科学与工程计算中心
Center for Computational Science and Engineering

SUSTech and SST

Hi, this is greeting from Southern University of Science and Technology (SUSTech) in Shenzhen, China. We are SUSTech Supercomputing Team(SST), and this is our first trip in SC Student Cluster Challenge. We consist of extraordinary undergraduates from mixed majors, enabling us respond to challenges from varied fields. Every member of us is experienced in supercomputing challenges and competed at least in two sessions of international supercomputing challenges. We won the first prize in ASC19 and many other prizes. Apart from contests, in daily practices, SST members are devoted in computational scientific researches and coding for revolutionary works. Meanwhile, members actively contribute to the maintenance of campus clusters, in which we gained intensive experience in deployment and debug.

SUSTech has optimal supercomputing platforms for students to practice skills in real scene. Two clusters are virtually unlimited available to student for test purposes, and one of them, Taiyi, is nominated as the 127th supercomputer on the TOP500 list. Moreover, other clusters including cryo-EM cluster with 100 Nvidia Tesla V100 GPUs and DGX workstations are also available on demand. Nevertheless, massive user groups in SUSTech is ready to give any instructions on most softwares and their underlying principles. Equipped with ideal computing resources and solid assisting backups, we SST members are fully lock and load for SC20 VSCC.

Technical Scheme

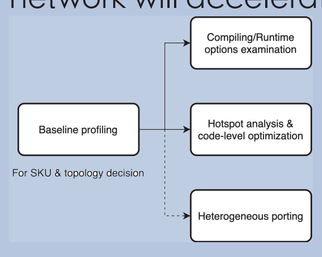
- Plan 1: NC24r x9
  - Hardware-spec
    - CPU: Intel Haswell Core x216
    - GPU: NVIDIA Tesla K80 x18
    - Memory: 2016 GB RAM + 432 GB VRAM
    - FP64 Performance: 8.98 TFLOPS CPU + 52.4 TFLOPS GPU
    - Regular Infiniband with RDMA capability
  - Pros
    - High Theoretical GPU Performance
    - Good at Compute-intensive Workload with GPU
    - High Cost-Effective due to Outdated Architecture
- Plan 2: HB120rs_v2 x4
  - Hardware-spec
    - CPU: AMD Zen2 Core x480
    - Memory: 1920GB RAM
    - FP64 Performance: 18.8 TFLOPS
    - 200 Gbps Infiniband with RDMA capability
  - Pros
    - Latest and Advanced Hardware
    - Centralized Resource and Rapid Intra-net
    - Good at Compute-intensive Workload with CPU
    - Good at IO-intensive Workload
    - High Cost-Effective due to AMD Chips

User Application	User Application
NVIDIA-Docker	NVIDIA-Docker
IBM Platform LSF	IBM Platform LSF
Guest OS	Guest OS
VM Supervisor	VM Supervisor
Azure Cloud Infrastructure	

Table 1 Software Architecture

The first plan is featured with remarkably high theoretical GPU performance. Eighteen NVIDIA Tesla K80 accelerator can provide up to 52.4 TFLOPs peak double-precision performance, while over two hundred Haswell cores at 2.6 GHz with two AVX2 engines per core can reach 8.98 TFLOPs. LINPACK, HPCG Benchmark and GROMACS will be significantly affected by theoretical peak performance while GPU is designed for this kind of application, and NVIDIA has fully optimized these software. Additionally, HPCG and Reproducibility Challenge benefit from large memory bandwidth, while the bandwidth of VRAM on K80 cards reaches 8.66 TB/s overall. 6.05 FLOPs/Byte is relatively ideal compared to the CPU platform with 9-12 FLOPs/Byte.

The second plan focuses on CPU-intensive applications. According to WikiChip, theoretically, the throughput of 480 Zen2 cores at 2.45GHz with two AVX2 engines per core can reach 18.8 TFLOPs. Another highlight of HB120rs model is 200 Gbps Mellanox HDR InfiniBand while the NC24r model only has an ambiguous guarantee of network performance, which means IO-intensive applications like IO500 Benchmark are also fit with this system. As for CESM, this is an extraordinarily complex and heavily coupled software running on distributed CPU platform, so powerful CPU core and high-quality network will accelerate CESM a lot.



The optimization approach depends on applications. For those three benchmarks, we will directly use the highly optimized official version first, and some team members have rich experience in tweaking the arguments of benchmark applications. We will also check the profiling result of these benchmarks generated by Intel VTune or NVIDIA Visual Profiler to adjust the arguments and cluster configuration. For regular applications, we will focus on profiling results which will contain detailed information about memory bandwidth, network activity, cache miss rate, NUMA, process time, vectorization, and then apply suitable strategy such as writing SIMD intrinsics, cache prefetch, utilizing shared memory or Tensor Cores on GPU, modifying compile flags, asynchronous transfer, adjusting precision, replacing dependent libraries with better imple-

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY
地址：广东省深圳市南山区学苑大道1088号
No. 1088,Xueyuan Blvd., Nanshan District,
Shenzhen Guangdong, P.R. China
www.sustech.edu.cn
+86-755-8801 0000