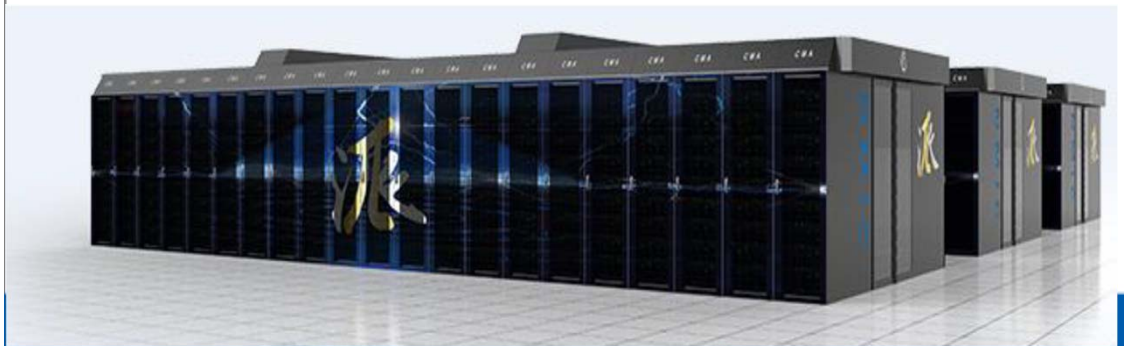
The 18th Workshop on high performance computing in meteorology

CMA HPC Update

Supporting meteorological service

Min Wei, Chunyan Zhao and many colleagues
National Meteorological Information Centre
China Meteorological Administration



国家气象信息中心
National Meteorological Information Center

Contents

- HPC Systems
- Model-Supportive Software Systems
- Conclusions

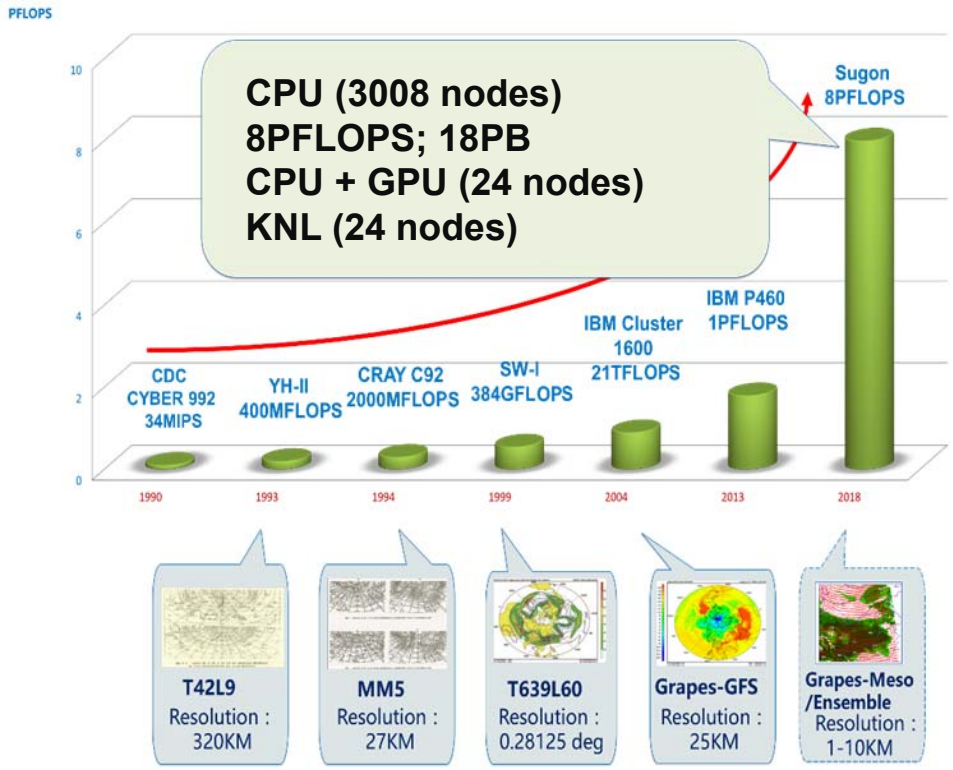
国家气象信息中心
National Meteorological Information Center

Contents

- HPC Systems
- Model-Supportive Software Systems
- Conclusions

国家气象信息中心
National Meteorological Information Center

History



国家气象信息中心
National Meteorological Information Center

IBM HPCS

Beijing : 1 PFLOPS

Shenyang : 75 **TFLOPS**

Shanghai : 50 **TFLOPS**

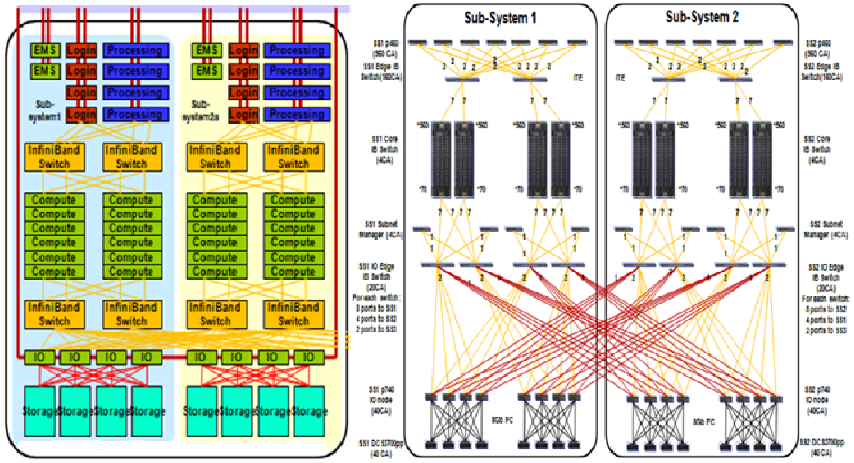
Wuhan : 75 **TFLOPS**

Guangzhou: 400 **TFLOPS**

Chengdu : 25 **TFLOPS**

Lanzhou : 25 **TFLOPS**

Urumqi : 25 **TFLOPS**



System	Installation Time	Peak Performance (TFLOPS)	Storage Capacity (TB)
IBM Flex System P460	2013	Production Subsystem: 527	2109.38
	2014	Research Subsystem: 527	2109.38

