



中国科学院超级计算中心  
Supercomputing Center of Chinese Academy of Sciences

# Applications, Software and Supercomputing Environment in SCCAS

---

Xue-bin Chi

SuperComputing Center, CNIC, CAS

August 27, 2010, Hong Kong University



# Outline

---

- **Brief Introduction of SCCAS**
- **Large Scale Applications**
- **Developed Software**
- **Supercomputing Environment**



# SCCAS-Computing Resources

---

## □ During 1996-2000 (9th 5 year plan)

– In 1996, SGI Power Challenge XL

✓ 6.4Gflops

✓ 16 CPUs

– In 1998: Hitachi SRR201

✓ 9.6GFlops

✓ 32CPUs

– **In 2000, Dawning 2000 II**

✓ 111.7Gflops

✓ 164 CPUs



# SCCAS-Computing Resources

## □ During 2001-2005 (10th 5 year plan)

- In 2003, Lenovo DeepComp6800
  - ✓ 5Tflops, 1024 CPUs
  - ✓ TOP500 : No.14 ; China TOP100 : No. 1

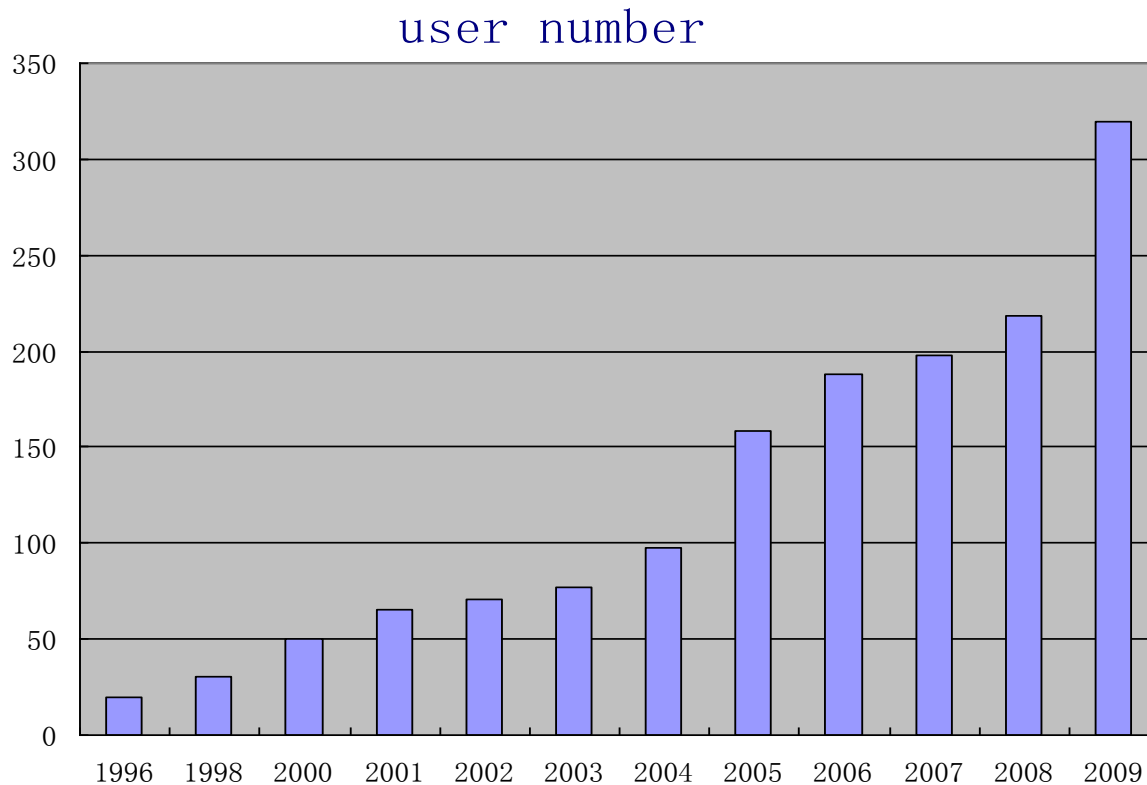
## □ During 2006-now (11th 5 year plan)

- In 2008, Lenovo DeepComp7000
  - ✓ 150Tflops, 13,000 cores
  - ✓ TOP500 : No.19 ; China TOP100 : No.2
  - ✓ 3 kinds of nodes, Altix 4700, IBM 3950, IBM Blades
- In 2009, 300TFLOPS GPU

# SCCAS-Services: Users

---

- Users in SCCAS has massively increased since 1996 (from 20 to more than 320)





