

---

# Big Data and Extreme Computing: Updates about China

---

Haohuan FU

March 9<sup>th</sup>, 2017



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

# Outline



HPC and Big Data Projects in China

Recent Efforts on Tianhe-2

Recent Efforts on Sunway TaihuLight

# MOST HPC Projects 2016

- Exa-scale pilot system (three systems in parallel)
  - Sunway successor
  - Tianhe successor
  - Sugon
- Computational physics model and numerical methods for exa-scale systems
- Programming framework
- HPC environment service and supporting system
- HPC numerical simulator
  - aircraft simulator
  - earth simulator
- Other typical HPC applications
  - complex EM environment simulation
  - ocean and wave modeling
  - mechanical engineering
  - material science

deep learning benchmark

310 Million, 19 Projects

# MOST Cloud Computing and Big Data Projects 2016

- Software defined cloud computing: basic theory and methods
- Big data storage technology and platform
- Dataflow oriented big data analytic system **389 Million, 15 Projects**
- Network operating system for cloud computing
- Scientific big data management system
- Big data management system for advanced manufacturing
- Big data based intelligent software development method and environment
- Big data knowledge engineering: basic theory and applications
- Human-computer interaction
- VR, AR



# Outline



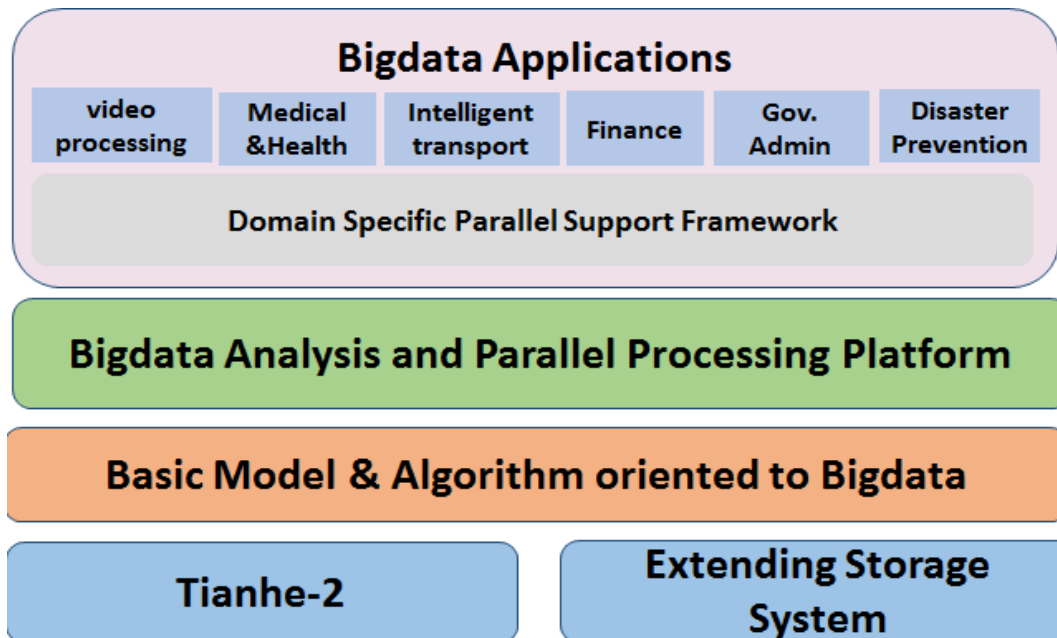
HPC and Big Data Projects in China

Recent Efforts on Tianhe-2

Recent Efforts on Sunway TaihuLight



- **NSFC-GuangDong Joint Project based on Tianhe-2**
  - 2016-2020, ¥300Million, steering by NSCC-GZ
- **Big data projects focus on Smart City**
  - Transportation 智能交通
  - Medical & Health 智慧医疗与健康
  - Disaster prevention 智慧防灾
  - Finance 智慧金融
  - Education 智慧教育
  - Social Management 智慧管理
- **Convergence of talent and technology resources, jointly solve key problems of big data science**



