Team SwanGeese

University of Science and Technology of China

A Brief Introduction to USTC

The University of Science and Technology of China (USTC) is a prominent university in China and enjoys an excellent reputation worldwide. It was established by the Chinese Academy of Sciences (CAS) in 1958 in Beijing.

Among domestic universities, USTC has the largest number of research achievements that have been acknowledged by distinguished academic authorities, such as “Top 10 Pieces of News in Sci & Tech in the World”, “The Top Physics Stories of the Year”, “Top 10 Pieces of News in Sci & Tech in China”, “Top 10 News Stories in Basic Research in China”, “Top 10 Research Advances in the Universities in China”, and so on. USTC is the only university in China that has been listed in the “Top 10 Pieces of News in Sci & Tech in China” for the recent eight consecutive years.

Team Information

Hao Zhang, Condensed Matter Physics
- Distributed Password Auditing/Recovery,
- Mystery Application

Jian Zeng , Computer Science
- Distributed Password Auditing/Recovery,
- Mystery Application

Shiming Zhuang , Computer Science
- ParConnect

Professor Hong An, Computer Architecture
- Adviser

Huanqi Cao, Computer Science
- Vice Team Captain,
- System Administrator,
- HPCG,
- ParaView

Dezhong Yang , Computer Science
ParConnect

Siyuan Zhuan , Computer Science
- Team Captain,
- Hardware System Administrator,
- Linpack,
- ParaView,
- Mystery Application

We guarantee our diversity by having one majoring in Condensed Matter Physics and two from School of Gifted Young, where the undergraduate students are all younger than others of the same year by 1 to 4 years. We are from all around the China, from the north-eastern mountains to the bustling city in the south. Also, though 5 out of 6 major in Computer Science, we all have interests in widely spread and highly differed areas.

System Configuration

Software

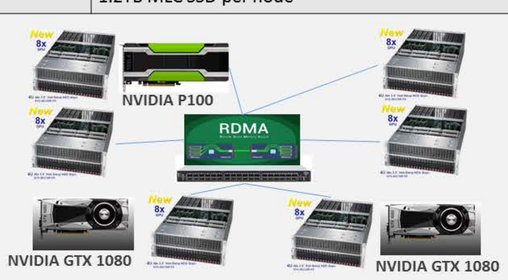
Operating System	GNU/Linux CentOS 7.0, x86_64
Compiler	GNU C/C++/Fortran Compiler 5.2; Intel C/C++/Fortran Compiler 17
Resource Managers	htop 2.0.2; NVIDIA-SMI 367.44
MPI	OpenMPI 1.10.0a; Intel MPI Version 2017
Power Monitor	Power Monitoring Toolkit (Self developed based on IPMI)

Hardware

System	SuperMicro SuperServer SYS-4028GR-TR
CPU	Intel Xeon E5-2695 v4 @2.1GHz, 18 Cores
Total Nodes	6
Total Cores	6 nodes * 2 sockets/node * 18 cores/socket = 216 cores
# of Acc.	8 Nvidia Tesla PCIe P100, 8 Nvidia GTX 1080 GPU
Memory	384GB per node, DDR4 RDIMMs@ 2400Hz
Interconnect	Mellanox EDR 100Gb/s Switch
IO	1.2TB MLC SSD per node

This is a highly heterogeneous system with CPUs, GPGPUs (NVIDIA Tesla P100) for HPC application and GPUs (NVIDIA GTX 1080) for graphic application. We choose two types NVIDIA GPU as accelerator, due to the high performance and power-efficiency of the new Pascal Architecture: Tesla P100 guarantees general purpose computing performance, and GTX 1080 for those OpenGL-based graphic applications, like ParaView. Also, we configure GTX 1080 on two nodes, one node with RAID to improve the ability of handling TB scale sized data.

We adopt Mellanox Infiniband EDR 100Gb/s Switch as system interconnection, which high bandwidth and RDMA technology can provide efficient communication.



Why We Will Win?

Team Work
- Each two work on one application
- Architecture trade off through all apps
- Weekly talk to train communication ability

Learn from Scientists
- ParaView: NAOC
- ParConnect: DOE JGI
- Password Recovery: IIE CAS

Work with Engineers
- Power Management
- GPU Utilizing
- Linpack & HPCG

Tools for Profiling
- Intel VTune Amplifier
- Allinea Profiler
- CUDA Profiler(nvprof)

Spirit
- We love, We enjoy!
- No perfect, but better!
- Never Give Up!

Benchmarks and Applications

Benchmarks: Linpack & HPCG

To show the computing ability of modern devices, we choose NVIDIA GPGPU as our main computing platform. Getting support on the newest Tesla P100 PCI-e version, and having tested with the help of engineers from NVIDIA, we are confident to have a good performance in both Linpack and HPCG.

