From that point onwards we were working on the applications together, communicating mainly via Slack and Discord. Of course, we also had weekly meetings with the students and supervisors in which we could show our progress, work on the applications, and discuss our problems.

Software

We chose the CentOS7 linux distribution as our operating system. It provides an rpm-based package system and is compatible with RedHat and all its supporting tools and drivers. Also the administrators of our university clusters are already familiar with this operating system. HPE gave us a trial license for Insight CMU which will be very useful, because it enables managing and monitoring the nodes remotely. Additionally, we will use the agent-less configuration management tool Ansible. Through writing playbooks it will enable us to propagate software and settings automatically to all nodes in the cluster. As a batch management system we decided to use SLURM, because some of our team members were already experienced in its use. The last software aspect to mention is Intel MPI, which makes sense to use on Intel processors. It will provide the cluster with low latencies and a high communication throughput.

Strategy

At the start of the competition we will be running Tersoff on our local cluster and MrBayes on the cloud. Ideally we will have the results of the Tersoff run a few hours after the beginning of the competition, giving us plenty of time to write the reproducibility report. We chose against having Tersoff run on the cloud as we would rather not risk getting inaccurate results due to using VMs for measuring scalability.
MrBayes is well suited to being run in a cloud environment as it has small input and output files meaning that the lower bandwidth of a cloud won't affect it too badly. It also scales almost linearly with the number of available processors which means we can achieve good performance without GPUs (which at the time of this writing are not available in the cloud, yet). Additionally it would not be difficult to migrate MrBayes to the local cluster in case the cloud budget runs out.
Once we have the necessary data for the reproducibility report we will start running Born on the local cluster. Born profits heavily from GPUs and processes large amounts of data leading us to believe that it is best served running on our local cluster with its 12 GPUs and Infiniband connections. Also, Born is run in the unit of 'shots' each of which should take only about 30 minutes meaning that we won't loose too much progress when the poweroff happens.

The Mystery Application isn't known to us at this point, so we'll have to adapt our strategy on the fly. At first we will have to compile it and have a close look at its input and output, which can be done while another application is running on our cluster. After that we will decide if it's better to run the Mystery Application on the local cluster or on the cloud, depending on the application, the cloud's budget and the progress of the other applications. We plan on having small teams for singular tasks like compiling, looking at the input and the applications uses. Members of these teams will be chosen by their experience in fields relating to the application and the amount of time they have available while working on their other tasks.

Hardware

HPE Apollo 6500 (iLo 4)

2x HPE ProLiant XL270d nodes with respectively

Power Intel ® Xeon ® E5-2690v4 2,6GHz	135 W
256GB DDR4 RAM	
6x NVidia V100 GPUs	6x 250 W
800GB SSD	

Interconnect: 40 Gbit Mellanox Infiniband EDR

Power consumption is an important issue in this competition and has influenced quiet a few of our hardware choices.
We decided to bring six NVidia V100 GPUs per node, giving us 12 GPUs in total. Earlier this year this setup provided the best FLOPS/power ratio for the P100 GPUs we were using at the/that time, leading to the FAU team achieving the best HPL performance at the SCC@ISC 2017. The main reason for us to have 512GB total RAM is to be prepared for the mystery application, which might require that much memory. We chose to forgo having cold storage on HDDs, instead taking 2x800GB SSDs, in order to save more of the power margin for benchmark runs.

The HPE Apollo 6500 allows us to hot-plug individual nodes enabling us to react quickly to hardware problems.The system also comes with integrated Lights-Out 4 out-of-band maintenance using HPE's own Insight CMU.

Why we will win

Our team has a good mix of experienced "veterans" and highly motivated "newbies". As our team members come from two universities we have access to extensive resources and we are being backed by three great advisors. One of them was leader of the FAU-team at SCC@ISC17 (that won the LINPACK there) and another was part of the TUM-team of SCC@SC15 (that won the LINPACK there). We also got support by Chair of Computer Architecture at the FAU and the Chair of Scientific Computing at the TUM.
Having students from different universities and different degree programs means that we have a broad range of skill sets in our team. The universities having different areas they focus their research on is also helpful, leading to a diverse inventory of courses available.
Independent problem solving is a skill that is highly valued and encouraged at both of our universities. We are very confident that we have all the prerequisite competence to handle the mystery application in an efficient manner.