



Exascale Applications and Software Conference 21st – 23rd April 2015, Edinburgh, UK

Introduction to National Supercomputing Centre in Guangzhou and Opportunities for International Collaboration

Xue-Feng Yuan CPhys FInstP FRSC

National Supercomputer Centre in Guangzhou (NSCC-GZ), Guangzhou, China Email: xuefeng.yuan@nscc-gz.cn Http:www.nscc-gz.cn









A Brief History of NSCC-GZ

The biggest science & technology facility of Guangdong Province and the No.1 science & technology project of Guangzhou - jointly sponsored by MOST, Guangdong Provincial Government, Guangzhou Metropolitan Government, Sun Yat-Sen University with total capital investment of US\$400 millions

Aim at developing Tianhe-2 supercomputer with peak performance over 100 petaflop/sec.

■ NUDT won the contract from Guangzhou Metropolitan Government and the project was initiated in November 2011, with simultaneous construction of the mainframe system, Tianhe-2 building and supporting facilities

In June 2013, the first phase Tianhe-2 supercomputer was ranked world's No.1 with a performance of 33.86 petaflop/sec under the Top500 Linpack benchmark test and retained its title in Nov 2013, June and November 2014, respectively

In October 2013, the first phase Tianhe-2 supercomputer was shipped from Changsha to Guangzhou

In November 2013, "National Supercomputer Centre in Guangzhou" was awarded by MOST

In April 2014, the first phase of Tianhe-2 system started trial service

July – January 2015, Tianhe-2 system was undergone upgrade but still provided limited service





Tianhe-2 Supercomputer (1st Phase)



| Items | Configuration |
|--------------|--|
| Processors | 32000 Intel Xeon (E5-2692 12C) CPUs + 48000 Xeon Phi 31S1P + 4096 FT1500A CPUs . The peak performance is 54.9PFlops |
| Interconnect | Proprietary high-speed interconnection network TH Express-2 |
| Memory | 1.2PB in total |
| Storage | Global shared parallel storage system, 15PB |
| Cabinets | 125+13+24+8=170 compute/communication/storage/service Cabinets |
| Power | 17.8 MW (1.9 GFlops/W) |
| Cooling | Closed Air/Water cooling system |





Software Stack of Tianhe-2 Supercomputer







Key Features of Tianhe-2 Supercomputer

- High performance the No.1 on the TOP500 2013 with a peak performance of 54.9 PFlops
- High energy efficiency 1.9GFlops/W
- Multipurpose heterogeneous architecture for HPC as well as for high throughput and big data
- Easy to use greatly simplified the complexity of application programming with OpenMC
- High availability an autonomous fault tolerant management system





Missions of NSCC-GZ

To facilitate interdisciplinary research on HPC applications and to carry out large scale scientific and engineering projects with strategic importance

National Supercomputer Centre in Guangzhou

To train postgraduates, post-docs and nonspecialists for HPC applications

To explore a selfsustained route for the growth of stateowned organization to meet the needs of HPC market. To integrate industrial sectors with relevant academic communities for accelerating industrial transformation toward high-end of sectors with high added values

To provide high throughput and high security information service and cloud computing applications for development of smart city





Organization of NSCC-GZ



Director and Leadership Team Council and Governing Board

Administration Support (HR, finance, project initiatives and planning, IP, commercialisation, international collaboration, *etc.*)

IT Enabling Support and Training (HPC, Big Data and Cloud Computing, *etc.*)

Materials Science and Engineering Life Sciences and Personalised medicine

Digital Design, Animation and Manufacturing Energy Related Technologies

Earth Science and Environmental Engineering Smart City – e-government, e-education, e-healthcare, e-finance...

End user groups led by PIs from academic and industrial R/D communities



An Integrated HPC & Big Data Research and Innovation Platform





Ε

10⁻³

Length

Ε

10⁻⁹



Immersed Boundary Method Two-fluid Model Microscopic kinetic models + LBM.

Lagrangian-Eulerian-Stochastic Method





GO Solution Bider Colored Colo

Two-fluid Model (SCFT+Reptation) Stochastic Entanglement Dynamics Lattice Boltzmann Methods (LBM) Smooth Particle Hydrodynamics (SPH)





10⁻⁹ sec

Theoretical approach: SCFT Kinetics of signalling and metabolic pathways Coarse-grained Monte Carlo (MC) and Molecular Dynamics (MD) and Non-equilibrium MD Car-Parrinello MD, Quantum MC

Time







An Integrated Ecosystem for Research and Innovation

Experimental Platform Analytical, Measurement and Fabrication Tools to reveal systems dynamics at various time and length scales

Simulation Platform Modularized simulation tools for multiple scale modelling of complex systems with molecular, mesoscopic and continuum dynamics

FULLY INTEGRATED PLATFORM

Big Data Platform Data and text mining, seamless data flow of various types between experimental and simulation platform and cross-correlation analysis tools





Development of Large-scale Virtual Drug Screening Software

Shanghai Institute of Pharmaceuticals and Beihang University

- Molecular docking (virtual screening) is an important tool for drug discovery
- A lot of software has poor scalability and low efficiency
- Drug screening software performance
 - Assessment of affinity for 750K small molecular compounds
 - In half hour using 150K CPU cores in Tianhe-2.









Mapping Gene Regulatory Networks using Million-cores Parallel Computing Georgia Institute of Technology

- The first ever genome-scale approach for construction of gene regulatory networks of Arabidopsis thaliana using Bayesian network structure learning
- Whole genome regulatory networks help us understand gene regulatory mechanisms on a genome-wide scale
- Large-scale CPU/MIC heterogeneous parallel computing
 - 8192 Nodes, 1.6 millions of cores
 - Measured performance: 4.81 PFlops
 - Scientific application with 220 millions of elements runs 3 hours
 - Nearly linear weak scaling, 83% strong scaling parallel efficiency





Arabidopsis thaliana , the first plant to have its genome completely sequenced