This year we have HPCG as a new benchmark for scoring the performance. As a model matching a broad set of modern applications, it requires much higher memory speed; and the CoWoS HBM2 Stacked Memory on Tesla P100 will highly satisfy this kind of requirement by HPCG, as well as those applications in real world.

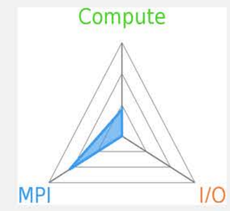
ParConnect

Competition Task

To analyze and recreate the graphs and tables in the publication “**A Parallel Connectivity Algorithm for de Bruijn Graphs in Metagenomic Applications**” using our own cluster.

Preparation for the Competition
- Consult and communicate with Professor Zhong Wang from JGI. Reading related publications and learning background knowledge under professor's instruction;
- Reading and understanding the source code;
- Building and running the application on different cluster with the original dataset and data extracted from bovine and ovine stomach. Profiling the application using Intel Vtune and Allinea Profiler.

Profiling and Understanding of the Application
- The application is compute and communication intensive.
- After the paralization, the application shows scalability and the task can be accelerated with the increasement of CPU cores.
- When the sequencing depth of input data is not enough, the application would be easily influeced by the error in the sequence and generate a result which can't be used in further research.



ParaView

Case Study

As a visualization application, we first focused on rendering workload. We’ve tested along CPU (OSMesa), GPGPU (NVIDIA Tesla series), and graphic oriented GPU (NVIDIA GeForce series).

Besides, Getting datasets from NAOC, we’ve done with input data of diffirent sizes, range from medium to large (20 ~ 700 GB). On the right there is a sample image of a partially rendered large dataset.

Strategy1 - Why using GTX 1080?

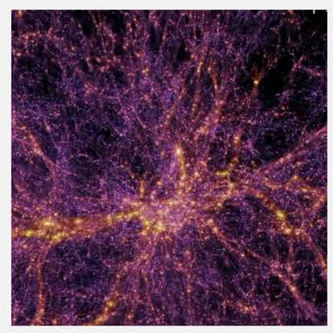
After testing with several rendering ways including CPU, GPGPU and GPU for graphic work, we find that GPUs designed for graphic workload like NVIDIA GeForce has much higher performance when using OpenGL, as ParaView does. Also, cards like Pascal Titan X by NVIDIA may have even better performance on this.

Strategy2 - Overcome Memory and I/O Bottleneck

For small to medium sized datasets that will fit into main memory, we can fully utilize the rendering performance of GTX 1080; for larger ones, we’ll use RAID of multiple SSD drive to speed up the I/O work.

More Over Consideration - NVIDIA IndeX Plugin for ParaView

For handling volume type data elegantly and efficiently, we introduce NVIDIA’s IndeX plugin. This is a plugin developed for volume data rendering and filtering in ParaView by NVIDIA, and it is powerful enough for realtime rendering on some simulation results.



Distributed Password Auditing/Recovery

Software Design Philosophy
- Dynamic computing devices control
  During the competition, the remaining power when other application running will be used to work on this task.
- Data saving in time
  All the password will be saved at the time recovered, so it is easy to resume after an unexpected shutdown.

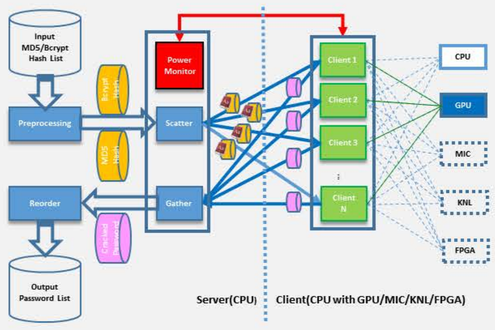
Software Architecture Design
- Client/Server architecture
  Server side distributes and arranges data for client side; client side gets data from server and starts to recover. This architecture make the application elastic and stable.
- Run on different platform
  With the highly flexible client/server architecture, it is possible to combine any different devices together, such as CPU, GPU, FPGA, and even MIC or KNL.

Strategy1 Well-Chosen dictionary

Our dictionary is well-chosen to get higher possibility to recover more passwords during the 48 hours.

Strategy2 Computing Filter

Besides our well-chosen dictionary, we also easily filter significant useless computation during competition owing to the software design philosophy and flexible software architecture which will minimize consumption of time and system resources.



Strategy3 Why using GPU not FPGA?

As for pure hashing performance, we have evaluated and concluded that FPGA > GPU > CPU, and both the former two show 10~100 efficiency than CPU. However, some flexible strategy can be easily used with GPUs but not FPGAs, and as a computing platform FPGA is not so compatible like GPU, so we choose GPU as an accelerator for this task at the last minutes.

Acknowledgements

Mellanox TECHNOLOGIES
NVIDIA
SUPERMICRO
蓝海创意云 www.vsochina.com
PARATERA 并行
AMAX 超集信息科技有限公司 www.amaxchina.com
JGI JOINT GENOME INSTITUTE DEPARTMENT OF ENERGY
NAOC