

# 并行计算

## Parallel Computing

主讲人 孙广中  
Spring, 2018

# 并行计算——结构·算法·编程

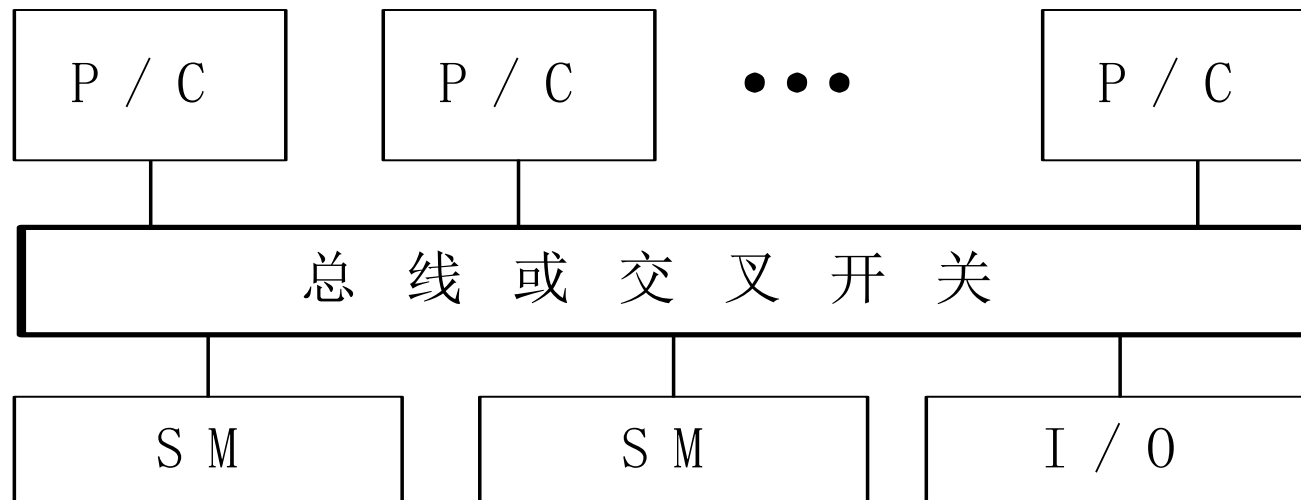
- 第一篇 并行计算的基础
  - 第一章 并行计算与并行计算机结构模型
  - 第二章 并行计算机系统互连与基本通信操作
  - 第三章 典型并行计算机系统介绍
  - 第四章 并行计算性能评测

# 第三章 典型并行计算机系统介绍

- 3.1 共享存储多处理机系统
  - 3.1.1 对称多处理机SMP结构特性
- 3.2 分布存储多计算机系统
  - 3.2.1 大规模并行机MPP结构特性
- 3.3 分布共享存储多计算机系统
  - 3.3.1 分布共享存储计算机系统特性
- 3.4 机群系统
  - 3.4.1 大规模并行处理系统MPP机群SP2
  - 3.4.2 工作站机群COW

## 对称多处理器SMP（1）

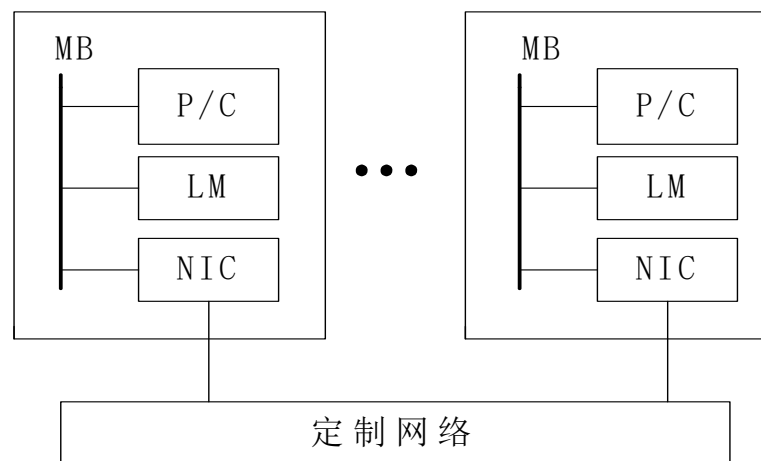
- SMP: 采用商用微处理器，通常有片上和片外Cache，基于总线连接，集中式共享存储，UMA结构
- 例子：SGI Power Challenge, DEC Alpha Server, Dawning 1



