

CUHK(SZ) Supercomputing Group Introduction

Jiekun Yang

The Chinese University of Hong Kong, Shenzhen

October 17, 2021

Agenda

- What is Supercomputing Challenge
- Team Overview
- How We Conduct Our Research/Study
- Miscellaneous FAQ

What is Supercomputing

• Definition

- A supercomputer is a computer with a high level of performance as compared to a general-purpose computer.
- The performance of a supercomputer is commonly measured in floating-point operations per second (FLOPS) instead of million instructions per second (MIPS).
- Since 2017, there are supercomputers which can perform over 10^{17} FLOPS (a hundred quadrillion FLOPS, 100 petaFLOPS or 100,000 teraFLOPS).

What is Supercomputing

- **TH-1A** (2009-2011)

- 32,000 Intel Xeon X5670 Processor
(6-Core on each die, 86,016 cores in total),
7168 NVIDIA Tesla M2050 GPU
- R_{Peak} : 4.7 PFlops
- R : 2.56 PFlops
- Site: Tianjin

- **TH-2** (2013-2016)

- 14,336 Intel Xeon E5-2692v2 Processor
(12-Core on each die), 48,000 Intel Xeon
Phi 31S1P (61-Core on each die)
- R_{Peak} : 54.9 PFlops
- R : 33.86 PFlops
- Site: SYSU, Guangzhou



What is Supercomputing

- **Sunway Taihulight** (2016-2018)
 - DEC Alpha 64 Architecture (RISC)
 - 40960 SW26010 Processor, 1.45GHz, 28nm (260 cores on each die, 10,649,600 cores in total)
 - R_{Peak} : 125.4359 PFlops
 - R : 93.0146 PFlops

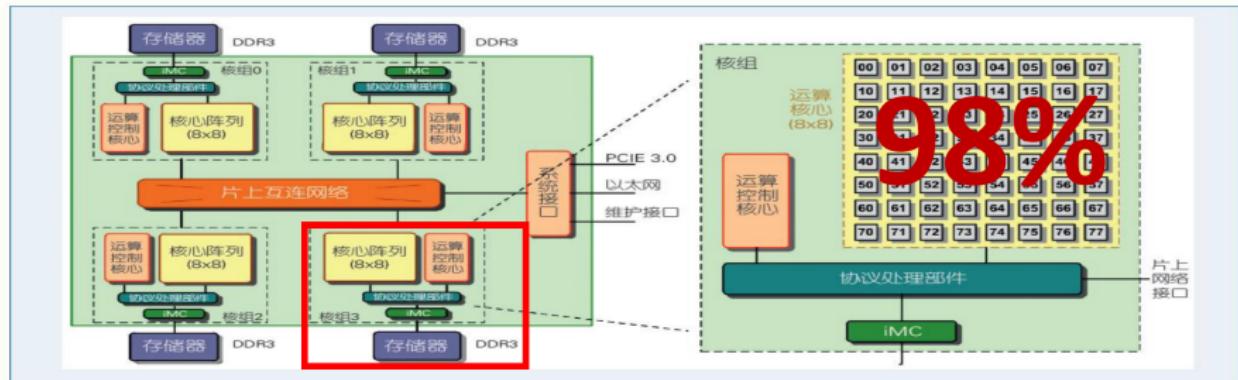


SW26010 处理器性能

主核浮点性能: $1 \times 1.45\text{GHz} \times 8(\text{DP}) \times 2(\text{FMA}) = \text{23.2GFLOPS}$

从核浮点性能: $64 \times 1.45\text{GHz} \times 4(\text{DP}) \times 2(\text{FMA}) = 742.4\text{GFLOPS}$

(浮点性能是指表示每秒钟计算乘法和加法的总次数, FMA表示乘加融合, 把乘法和加法融合为一条指令)



从核优化要点一：充分发挥从核计算性能



国家超级计算无锡中心
National Supercomputing Center in Wuxi

What is Supercomputing

- **Summit** (2018-2020)
 - CPU (POWER9) + GPU (NVLink)
 - 4608 Computation nodes
 - 9216 POWER9 CPU (22 cores on each die)
 - 27648 NVIDIA Tesla V100 GPU
 - R : 200 PFlops



What is Supercomputing

- **Fugaku** (2020-)
 - ARM Architecture (A64 SoC)
 - 158976 computation nodes
 - R_{Peak} : 1.42 exaFlops, 14200 petaFLOPS
 - R : 415 PFlops