- **Bigdata + HPC**
- **Pearl River Delta National Bigdata Integration Test Zone**

# Outline



HPC and Big Data Projects in China

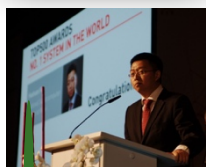


Recent Efforts on Tianhe-2



Recent Efforts on Sunway TaihuLight

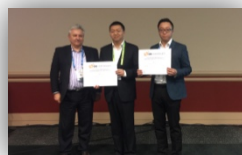
# 2016 Highlights



ISC 2016



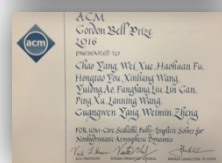
World Internet Conference



SC 2016



International Workshop on HPC Architecture, Software, and Application at an Extreme Scale



Gordon Bell Prize

Over 60 large-scale applications from over 100 research institutes covering 19 application domains, 6 full-scale applications, 18 half-scale, 22 million-core-scale, 3 Gordon Bell Finals, and 1 Gordon Bell Prize.

## Sunway TaihuLight: Overview

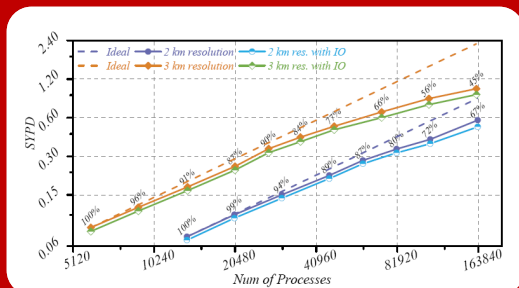
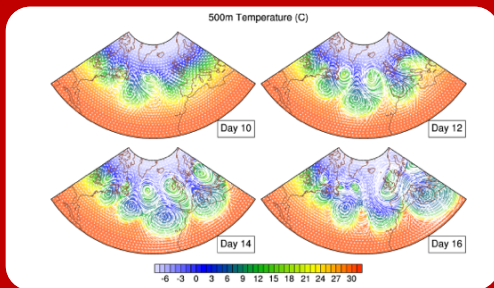
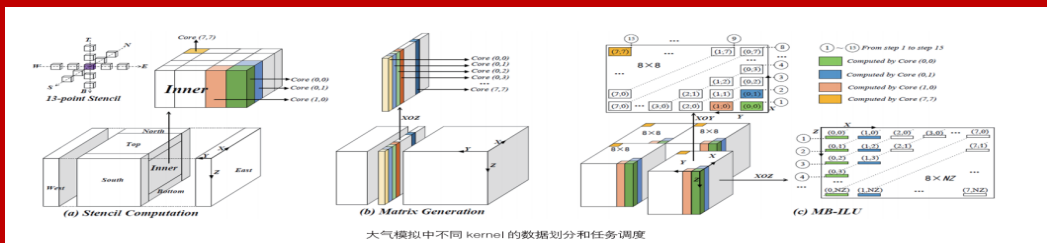
System	TaihuLight	Tianhe-2	Titan	Sequoia	Cori
Peak Performance (PFlops)	<b>125.4</b>	54.9	27.1	20.1	27.9
Total Memory (TB)	<b>1310</b>	1024	710	1572	879
Linpack Performance (PFlops)	<b>93.0(74%)</b>	33.9(62%)	17.6(65%)	17.2(85.3)	14.0(50%)
Performance/Power (Mflops/W)	<b>6051.3</b>	1901.5	2142.8	2176.6	3266.8
GTEPS	<b>23755.7</b>	2061.48	###	23751	###
HPCG (Pflops)	<b>0.3712</b>	0.5801	0.3223	0.3304	0.3554
Rank of Top500	<b>1</b>	2	3	4	5
Rank of Green500	<b>4</b>	135	100	90	26
Rank of Graph500	<b>2</b>	8	###	3	###
Rank of HPCG	<b>4</b>	2	7	6	5