Resource utilization

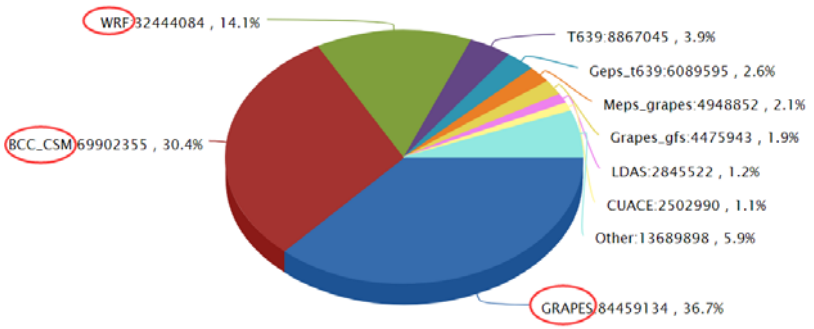
IBM HPCS accounts

- 578

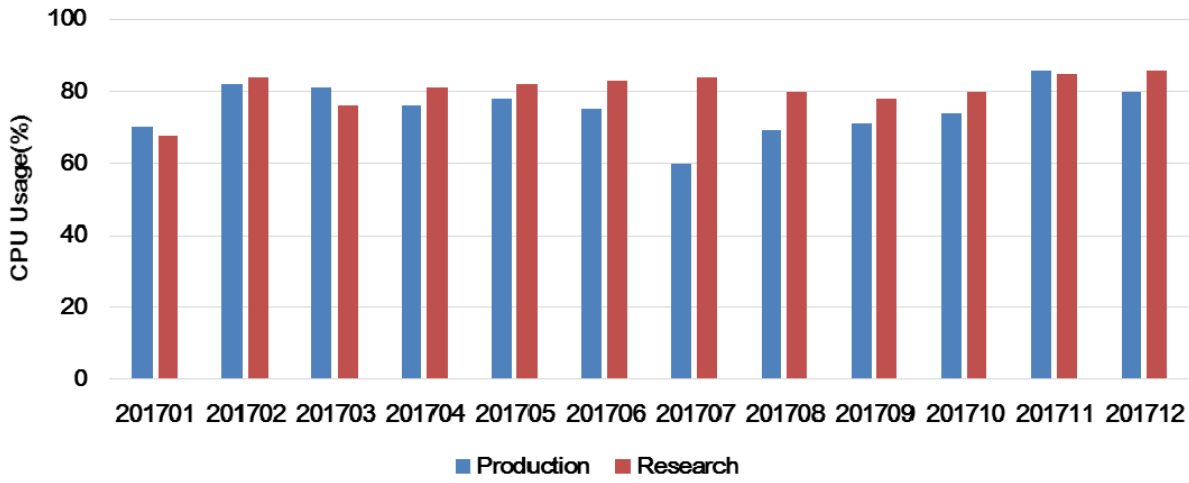
IBM HPCS utilization

- Maintain high both in system availability and CPU utilization, 70% to 80% on average peaking at 95%.

Computing resources statistics



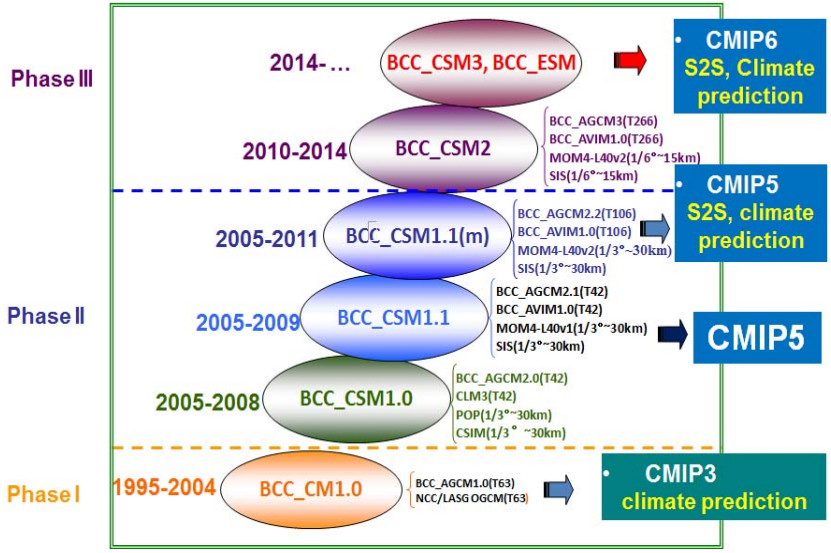
Average CPU utilization rate



GRAPES & BCC_CSM

- GRAPES = Global/Regional Assimilation PrEdiction System
- BCC_CSM = Beijing Climate Center Climate System Model

	GRAPES-GFS	GRAPES-MESO	GRAPES-TYM	GRAPES-MEPS
Forecast range	10d	3d	5d	3d
Domain	Global	East Asia	West Pacific	East Asia
H-resolution	0.25°	0.1°	0.12°	0.15°
V-resolution	60L 3hPa	50L 10hPa	50L 10hPa	50L 10hPa
Forecast time	00, 12 UTC 240 h	00, 12 UTC	00, 12 UTC	00, 12 UTC 15 members