# 大规模并行机MPP

- 成百上千个处理器组成的大规模计算机系统，规模是变化的。
- NORMA结构，高总计带宽，相对低延迟，定制互连。
- 可扩放性：Processors, Memory, Bandwidth, I/O, 平衡设计
- 系统成本：商用处理器，相对稳定的结构，SMP节点，分布
- 通用性和可用性：不同的应用，PVM，MPI，交互，批处理，互连对用户透明，单一系统映象
- 通信要求：高于标准的LAN
- 较大存储器和I/O能力
- 现在MPP与Cluster难以区别
- 例子：Intel Option Red

IBM SP2, Dawning 1000



# DSM计算机系统特性

- DSM结构特性
  - 共享存储系统采用分布共享，减少集中共享的冲突
  - 采用高速缓存来缓和由共享引起的冲突和分布存储引起的长延迟
  - 保持了共享编程的方便性和软件的可移植性
- 存储一致性问题
  - 非均匀存储访问和高速缓存一致性问题
  - 影响了一些技术的应用和系统的可扩放性
- DSM系统分类
  - 硬件实现的共享存储：CC-NUMA、NCC-NUMA、COMA
  - 软件实现的共享存储：共享虚拟存储(SVM)
- 典型机器：SGI Origin 2000
- 典型软件：JiaJia（佳佳）（实现锁的高速缓存一致性协议）

- 国家最高科学技术奖获奖人金怡濂
- “最强大脑”！它的国防意义堪比“两弹一星”
- “中国超算冠军宝座或难保 应用远滞后于能力发展”
- “全国人大代表吴立新：应尽快启动E级超算建设”

超级计算机

天河二号



## 天河二号 (Tianhe-2/MilkyWay-2) · 建造情况

网站	<a href="http://www.nscg-gz.cn/">http://www.nscg-gz.cn/</a>
制造商	国防科大 (NUDT) & Inspur
位置	国家超算中心 (广州) NSCC-GZ
占地面积	720 m <sup>2</sup>
研发人员	约1300人
耗资	\$390 million
首次进入Top500	2013.6
概况	170个机柜, 包括125个计算机柜、8个服务机柜、13个通信机柜、24个存储机柜



## 天河二号 (Tianhe-2/MilkyWay-2) · 机器情况

Hardware		Software	
Architecture	TH-IVB-FEP Cluster	Operating System	Kylin Linux
Processor	Intel Xeon E5-2692v2 12C 2.2GHz	Compiler	icc
Accelerator	Intel Xeon Phi 31S1P	Math Library	Intel MKL-11.0.0
Cores	3,120,000	Power Consumption	
Memory	1,024,000 GB		
Interconnect	TH Express-2	Power	17,808.00 kW

Performance	
Linpack Performance(Rmax)	33,862.7 TFlop/s
Theoretical Peak (Rpeak)	54,902.4 TFlop/s
Nmax	9,960,000