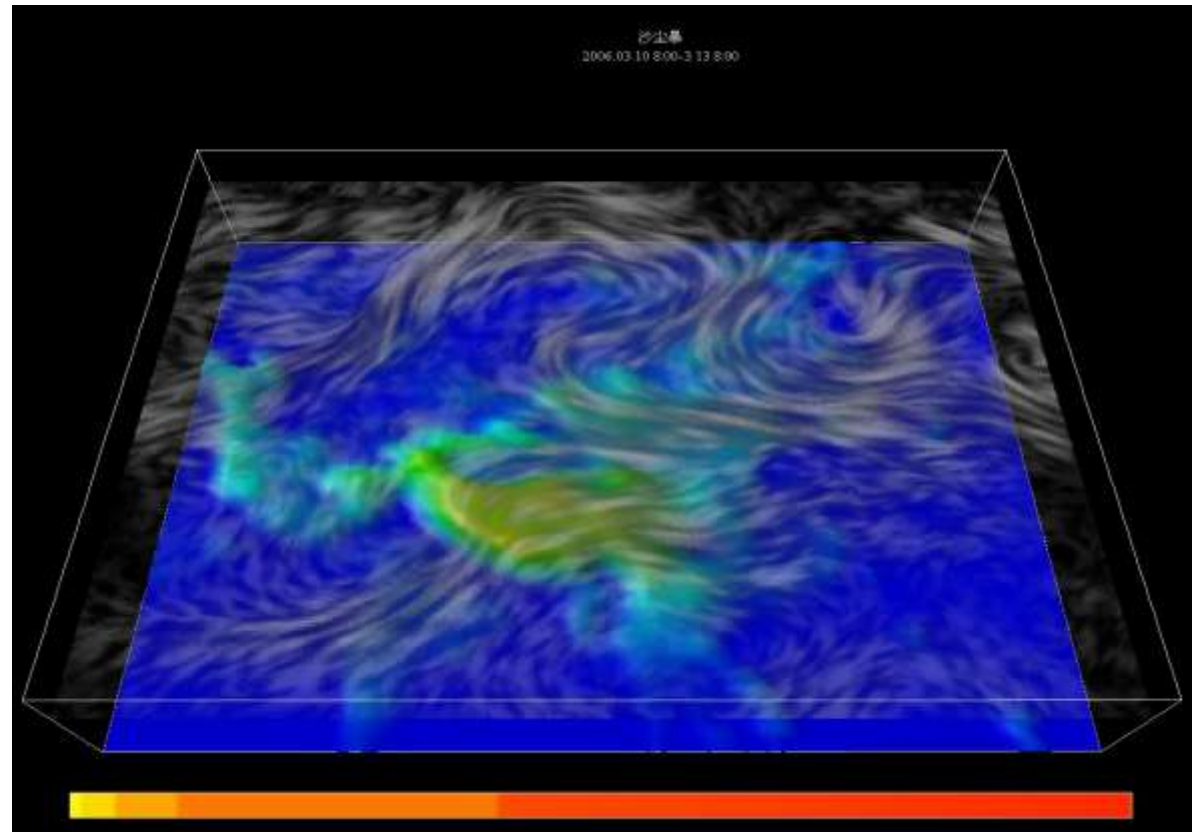
# SCCAS-R&D of Algorithms and software

---

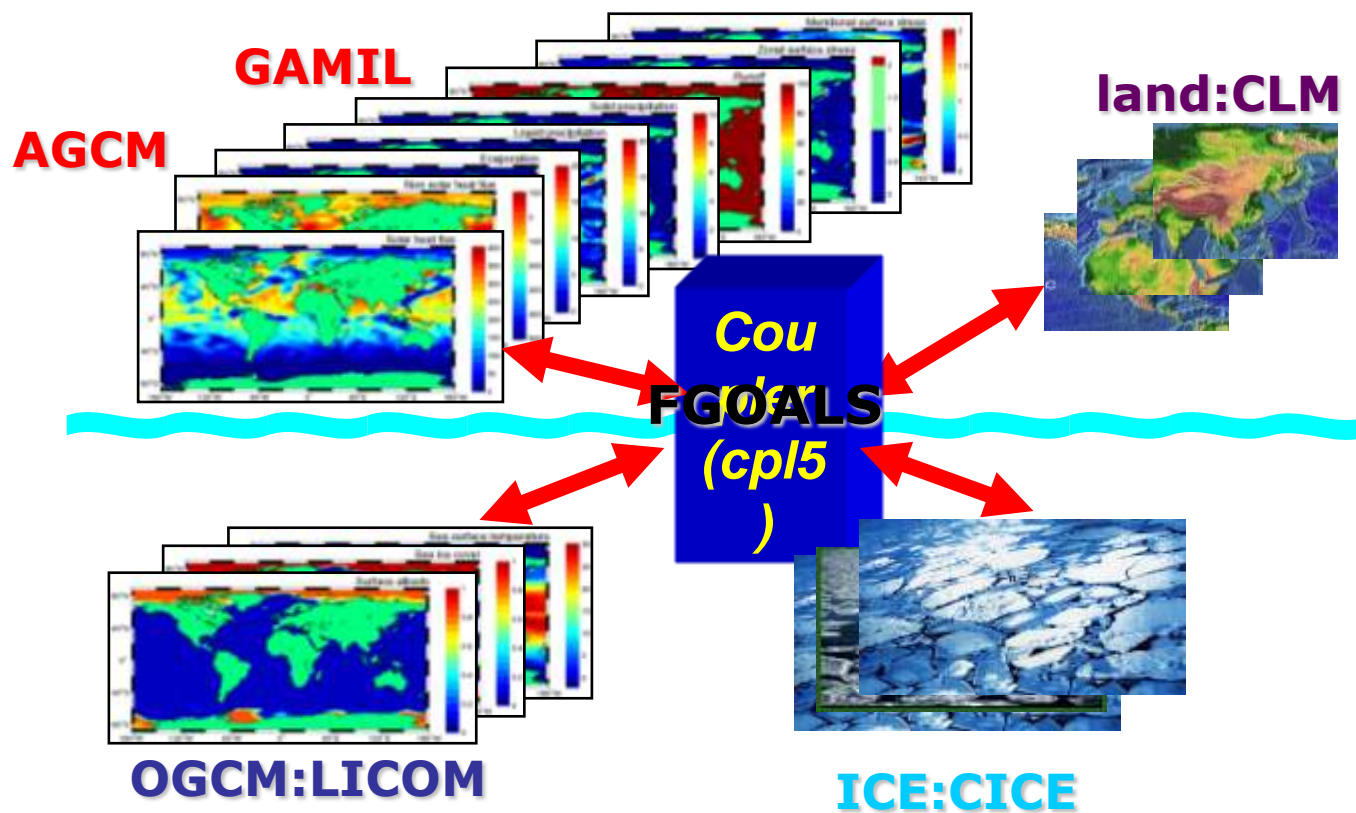
- ❑ **Parallel AMR (Adaptive Mesh Refinement) method**
- ❑ **Parallel Eigenvalue Problem**
- ❑ **Parallel Fast Multipole Method**
- ❑ **Parallel Computing Model**
- ❑ **Gridmol**
- ❑ **ScGrid middleware**
- ❑ **PSEPS**
- ❑ **FMM-radar**
- ❑ **Transplant many open source software**

# Prediction of Sandstorm

- ❑ Real-time prediction system of Sandstorms in China Meteorological Bureau
- ❑ DeepComp6800
  - 256 CPUs
  - from 15hours down to 8mins