Benchmark

- GRAPES-GLOBAL model (Parallel)
- GRAPES-MESO model (Parallel)
- GRAPES-4DVAR four-dimensional variational model (Parallel)
- BCC_CSM climate system model (Parallel)
- BCC_AGCM atmosphere model (Parallel)
- GRAPES-SVD singular vector analysis of regional ensemble forecast system (Serial)
- WRF model (Parallel)

- IOzone Benchmark
- IMB Benchmark
- Job scheduling
- Public domain meteorological software packages

PI-Sugon

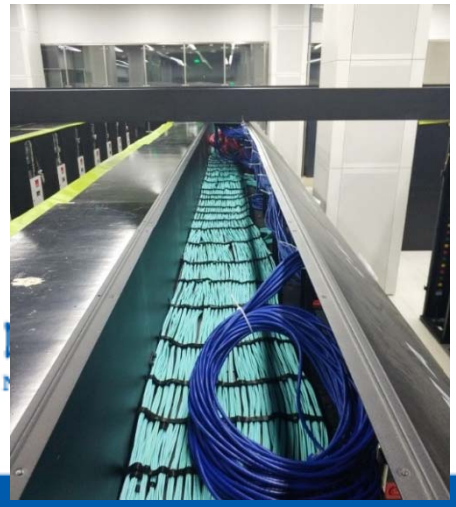
2017.9 2018.1 2018.6

1 Contract 2 Arrival 3 Installation 4 Power up 5 Service 6 Pre-Operation





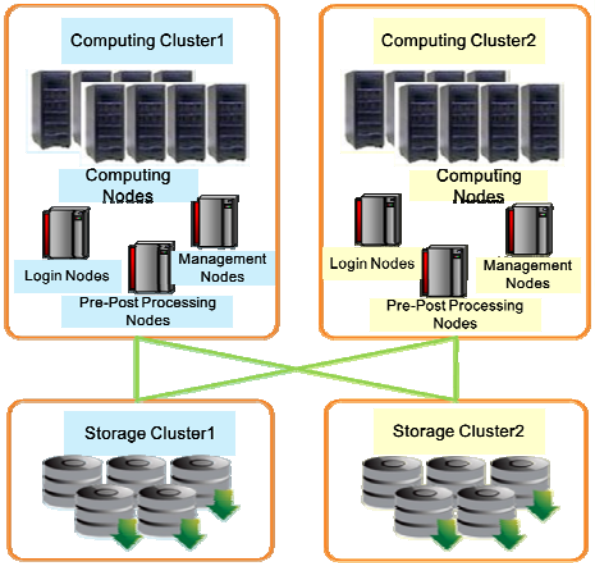


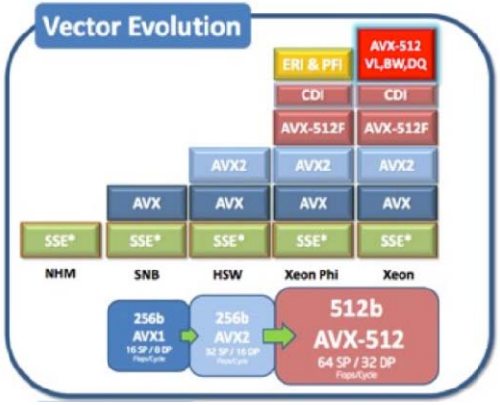


Architecture

Two subsystems: hot standby

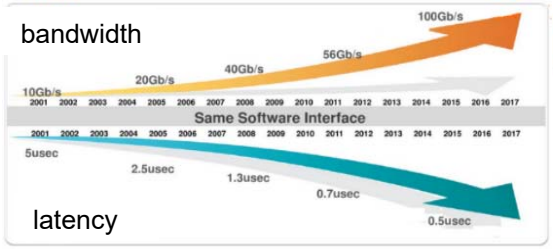
- Independent computing, shared storage
- General processor, for each system
  - Computing nodes: ~1500
  - Total CPU cores: ~50000
  - Intel Xeon Gold 6142 (16 Core, 2.6GHz)
- 8 PFLOPS peak performance
- 18 PB storage capacity
- 100Gb/s InfiniBand EDR network
- Parastor 300 parallel file system