Upper bound on performance

Measure of floating-point rate of execution

G T P E Z Y

# 天河二号 (Tianhe-2/MilkyWay-2) · in the Top500

Performance  
&  
Rank

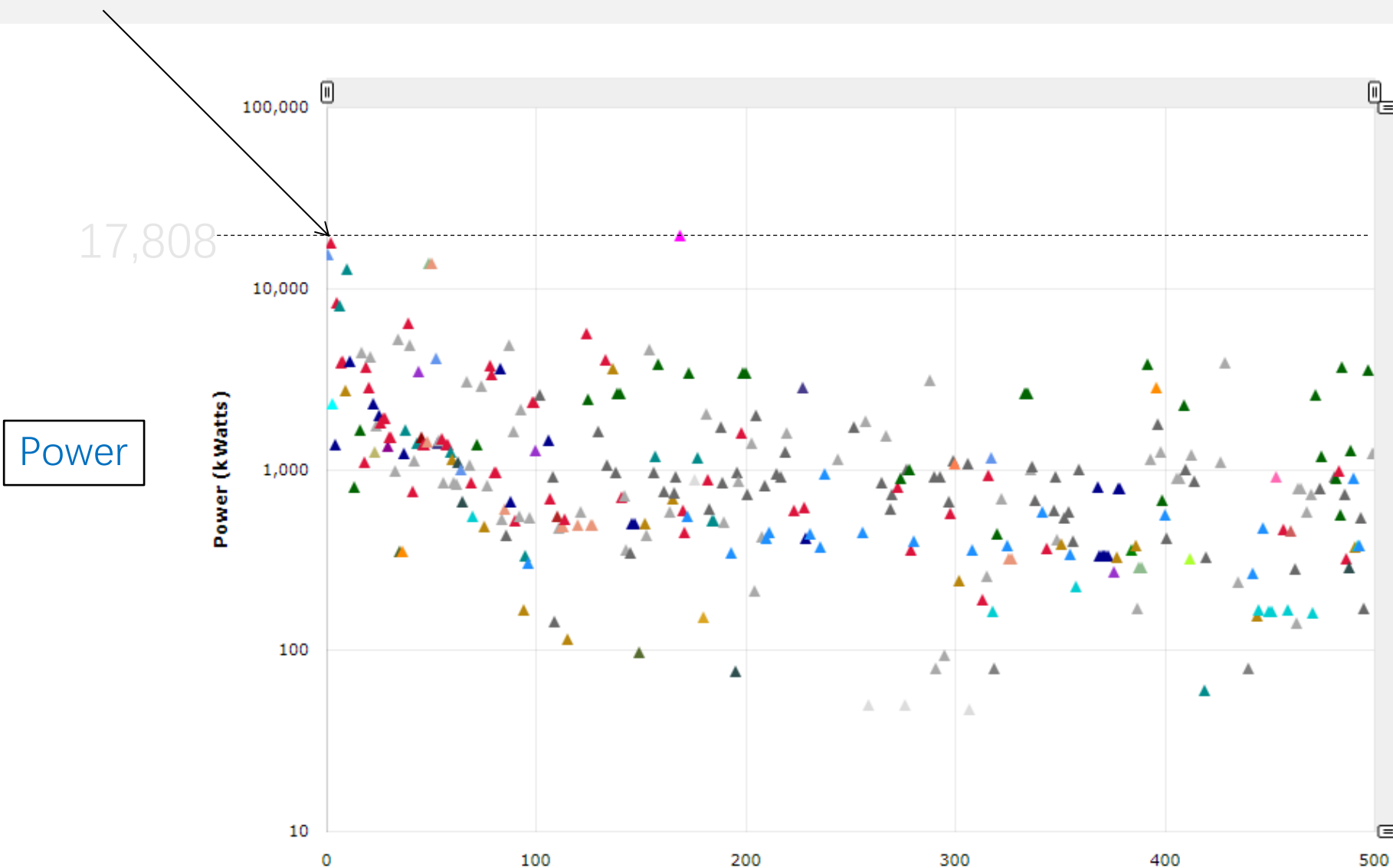
#1

2013.6~2015.11

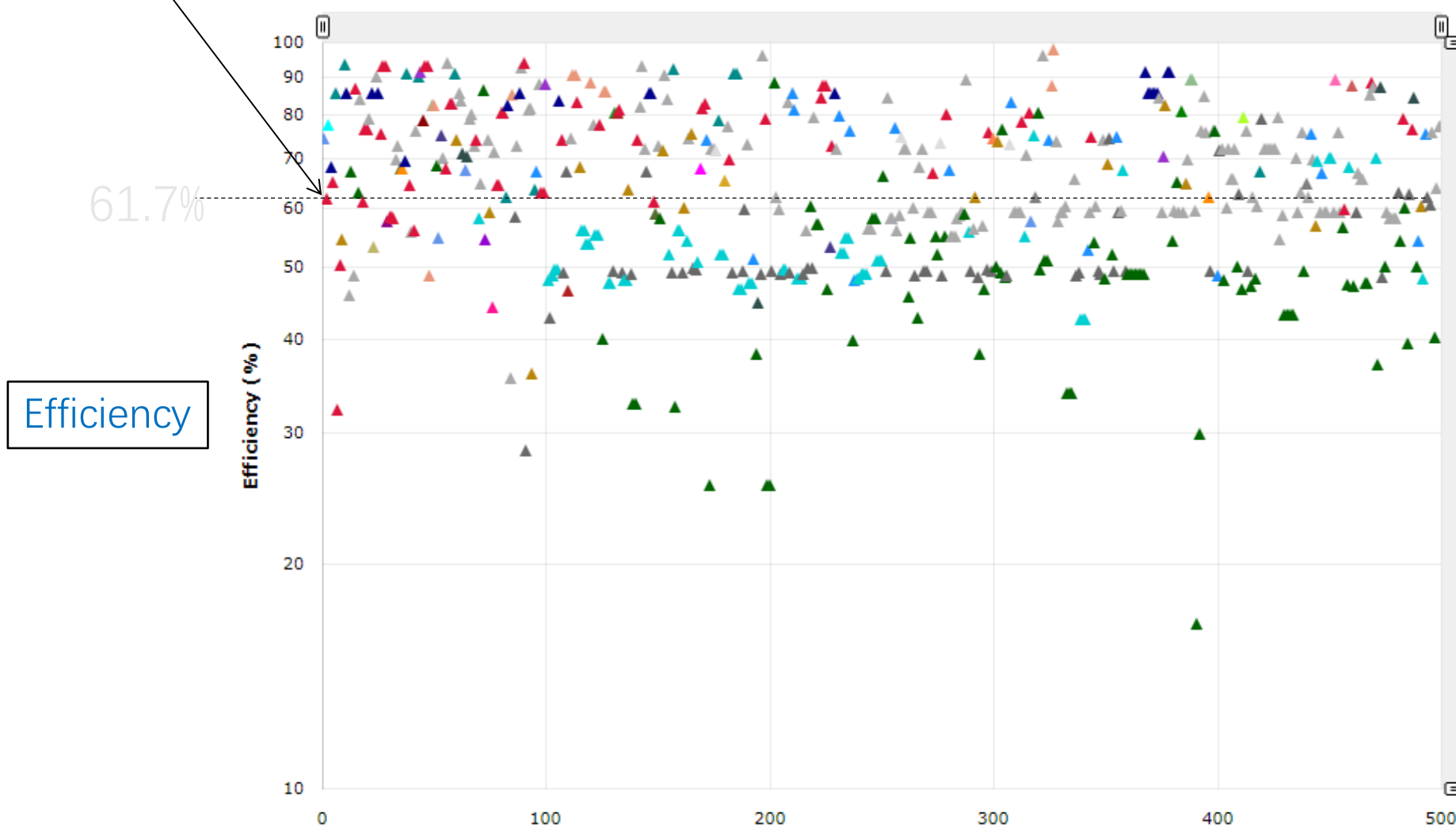
G T P E Z Y