# Global Climate Model-FGOALS



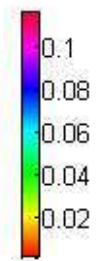
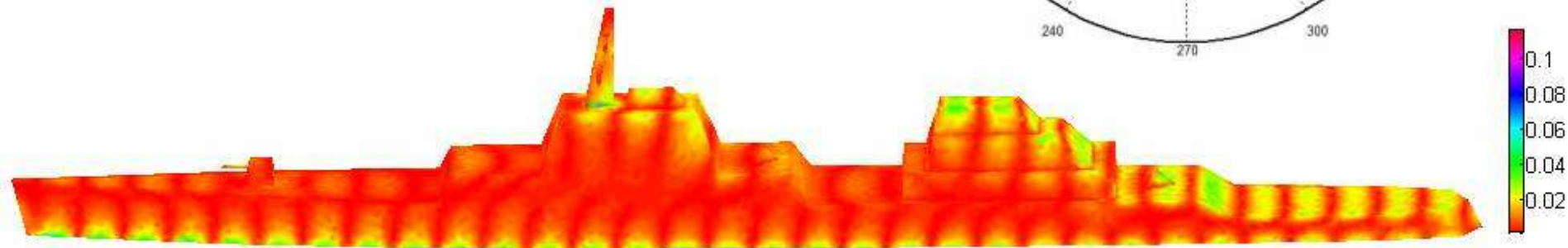
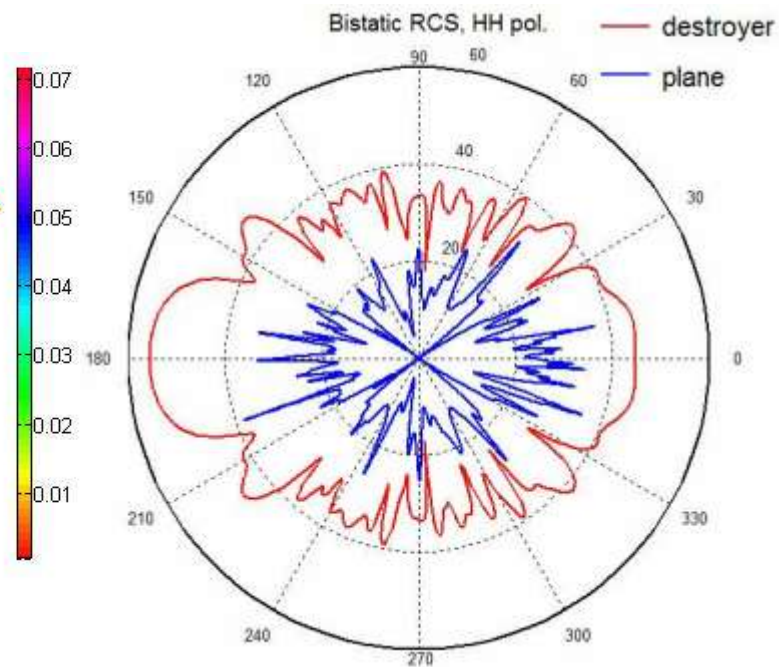
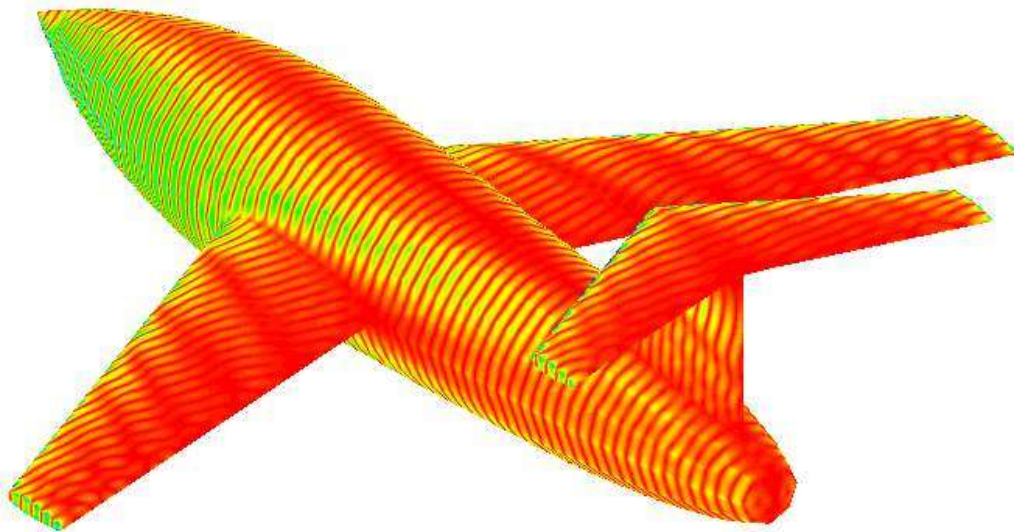
CPU	AGCM	LAND	OGCM	ICE	Coupler
1620	720	90	540	180	90



# EM Scattering

## □ The surface currents and RCS of plane and destroyer models

- Used for antenna design, RCS analysis
- Stealth design of airplanes, etc.



# P\_InsPecT/cuda-InsPecT Software

## Software Introduction

- Both are optimized InsPecT software
- P\_InsPecT is open source and can be downloaded from SCBG
- Cuda-InsPecT will be open source

## Software Function

- InsPecT is an unrestricted identification software of PTMs(post-translational modifications)

## Software Characteristics

- P\_InsPecT
- via MPI
  - run on CPU cluster or CPU nodes of HPC

- cuda-InsPecT
- via MPI+cuda C
  - run on GPU cluster

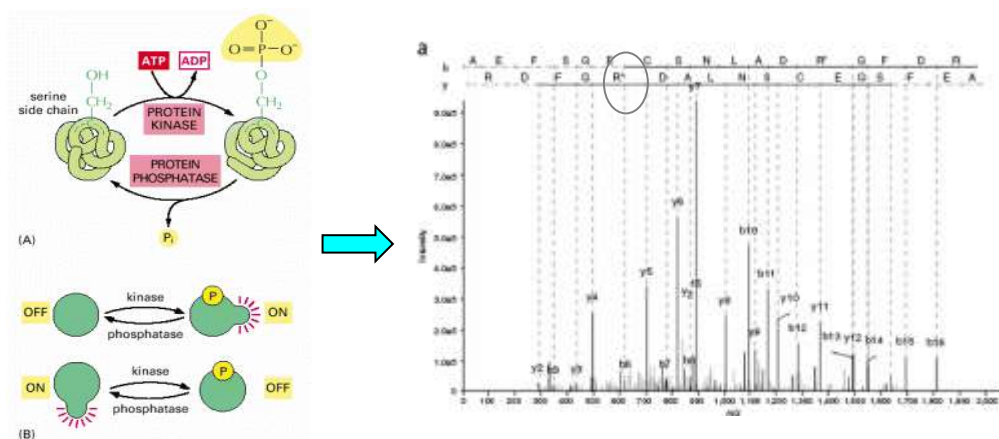
## Software Performance

Software: P\_InsPecT (one modification)  
 Database: 36547 mass spectrometric;  
 107962 protein sequences  
 Environment: DeepComp7000