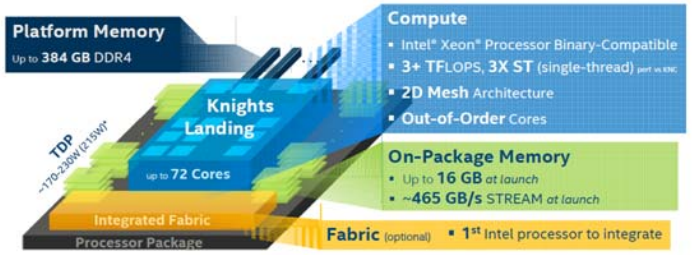




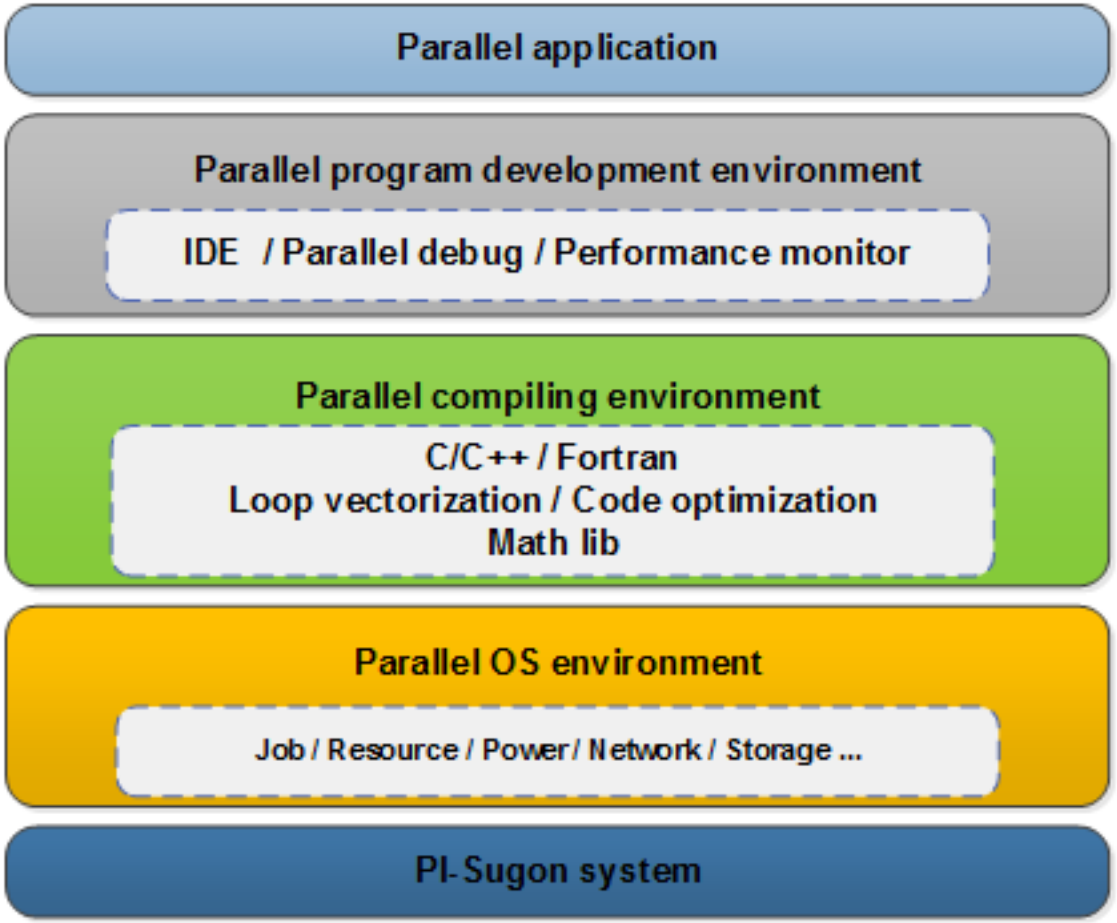
New technology test and development subsystem

- CPU + GPU (24 nodes)
- Intel KNL (24 nodes)





Software Stack



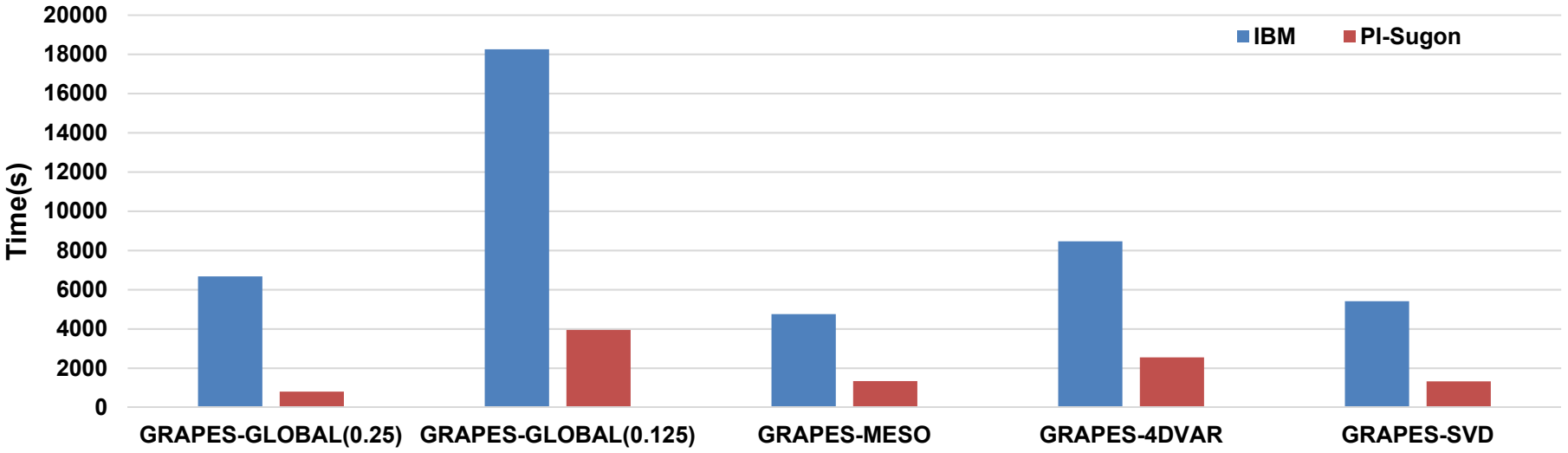
国家气象信息中心
National Meteorological Information Center

IBM & PI-Sugon

	IBM	PI-Sugon
Improvement		
Peak Performance	~1PFLOPS	~8PFLOPS
Storage Capacity	~4PB	~18PB
Inter-Connection	QDR 40Gb/s	EDR 100Gb/s
Difference		
OS	AIX 7.1.0.0	RedHat Enterprise 7.4