# Atmospheric Modeling

Institute of Software, CAS  
Tsinghua University  
Beijing Normal University

10-million-core scalable full implicit solver for non-hydrostatic atmospheric dynamics

- support up to 500m resolution
- sustained performance of 7.95 Pflops
- 2016 Gordon Bell Prize winner

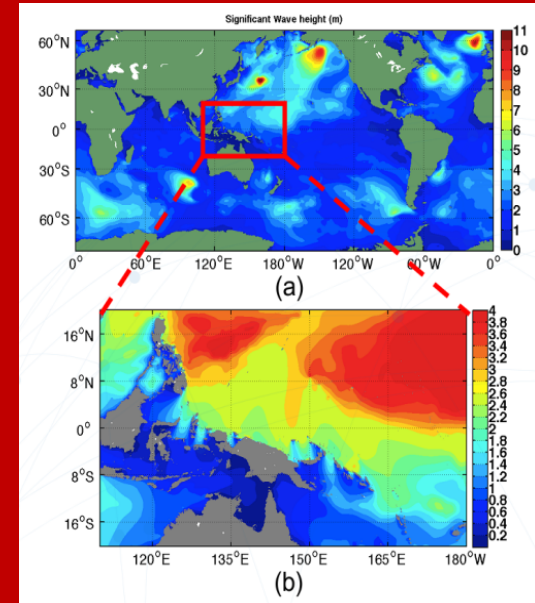
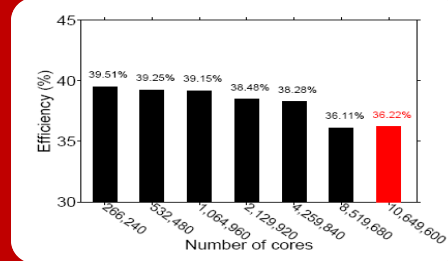
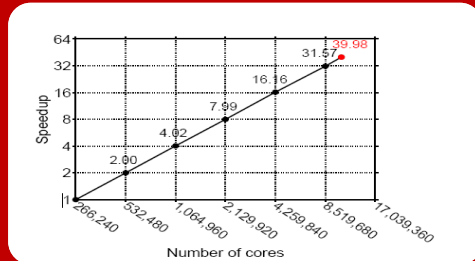
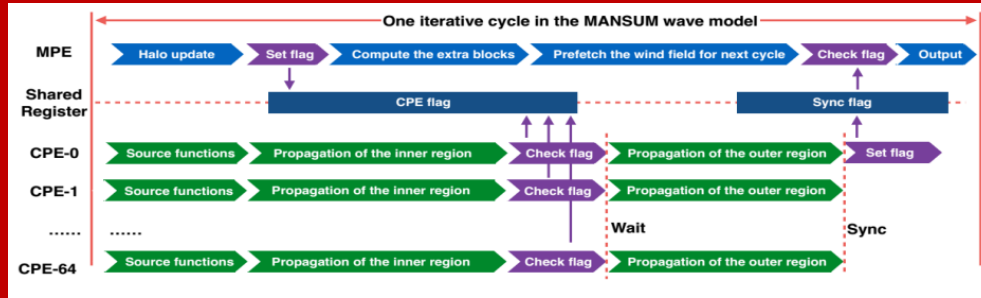


# Wave Model

First Institute of Oceanography (FIO)  
Tsinghua University

## MASNUM (Key laboratory of MARine Science and NUMerical Modeling) wave model

- sustained performance 45.43 Pflops
- global 1km wave simulation
- 2016 Gordon Bell Prize finalist