	One Core (estimate)	2048 Cores
Time	1177.7 h	0.4 h

Software: cuda-InsPecT (two modifications)  
 Database: 62346 mass spectrometric;  
 107962 protein sequences  
 Environment: Dawn 6000A

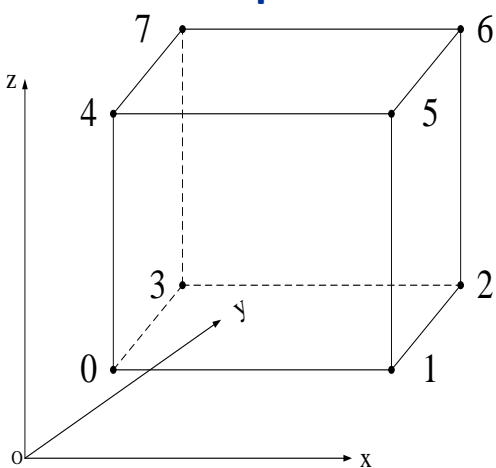
	One Core (estimate)	677 Fermi C2050
Time	6 years	2.034 h



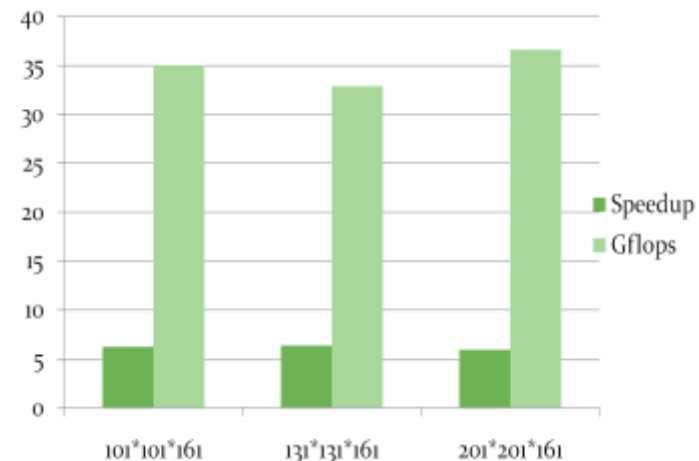
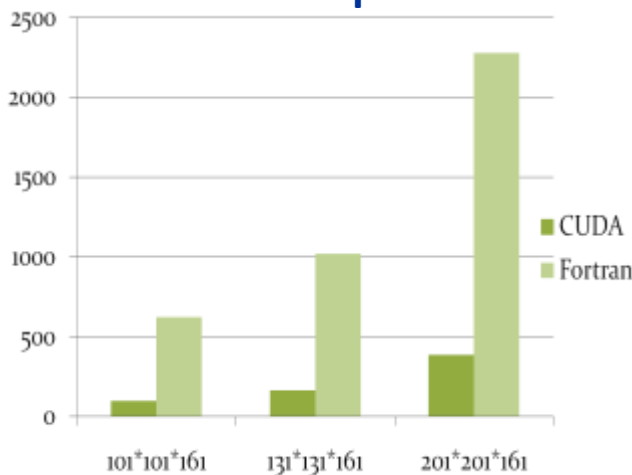
# Seismic Wave



- ❑ SORD is an open-source software developed by Geoffrey Ely from USC, to simulate 3D wave propagation and spontaneous rupture on hexahedral mesh.
- ❑ Supports many types of BC including PML.
- ❑ The test we implemented involved two layers, one (7km) over another (1km) with rectangular mesh. A double-coupled point is set at (0, 0, 2km) and last for 0.1 second to provide seismic waves. Seismic wave propagates in the whole space during the test. Two types of waves will be observed, P wave and S wave.
- ❑ Based on Emmett and David Wang's work which outlined a framework. Tim Dong do some optimization and MPI+CUDA implementation.