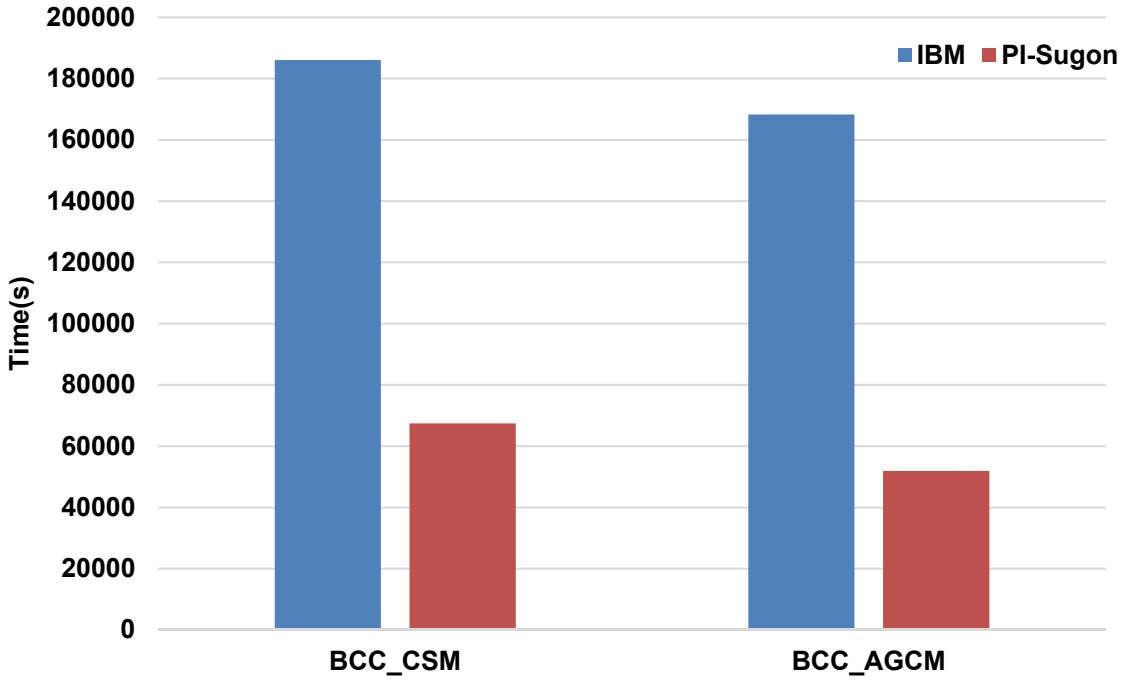
adios	ferret	grads	hdf5	ioapi	ncl_ncarg	nlopt	plapack	udunits
blas	fftw	grib_api	hdfeos	jasper	nco	openblas	plasma	wgrib
boost	geos	gsl	hdfeos5	lapack	ncview	parallel-netcdf	proj	wgrib2
esfm	GotoBlas2	hdf	hypre	libpng16	netcdf	petsc	scalapack	

GRAPES model suite



国家气象信息中心
National Meteorological Information Center

Climate models



国家气象信息中心
National Meteorological Information Center

Contents

- HPC Systems
- Model-Supportive Software Systems
- Conclusions

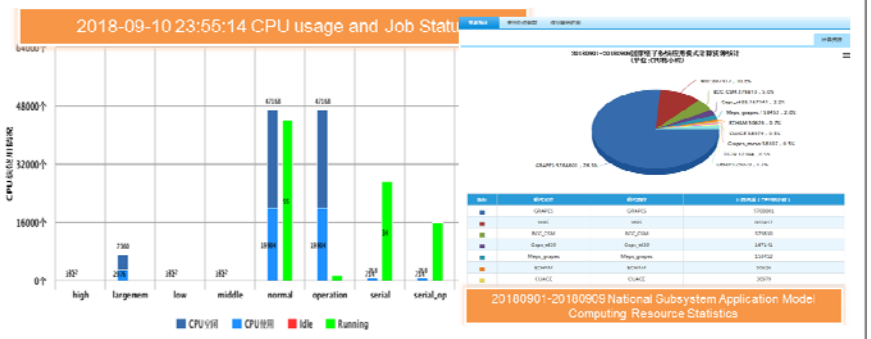
国家气象信息中心
National Meteorological Information Center

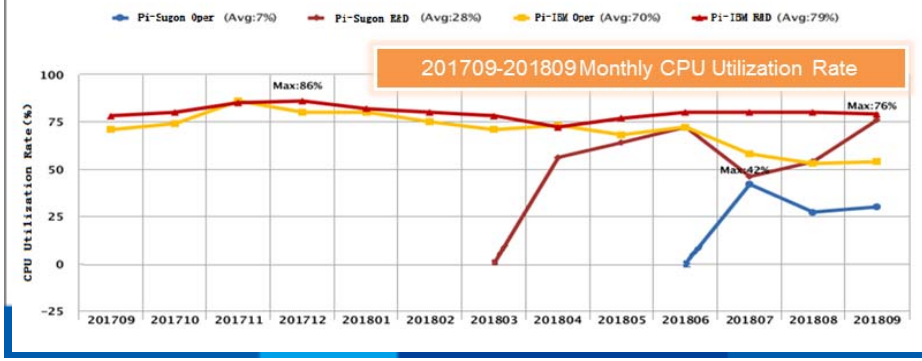
- **High performance computer management software**
  - Refined resource management system
  - Operational monitoring system
- **Numerical model supporting software**
  - Code management system
  - GRAPES Integrated Setting Experiment Tool(GISET)
  - GRAPES Interactive Data Analytics Tool (GIDAT)