What is Supercomputing

- **The Grand Challenges in Computing**

- Physical Simulations
 - N-body Simulation
 - Weather Forecasting
 - Quantum Simulation
 - Molecular Model
 -
- Machine Learning and Deep Learning Tasks
- Cryptography and Cryptoanalysis
-

- **Criteria: Performance and Power**

What is Supercomputing

- Main Challenge: How to accomplish computation tasks as fast as possible
 - Level 1: Implementation level improvement
 - ▷ Reduce the computation complexity, e.g. $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log n)$
 - ▷ Improve data locality (Register, L1/L2/L3 Cache, Memory, Disk, etc.)
 - ▷ Exploit advanced hardware APIs

What is Supercomputing

- Main Challenge: How to accomplish computation tasks as fast as possible
 - Level 1: Implementation level improvement
 - ▷ Reduce the computation complexity, e.g. $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log n)$
 - ▷ Improve data locality (Register, L1/L2/L3 Cache, Memory, Disk, etc.)
 - ▷ Exploit advanced hardware APIs
 - Level 2: System level improvement
 - ▷ Compiler (improve how the compiler compiles high level codes into low level)
 - ▷ Operating System (improve how the OS manages memory and processes)
 - ▷ Architectural Design (improve the implementation of ISA)

What is Supercomputing

- Main Challenge: How to accomplish computation tasks as fast as possible
 - Level 1: Implementation level improvement
 - ▷ Reduce the computation complexity, e.g. $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log n)$
 - ▷ Improve data locality (Register, L1/L2/L3 Cache, Memory, Disk, etc.)
 - ▷ Exploit advanced hardware APIs
 - Level 2: System level improvement
 - ▷ Compiler (improve how the compiler compiles high level codes into low level)
 - ▷ Operating System (improve how the OS manages memory and processes)
 - ▷ Architectural Design (improve the implementation of ISA)
 - Level 3: Hardware level improvement
 - ▷ More CPU Cores, more GPU cards, use InfiniBand, etc.
 - ▷ Equip better cooling system
 - ▷ Tuning on BIOS
 - ▷

What is Supercomputing

- Main Challenge: How to accomplish computation tasks as fast as possible
 - Level 1: Implementation level improvement
 - ▷ Reduce the computation complexity, e.g. $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log n)$
 - ▷ Improve data locality (Register, L1/L2/L3 Cache, Memory, Disk, etc.)
 - ▷ Exploit advanced hardware APIs
 - Level 2: System level improvement
 - ▷ Compiler (improve how the compiler compiles high level codes into low level)
 - ▷ Operating System (improve how the OS manages memory and processes)
 - ▷ Architectural Design (improve the implementation of ISA)
 - Level 3: Hardware level improvement
 - ▷ More CPU Cores, more GPU cards, use InfiniBand, etc.
 - ▷ Equip better cooling system
 - ▷ Tuning on BIOS
 - ▷
 - Level 4: Regression (Perf and restart from Level 1)

What is Supercomputing

- Previous Case in ASC20-21:

- HPL: Linkpack based, solver of a random dense linear system in double precision FP64 arithmetic
- HPCG(High Performance Conjugate Gradients): IO-bounded, model the data access patterns
- TASK 1. Language Exam (LE) Challenge
- TASK 2. Quantum Circuit Simulation (QUEST)
- TASK 3. Pulsar Exploration and Search Toolkit (PRESTO)



What is Supercomputing

- Crucial Knowledge in HPC
 - Being familiar with C/C++ programming(**CSC3002**)
 - Being familiar with Hardware API / Libraries (Assembly, SIMD, CUDA, ...)
 - Basic Knowledge on Data Structures and Algorithms(**CSC3100**)
 - Solid Knowledge on Computer Arch(**CSC3050**), OS(**CSC3150**), Principles of Compilation(**CSC4180**)
 - Parallel Algorithms Design(**CSC4005**)
 - Computer Networks(**EIE4007**)
 - Experience in basic configuration and tools on Linux
 - other domain knowledge