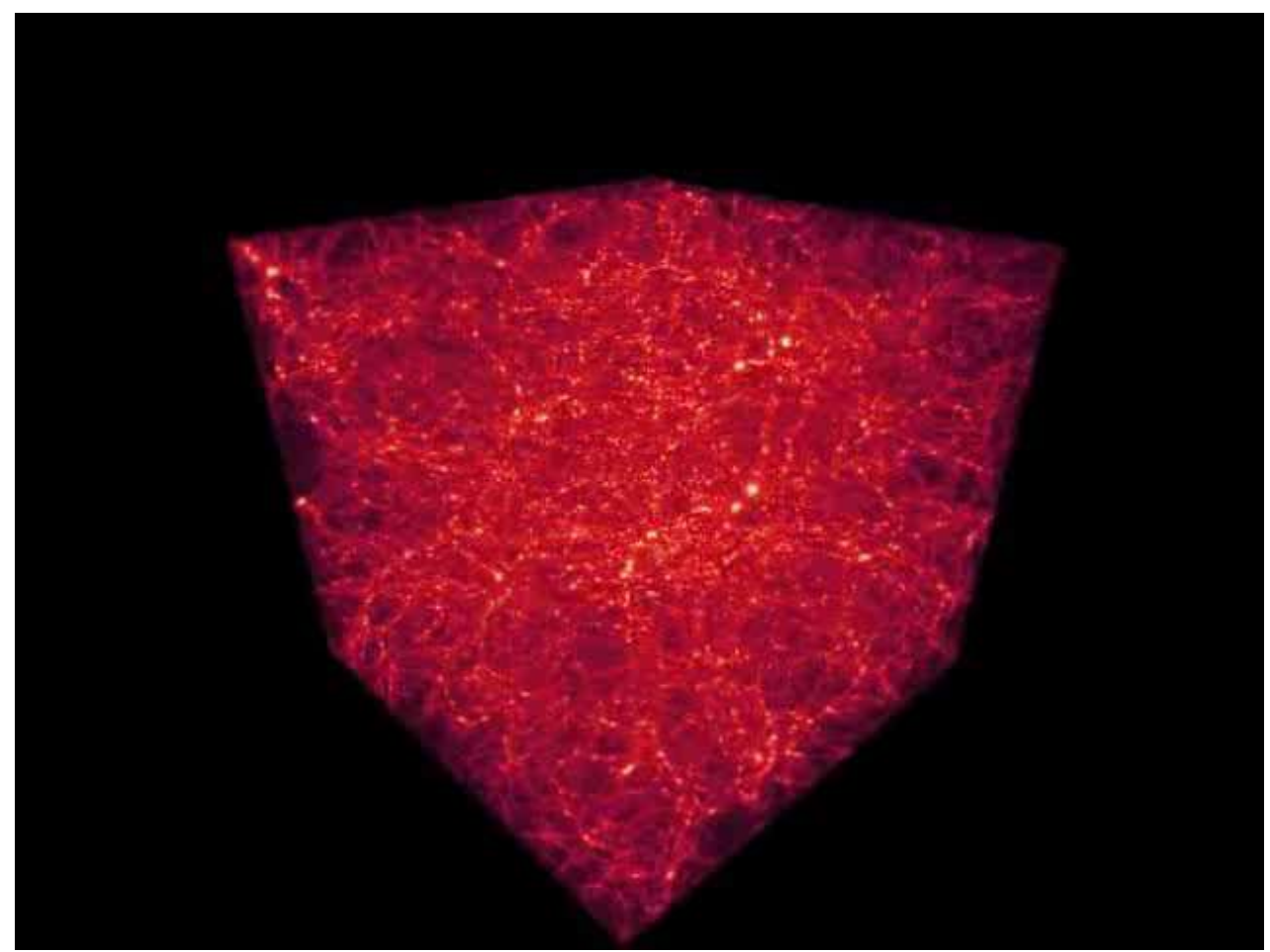


cubic stencil



1000 steps on NIVIDIA Tesla 1060C and Xeon X5570 2.9GHz

# C4: Computational Cosmology Consortium of China





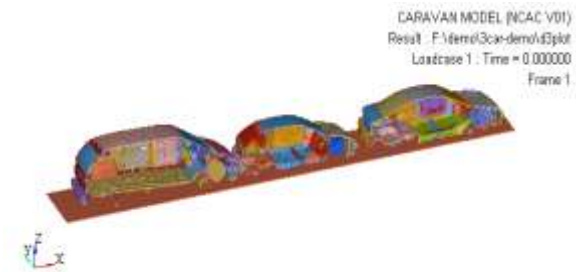
## CAD/CAE Platform in Automation Industry

### □ Who we provide our service to

- CAD/CAE engineers

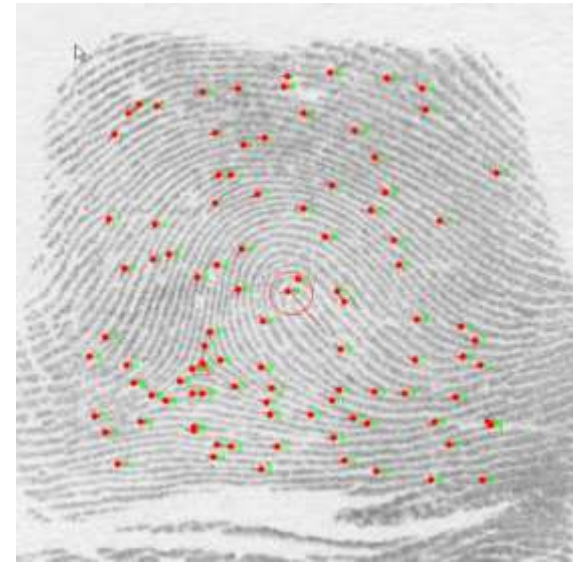
### □ What we deliver to them

- A high-performance service
  - ✓ DeepComp 7000 for parallel solver
  - ✓ High-performance graphic workstations for pre/post processing
- A high-usability cloud computing solution via internet
  - ✓ Remote graphical access to workstations
  - ✓ Grid-computing based resources
  - ✓ Web portal integrated with CAE software
- A high-safety environment
  - ✓ Secure access through VPN and firewall
  - ✓ Isolate unauthorized access from cluster by web portal



# Fingerprint Recognition

---



- **Fingerprint database : 19,600,000 fingerprints**
- **HPC: 1024 cores ( 16 nodes, 64 cores/node )**
- **Speed: 4000 fingerprints/second/core**
- **Run time: 3 months (linear Speedup)**



## SCE Introduction

---

- SCE is an acronym for
  - Scientific, Super, Scalable
  - Computing
  - Environment, Easy, Yi



- Yì: have 10+ meanings in chinese
- change (改变)
  - easy (容易)
  - the Book of Changes (周易)
  - ...

## SCE Introduction

---

- **Change we need**
  - Simple
    - ✓ many HPCs as one
  - Easy
    - ✓ bring input, take result
  - Stable
    - ✓ jobs can run anytime



# GridMol

**A Molecular Visualizer and Builder by SCCAS**



## Overview

---

- ❑ **Developed by Virtual Laboratory for Computational Chemistry, SCCAS**
- ❑ **A typical application of SaaS (Software as a Service) for molecular modeling and visualization**
- ❑ **An extensible tool for building a high performance computational chemistry platform in Grid environment**
- ❑ **Providing users one-stop calculation service**