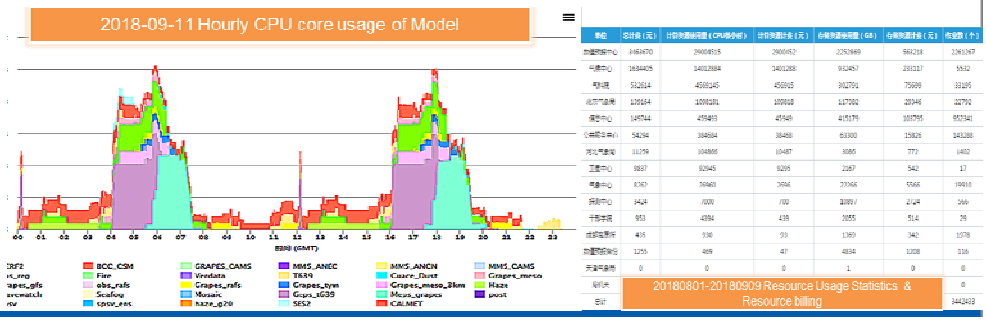
国家气象信息中心
National Meteorological Information Center

Refined resource management system

- Resource management of IBM & PI-Sugon systems
- Unified management of national and regional resources
- Real-time and historical statistical analysis of system resource usage and utilization
- Computing resource and storage resource usage accounting
- Model & job statistical analysis
- Planning: intelligent resource management
  - Resource data mining
  - Decision support analysis

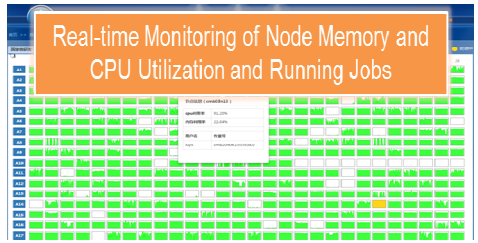


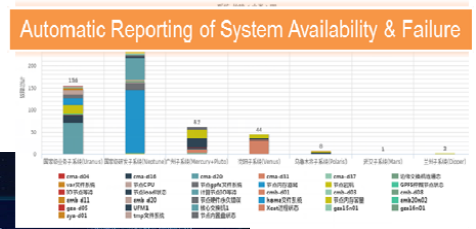


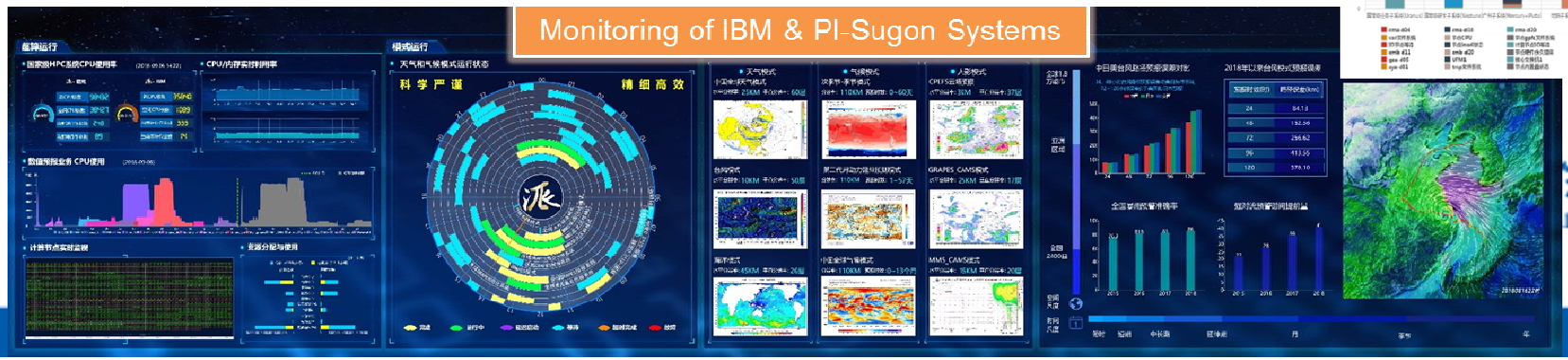


Operation monitoring system

- Monitoring of IBM & PI-Sugon systems and software，audio alarm
- Unified management of national and regional resources
- Real-time monitoring and historical statistical analysis of failure
- Automatic reporting of system availability & statistical analysis of failure
- Fault handling workflow & fault knowledge database
- Model job monitoring
- Real-time monitoring of memory, CPU utilization and jobs
- Planning: intelligent job management
  - Model application feature analysis and data mining
  - Decision support analysis