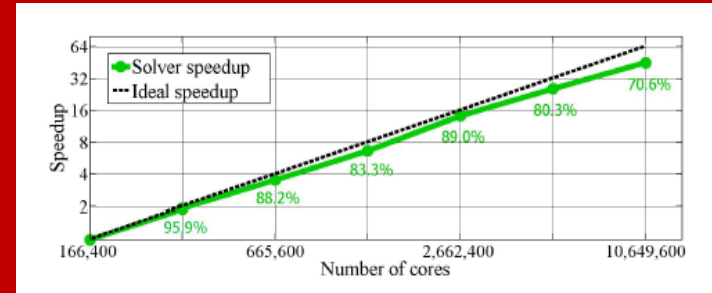
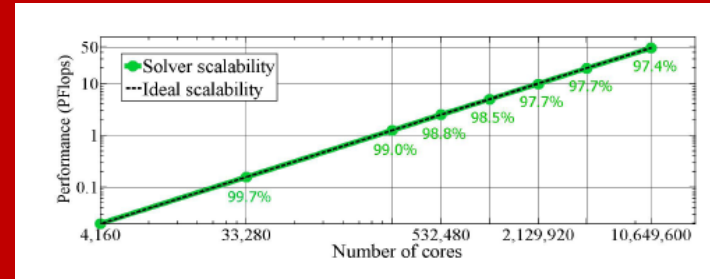
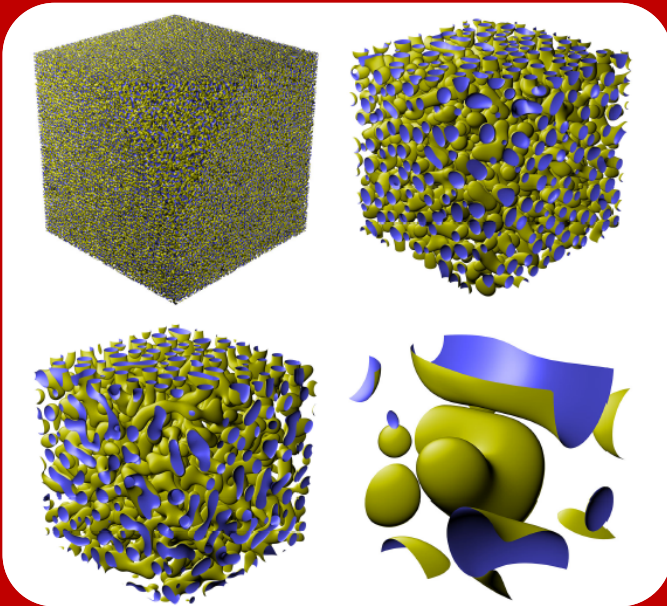


# Phase Field Simulation

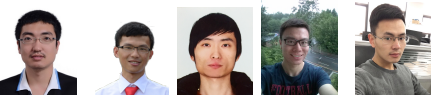
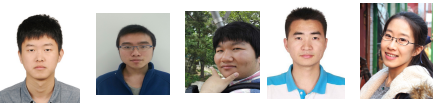
Computer Network  
Information Center, CAS

## Large Scale Phase Field Simulation for Coarsening Dynamics Based on Cahn-Hilliard Equation with Degenerated Mobility

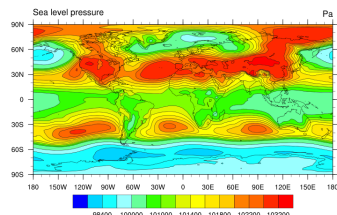
- over 50 Pflops sustained performance
- highly scalable, large time-step integrating algorithm



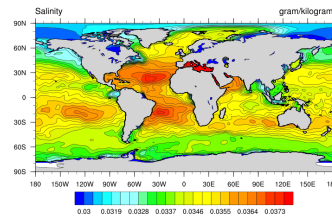
# The CESM Project on Sunway TaihuLight



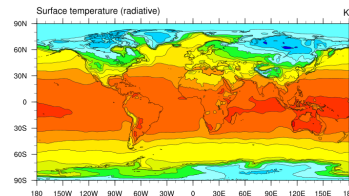
Tsinghua + BNU  
30+ Professors and  
Students