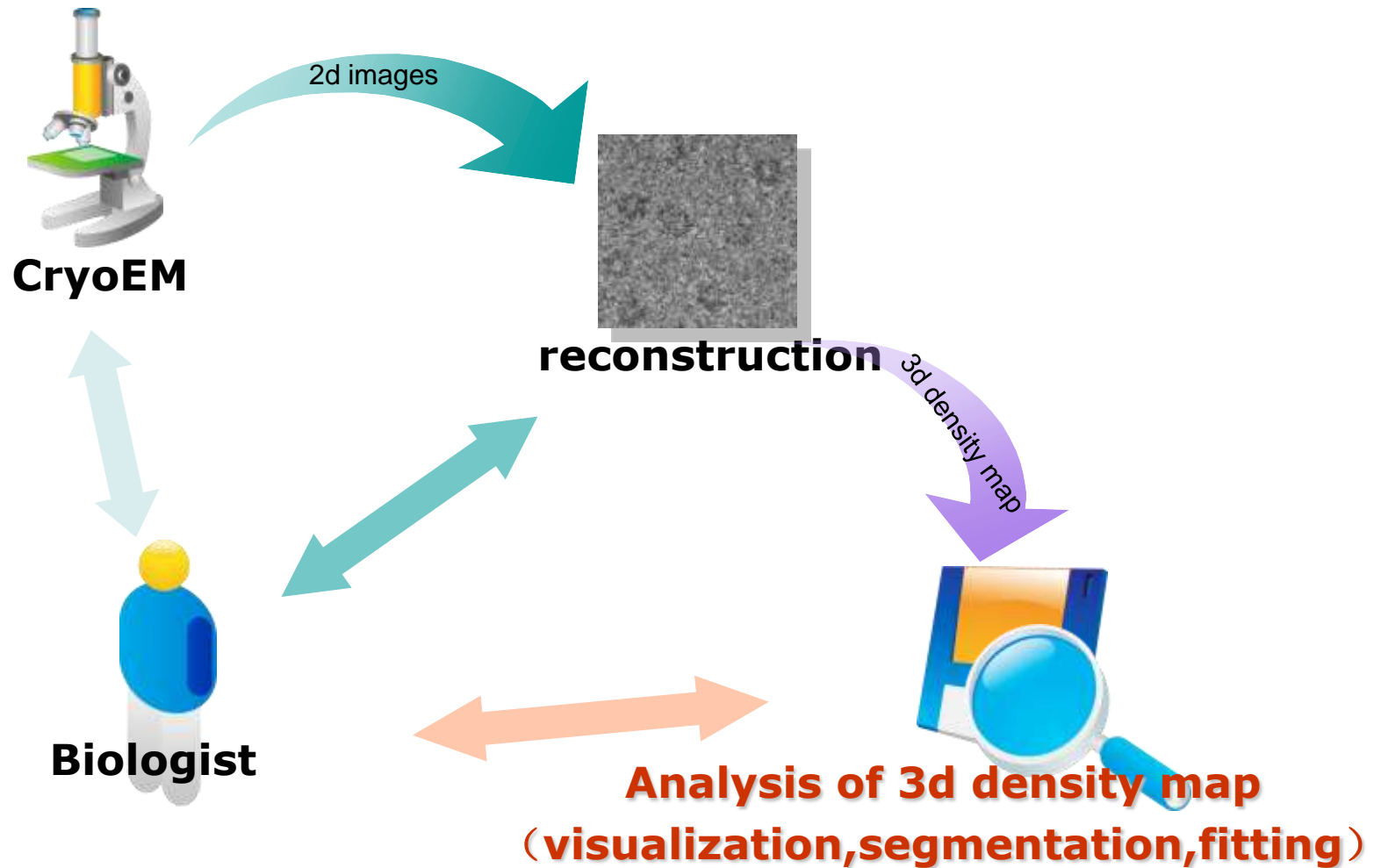
## Functions

---

- ❑ **Molecular Modeling**
- ❑ **Molecular Visualization**
- ❑ **Calculation results analysis**
- ❑ **Job submission and monitoring**

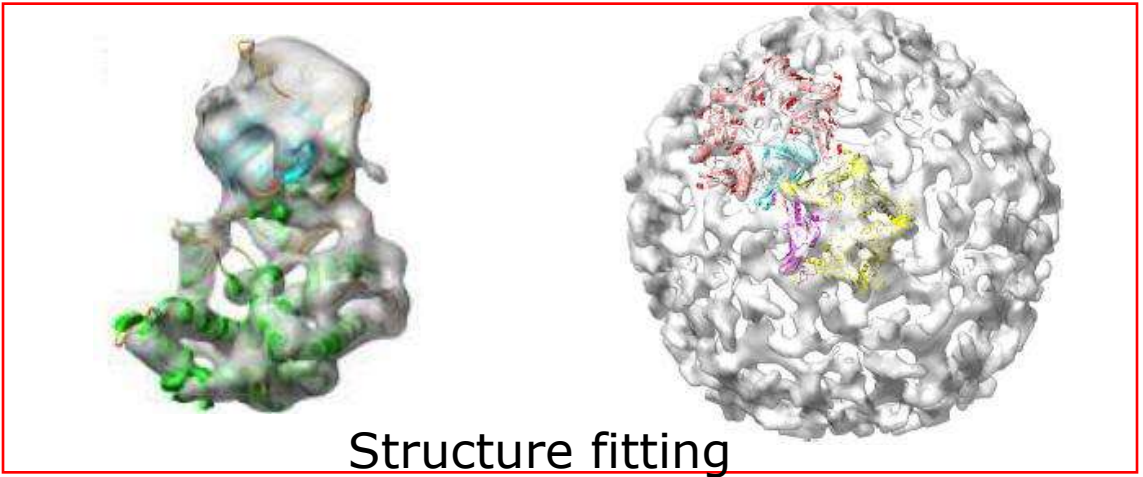
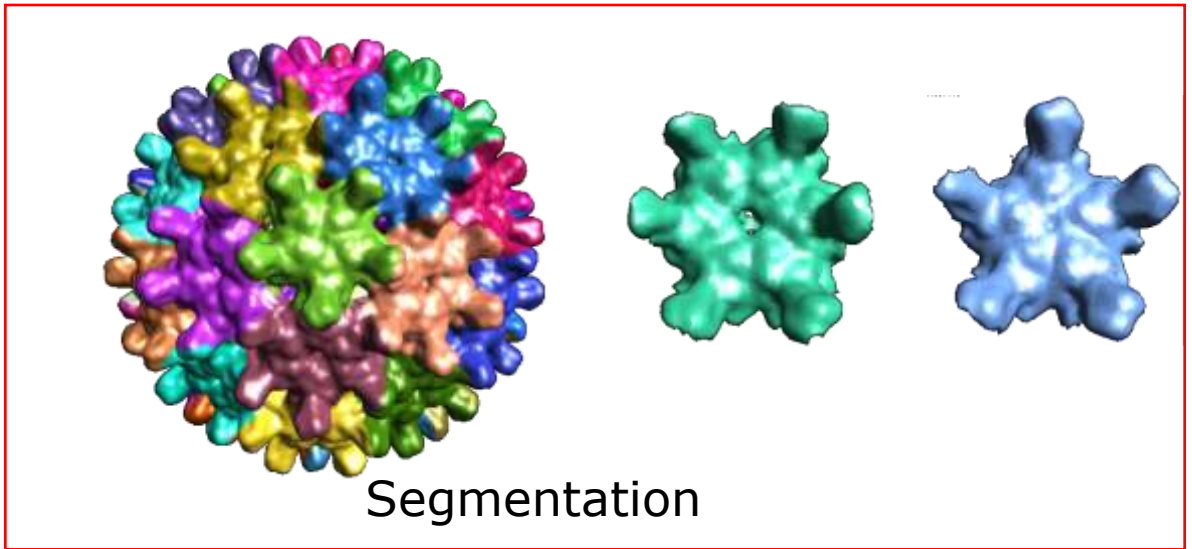
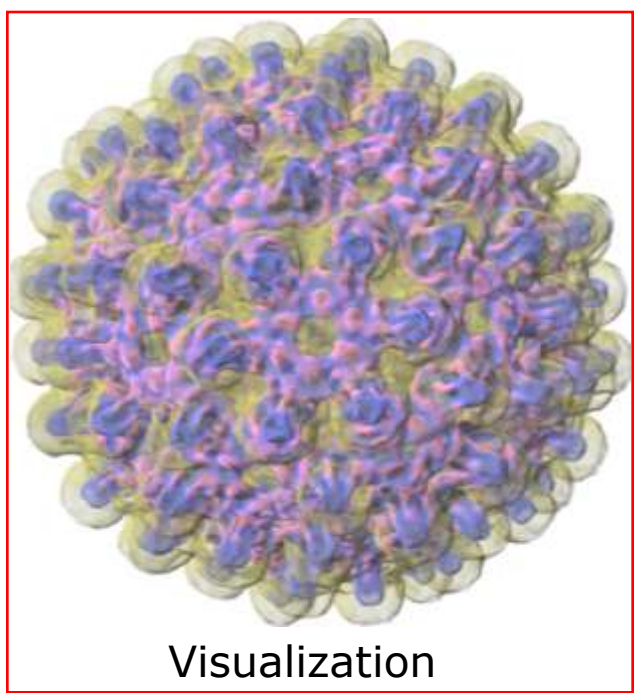
## Workflow of CryoEM technique