Team Overview

• Team Structure

- Director: Prof. Yeh-Ching CHUNG, Director of CSE Major, Lecturer of CSC3150 and CSC4005
- 6 Alumni, 5 Senior, 2 Junior (most of us are **CSE** major)
- Info-Group: Organizing Group Meeting & Setting up Timeline/ddl;
- AP-Group: Exploring Applications on Cluster & Proposing Training Projects;
- SysOps-Group: Managing Cluster System & Exploring New Configurations and Performance bonus;
- <https://cuhksz-hpc.github.io/>

The collage consists of three images:

- Left Image:** A group of six people in yellow jackets standing behind a table. The table has a white cloth with the "國立清華大學" logo and "National Tsing Hua University" text. In the background, there's a banner for the "SC10 NTHU-Score" competition.
- Middle Image:** A table titled "SC10 NTHU-Score". It displays two tables of results:
 - Top Table:** Application LINPACK RESULTS. It shows HPC Score - percentage of high, High Score, and Percentage. The data includes: MPI-RandomAccess (0.48717, 83.61%), MPI-FFT (49.8039, 83.61%), HPL (1.06821, 93.27%), PTRANS (17.155, 100.00%).
 - Bottom Table:** HPC Points. It shows Task 1 (2, 2), Task 2 (6, 3.0), Task 3 - Ubiquitin A (4, 4), Task 3 - Ubiquitin B (4, 4).
- Right Image:** Five people standing on a stage, holding certificates and awards. They are dressed in yellow jackets, and one person is wearing a purple shirt.

Team Overview

- **Lab and resources**

- Two computing nodes, each with 2x Intel Xeon Silver 4210. 40 CPU cores total.
- Nx NVIDIA Tesla V100 32GB PCIe GPU from ITSO
- Lab Environment



How We Conduct Our Research/Study

- **Objective:**

- ① Participate in Top-tier Student Cluster Competitions
- ② Study on state-of-the-art HPC knowledge (HW/SW APIs, parallel programming paradigms, optimization techniques, etc.)
- ③ Learn how to assemble and configure a cluster system
- ④ (extra) Share valuable ideas and conduct research work

How We Conduct Our Research/Study

- **Target Competitions**

- ASC Student Supercomputer Challenge (East Asia, year-round, 1-5, **has hardware support**)
<http://www.asc-events.org/>
- ISC Student Cluster Competition (Europe, year-round, 6-9, **no hardware support**)
<https://www.isc-hpc.com/student-cluster-competition.html>
- SC Student Cluster Competition (US, year-round, 11-12, **no hardware support**)
<https://sc20.supercomputing.org/>

- **General Procedure of Supercomputing Competition**

- ① Setup a cluster (Configure CPU, GPU, memory, disk, network, OS, etc)
- ② Run 2-3 widely-used benchmarks (HPCC, HPCG, HPL, etc)
- ③ Run and optimize 2-6 HPC applications
- ④ Obtain a score according to given criteria

How We Conduct Our Research/Study

- **Group Activities**

- **Hosts:** Group Members in Campus

- **Topics:**

- ① Tutorial forum associated with CSC4005 projects.
 - ② HPC Case Study (on specific problem or the state-of-the-art implementation)
 - ③ Research Panel (relates to HPC and computer system)

- **Incoming Seminars**

- ① Introduction to basic tech stack(C++, Linux and compiling) (10.17)
 - ② Hands on lab on AWS cloud application
 - ③ Optimizing on the Cache Performance
 - ④ Transformer modeling in natural language processing
 - ⑤

- **Incoming ASC22, Registration Open soon in holidays.**

Misc FAQ

- **Q: Can I get a 4.0 GPA by joining this group?**

A: No. We aim to offer complementary knowledge other than content taught by lectures.

- **Q: Am I be expected to spend considerable amount of time during regular semester weeks?**

A: No, **BUT** we expect your active participation during Competition period.

- **Q: Can I get a girlfriend/boyfriend by joining this group?**

A: No. The current intragroup sex ratio is far from such demand.

- **Q: Can I publish papers by joining this group?**

A: Generally, the answer is no.

Misc FAQ

- **What we could offer:**

- Complementary CS/CE stack knowledge that may be practical in future employment or serve as project experience in your CV.
(computer arch, computer network, parallel programming on heterogeneous systems,)
- Architectural thinking to leverage problem solving and system design.
- Rich opportunity to access state-of-the-art implementations on diverse and active research domains.
(include but not limit to, mathematics, physics, aerospace engineering, quantum computer science, cryptography and cryptoanalysis,)
- Friends and peers who are able to establish collaborative connections.