F Compset



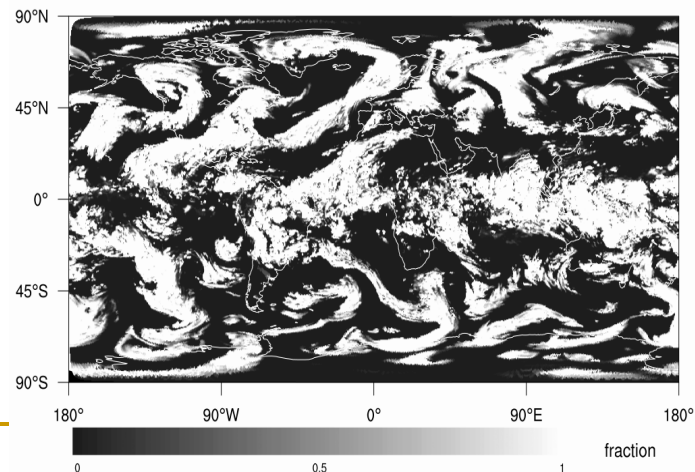
G Compset



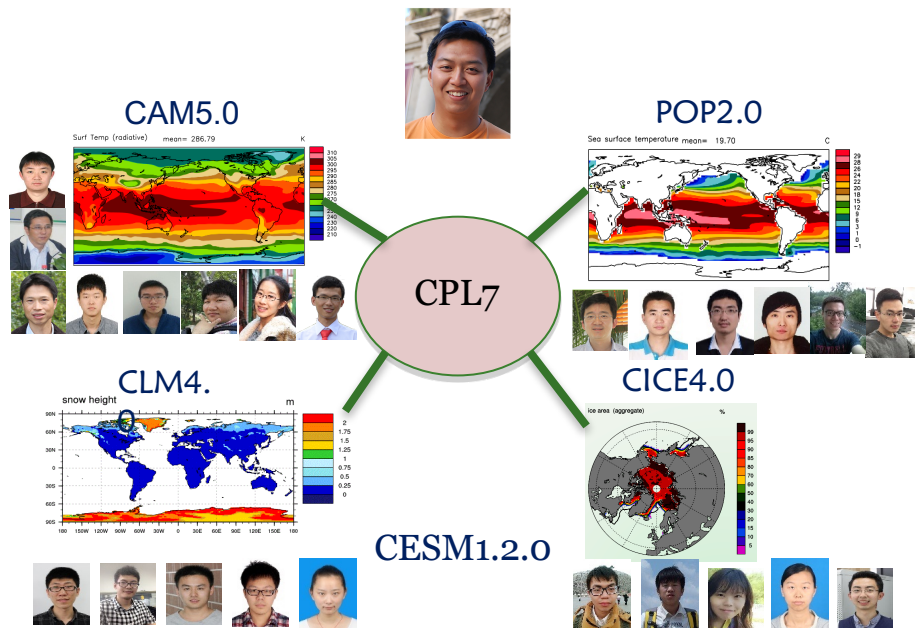
B Compset

- Four component models, millions of code

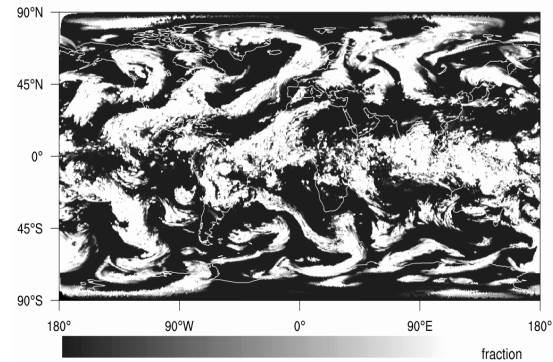
- Large-scale run on Sunway TaihuLight
- 24,000 MPI processes
- Over one million cores
- 10-20x speedup for kernels
- 2-3x speedup for the entire model