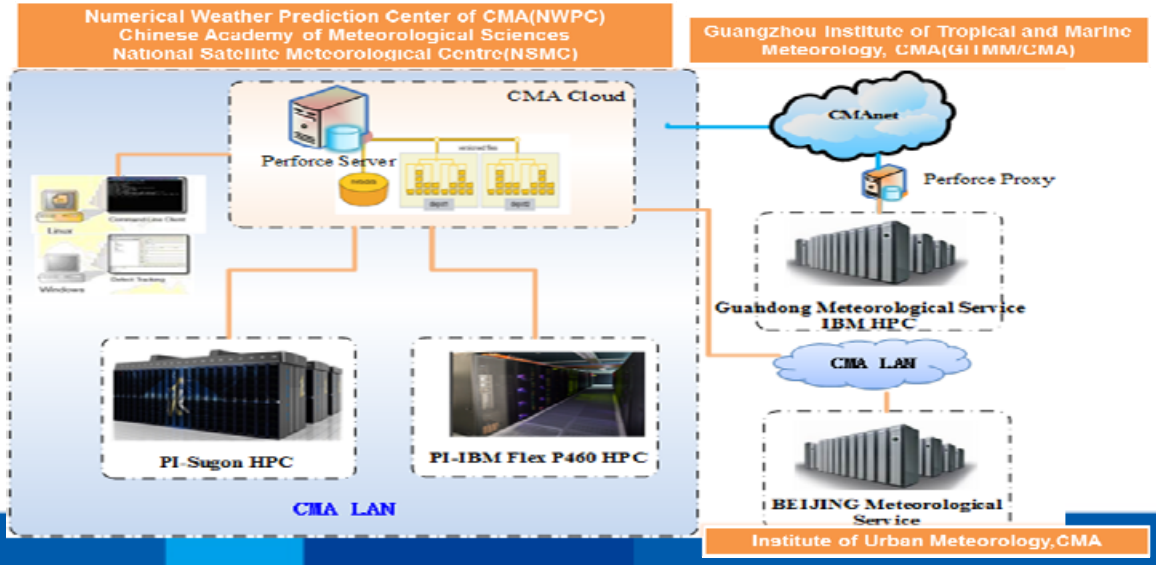




信息中心
l Information Center

Code management system

- Code management system (since 2010)
  - Perforce application on IBM and PI-Sugon HPC
  - GRAPES-GFS,GRAPES-MESO,BCC_CSM code repository
  - National & regional distributed design for GRAPES-MESO collaboration
  - Code version control and integration control
  - Code updates 17,000+, code integration 1,000+, version release and bug fix 500+
- Planning: git-based code management

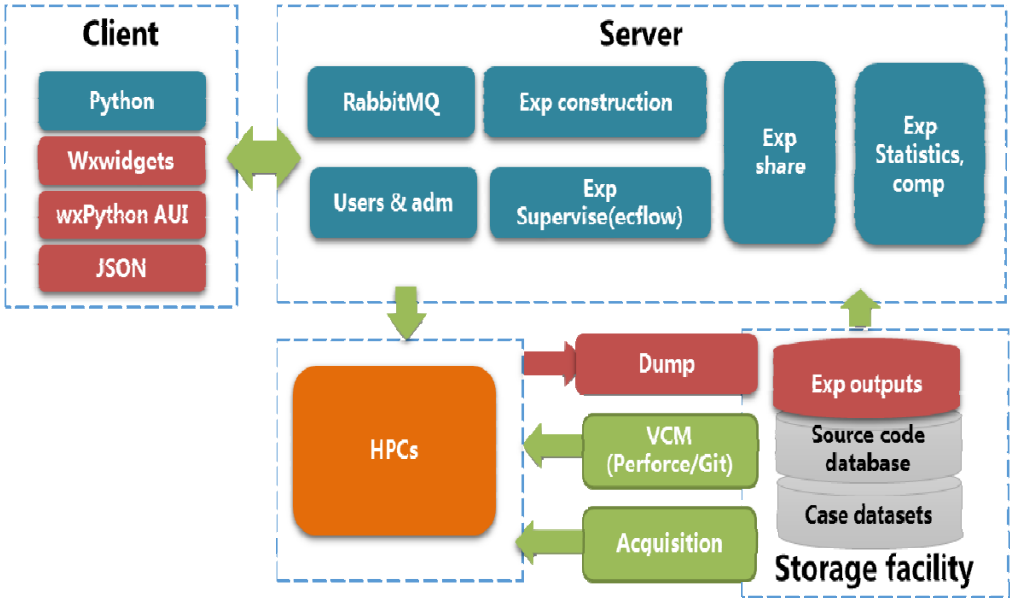


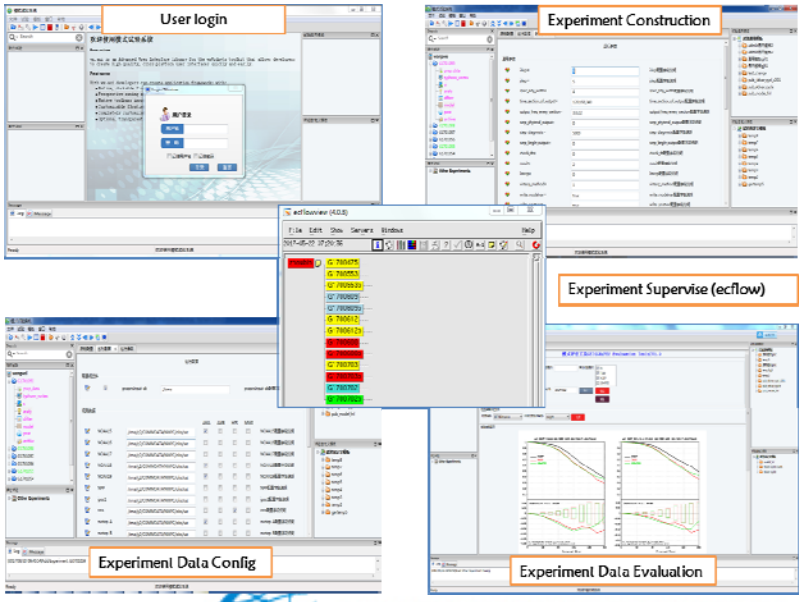
国家气象信息中心
National Meteorological Information Cent

GRAPES Integrated Setting Experiment Tool(GISET)

- Experiment construction
- Experiment scheduling (ecFlow)
- Experiment sharing, statistics, compare
- Integrated code and experiment data management

- Design and implementation based on C/S mode
- Coded by python
- Back-end services run on servers