---



# Motivation

---



## Features

---

- ❑ **Volume rendering for density maps**
- ❑ **Automatic Segmentation**
- ❑ **Automatic fitting for multi-crystal structures**
- ❑ **Easy movie made based on key frame**

# Applications

中国科学院生物物理研究所  
Institute of Biophysics, Chinese Academy of Sciences

## Structure research of Group II Introns

Yanxin Mao, Shengjun Hu, Yong Zhai, Yongli He, Zhong  
1. National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese  
2. Institute of Microbiology, Chinese Academy of Sciences

**Introduction**  
Group II (G2) is a unique class of the protein-independent catalytic introns. It is the first class of self-splicing introns and represents a unique Group II found in a wide range of organisms. Group II introns are widely distributed in various organisms, including bacteria, archaea, fungi, plants, and animals. They have been shown to catalyze their own excision from the DNA of their hosts. The catalytic mechanism of G2 introns is still unclear. The catalytic mechanism of G2 introns is still unclear. The catalytic mechanism of G2 introns is still unclear.

**Crystalization, diffraction and symmetry analysis**

Crystals of G2 and G2P were grown by microbatch and microbatch methods. The crystals were collected using Quantum 210 D DSI detector and analyzed on a Siemens diffractometer. The data were collected with a Siemens DSI 210 DSI detector. The data were collected with a Siemens DSI 210 DSI detector.

Parameter	Value
Resolution	2.30 (4.00) Å (3.44 Å)
R-factor	0.1920 (0.1727)
R-free	0.2400 (0.2100)
Space group	P21
Completeness (%)	99.99 (99.99)
Redundancy	13.50 (3.00)
Resolution	11.8 (3.1)
Map quality	4.3 (1.7)

**Negative stained EM, image processing and 3D reconstruction**

EM was used to study the structure of G2 and G2P. The images were processed and analyzed using EM software. The 3D structure was reconstructed using cryo-EM techniques.

The Laboratory of Biological Electron Microscopy  
Tel/Fax: +86 (10) 64008419 or 64008420  
Email: emlab@ibp.ac.cn or feibio@ibp.ac.cn