# The CESM Project on Sunway TaihuLight



Tsinghua + BNU **30+ Professors and Students**



- Four component models, millions lines of code
- Large-scale run on Sunway TaihuLight
  - **24,000** MPI processes
  - Over **one million** cores
- **10-20x** speedup for kernels
- **2-3x** speedup for the entire model

# Library for Deep Learning (swDNN)

- swDNN: Provide interface for optimized basic operators
  - Fully-connected layer (BLAS); Pooling layer
  - Activation function; Batch Normalization
  - **\*Convolutional Layer(90% time for CNN)**

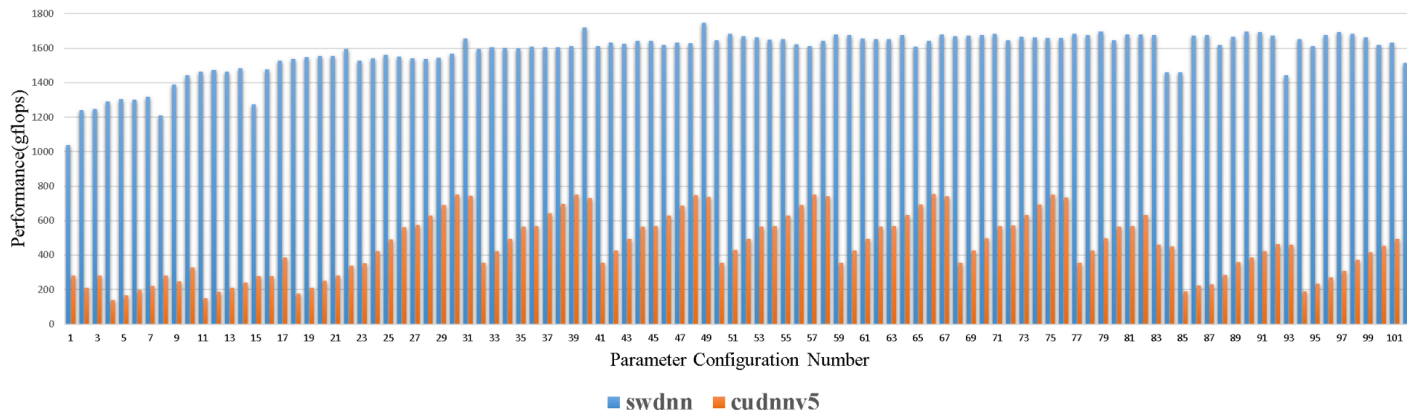
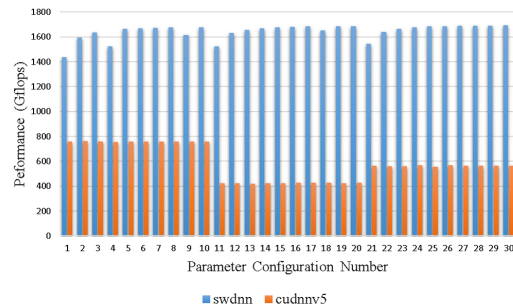
## Related Works on other architectures

Work	Platform	Method
cuDNN(2014)	GPU	GEMM
fbtfft(2014)	GPU	FFT
Andrew Lavin (2015)	GPU	Winograd
Chen Zhang (2015)	FPGA	Direct Conv
swDNN	SW26010	Blocking GEMM