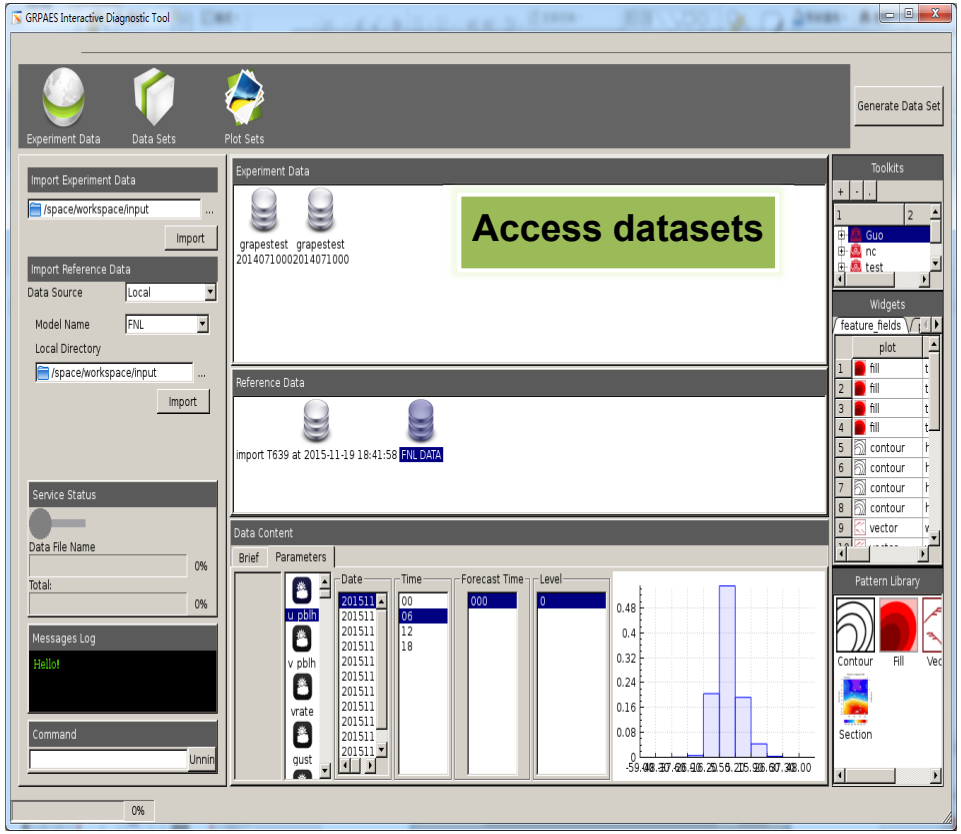


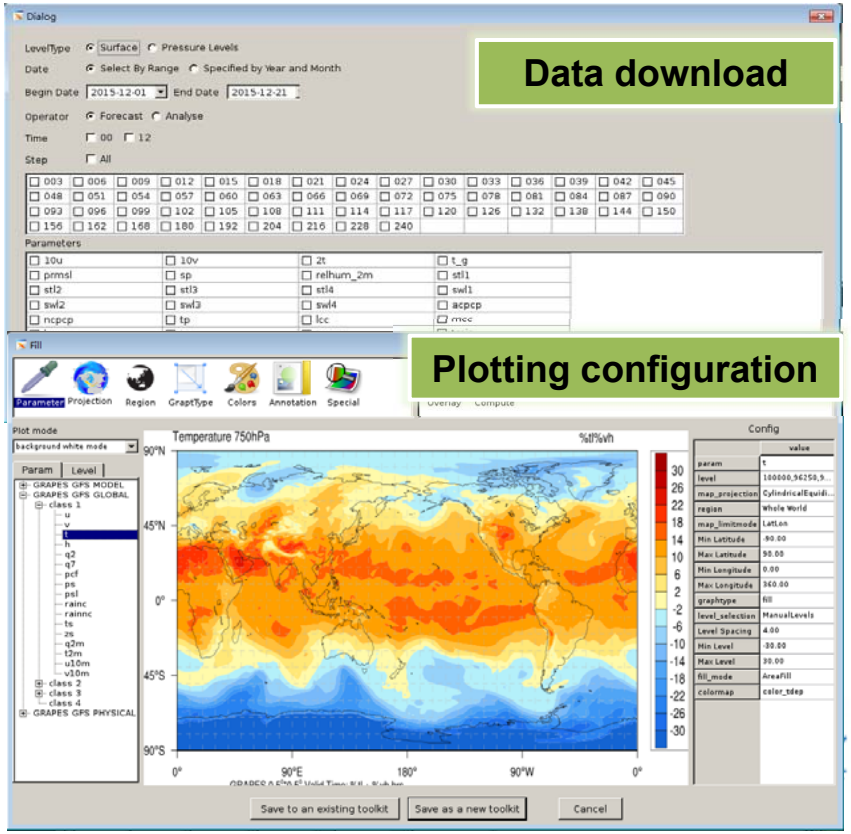


国家气象信息中心
National Meteorological Information Center

GRAPES Interactive Data Analytics Tool (GIDAT)

- On-line plotting of diagnostic data
- Interactive analytics function
- Access the datasets by data service API





Contents

- HPC Systems
- Model-Supportive Software Systems
- Conclusions

国家气象信息中心
National Meteorological Information Center

What have we achieved?

A new HPC solution has been deployed

- Architecture: CPU cluster; GPU cluster; Intel KNL cluster; network, environment, redundancy……
- Majority of migration work completed
- Testing novel architectures
- Collaboration: CMA members; vendors; universities

国家气象信息中心
National Meteorological Information Center

Next steps

- Efficient and portable code
- Test new architectures and programming models
- Software support services

国家气象信息中心
National Meteorological Information Center

Thank You for listening !

国家气象信息中心
National Meteorological Information Center