2010.5

# 生物物理学

## ACTA BIOPHYSICA SINICA

中国生物物理学会 中国科学院生物物理研究所

$$F(x, y) = \iint f(x, y) e^{-i2\pi(xu + yv)} dx dy$$

$$f(x, y) = \iint F(u, v) e^{i2\pi(xu + yv)} du dv$$

$$F(f+g) = F(f) + F(g)$$

## International Workshop of 3D Molecular Imaging by Cryo-Electron Microscopy

### 冷冻电镜三维分子成像国际研讨会

ISSN 1674-800X

Volume 1 Issue 1

# Protein & Cell

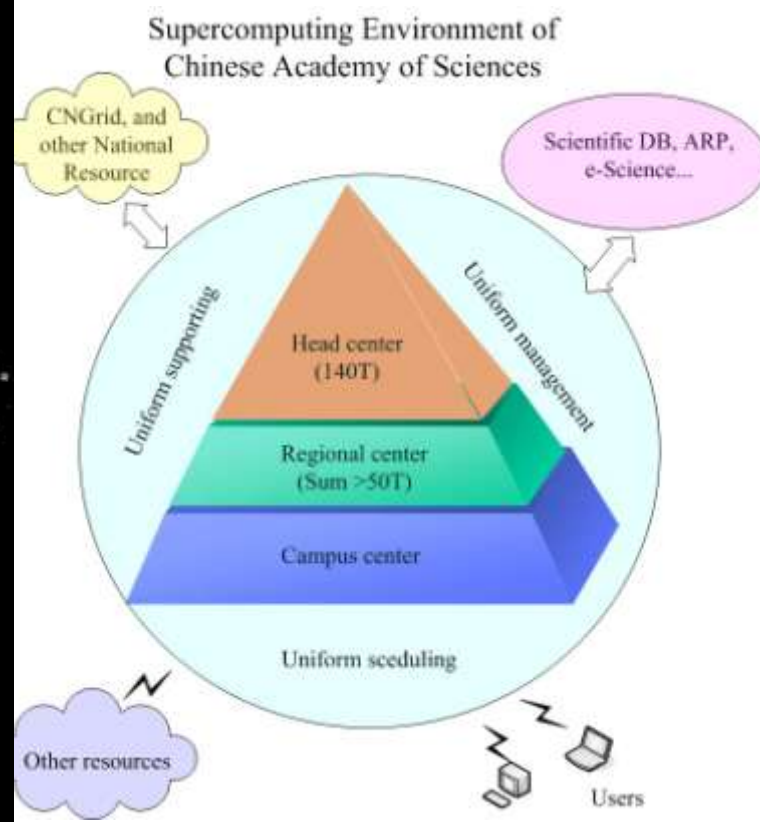
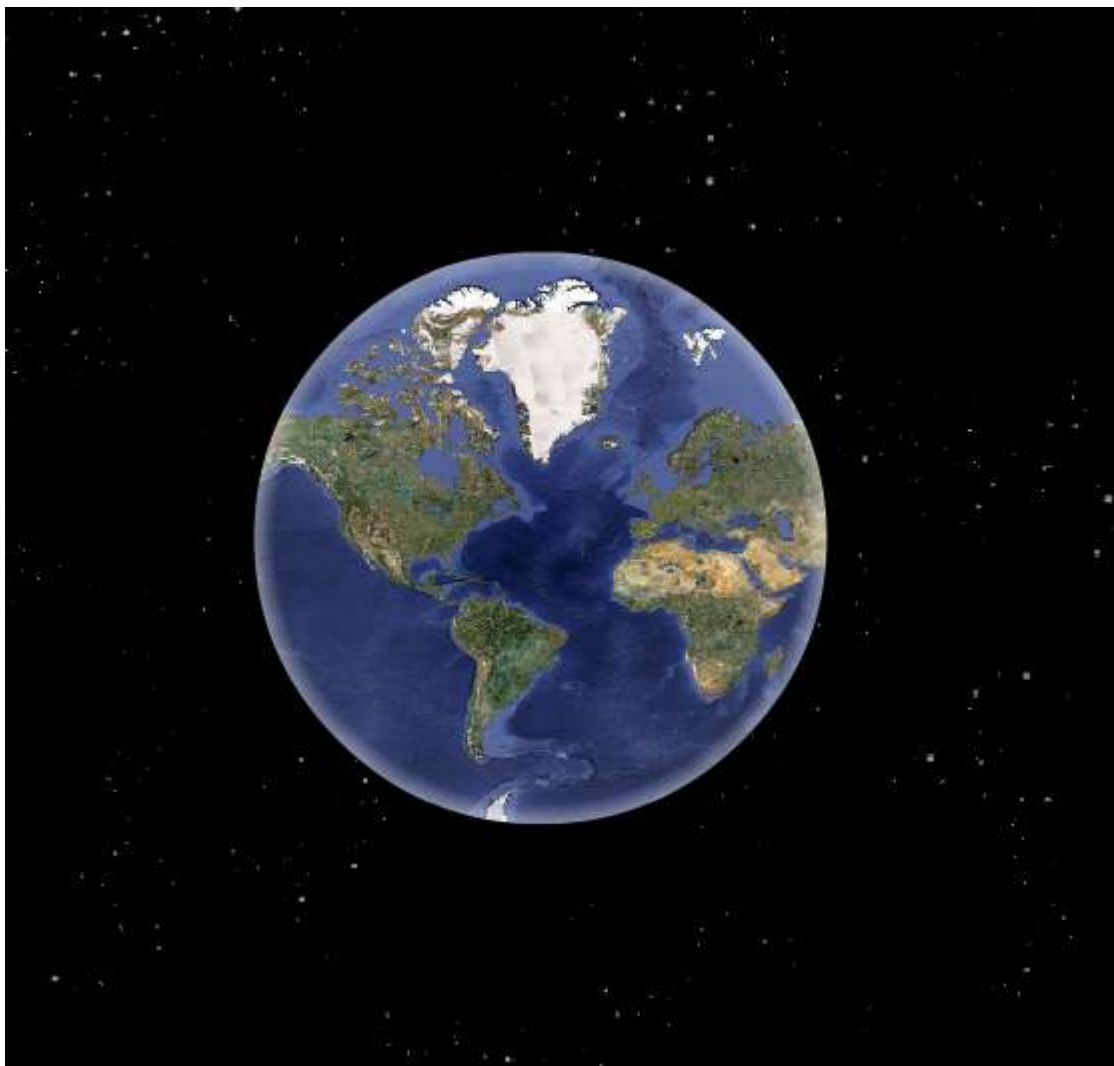
January 2010

RHOV Structures Reconstructed by Cryo-Electron Microscopy

高等教育出版社 Higher Education Press

Springer

# China Scientific Computing Grid (China ScGrid)







## GPU Clusters within CAS

Site	Vendor	R <sub>peak</sub> /Tflops
Institute. of Electrical Engineering	Lenovo	112
Shenzhen Institutes of Advanced Technology	Lenovo	200
USTC	Lenovo	183
CNIC	Lenovo, Dawning	300
National Astronomical Observatories	Lenovo	158
Institute of Geology and Geophysics	Lenovo, Dawning	200
Institute of Modern Physics	Lenovo	202.5
Institute of High Energy Physics	Dawning	195.5
Institute of Metals Research	Dawning	183
Purple Mountain Observatory	Dawning	180
SUM		1.914 Pflops

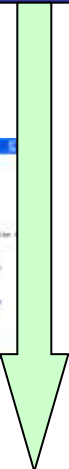


Users

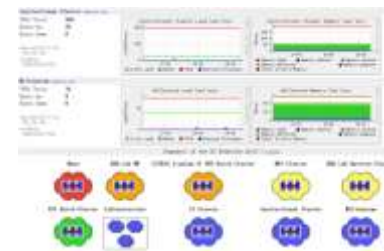
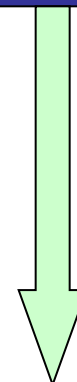


Administrator

Windows / Linux Clients



Web Portal



SCE Middleware



HPC, Cluster, Workstation, Storage

## Acknowledgements

---

- **My colleagues**

- Long Wang, Haili Xiao, Zhong Jin, Guihua Shan, Zhonghua Lu, etc.

- **My Students**

- ...

---

# Question & Suggestion

# Thank you very much!