# Library for Deep Learning (swDNN)

## ■ Performance

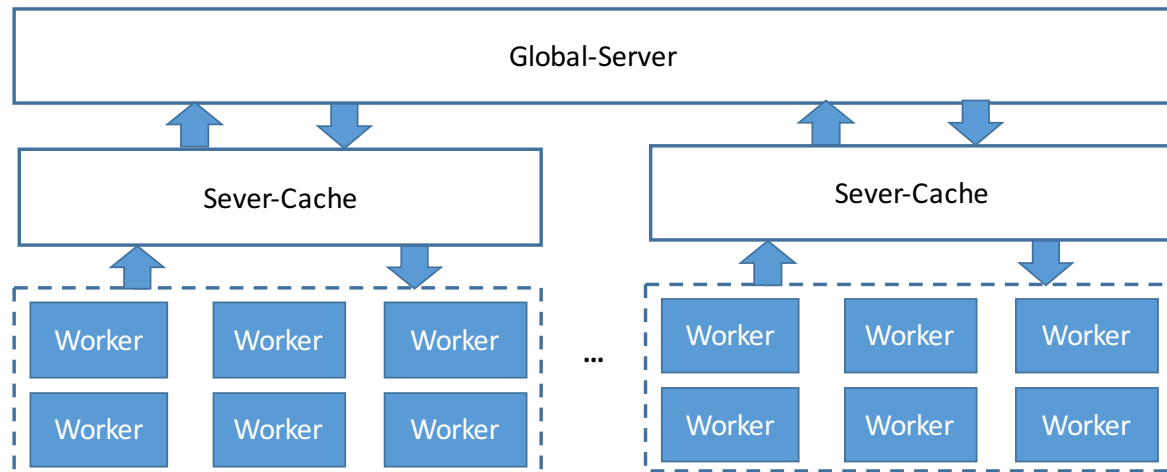
- Convolutional performance above **1.6 Tflops** with double-precision
- Speedup ranging from **1.91x** to **9.75x** compared with cudnn5.1.



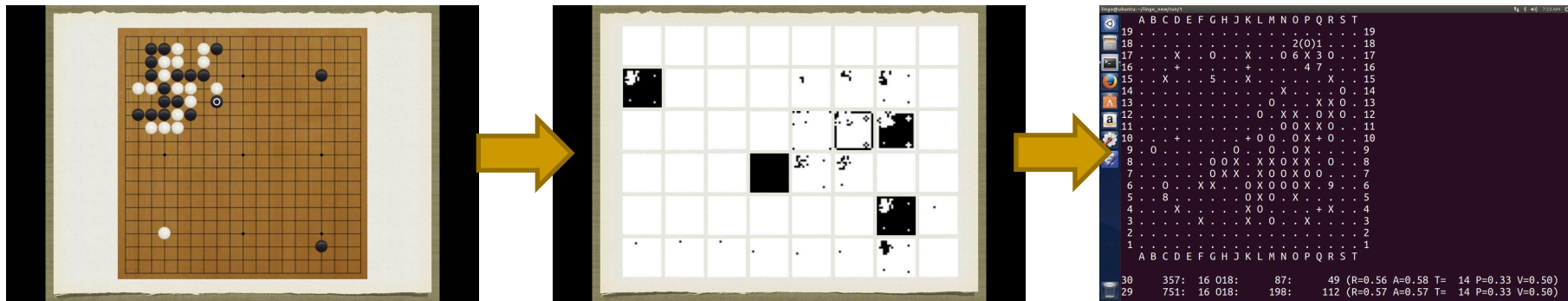
# Framework for Deep Learning (under development)

## ■ Distributed framework

- Customized from *Caffe* with less dependencies
- Two-level Parameter Server Based-on MPI



# swDNN Supported Project: Sunway-Lingo collaborated with Prof. Zhiqing Liu, BUPT



- Original go board to be processed
- Converted to a 48-channel image fed to deep CNN with essential go features such as liberties
- Order of probabilities of plausible moves as outputted by policy network

# Long Term Plan

- Traditional HPC Applications
  - weather / climate service
  - seismic data processing service
  - CFD simulation framework for Advanced Manufacturing
- Deep Learning Related Applications
  - the swDNN framework
  - collaborating with face++ for face recognition applications
  - collaborating with Sogou for voice recognition and translation
  - customized DNN Sunway chip?
- Big Data Center
  - National Health and Medical Big Data Center at Nanjing