# 天河二号 (Tianhe-2/MilkyWay-2) ▪ in the Top500



# 天河二号 (Tianhe-2/MilkyWay-2) ▪ in the Top500

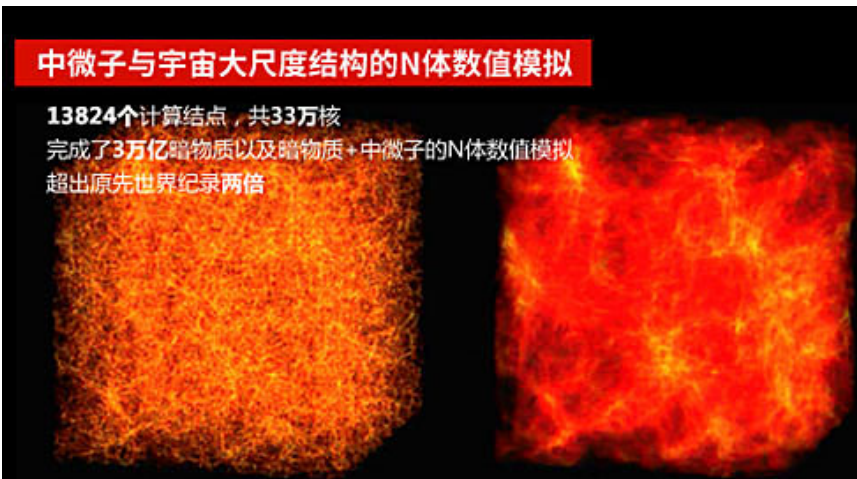




# 天河二号 (Tianhe-2/MilkyWay-2) · 应用

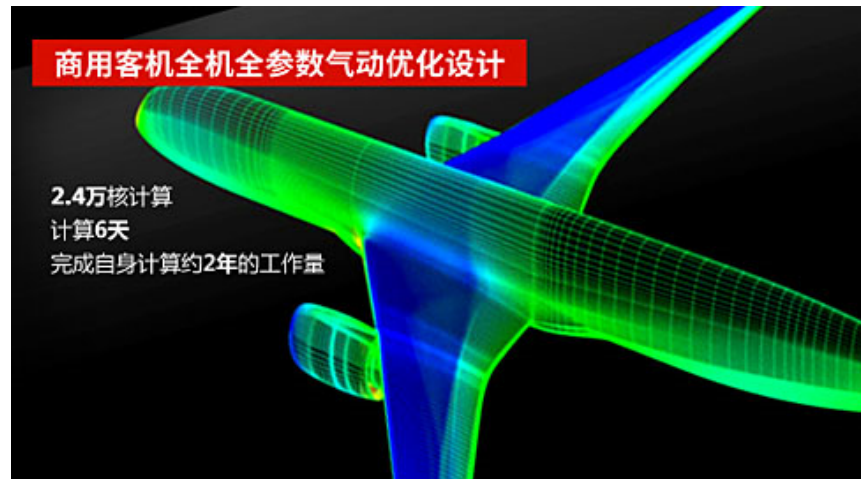
## 中微子与宇宙大尺度结构的N体数值模拟

13824个计算结点, 共33万核  
完成了3万亿暗物质以及暗物质+中微子的N体数值模拟  
超出原先世界纪录两倍



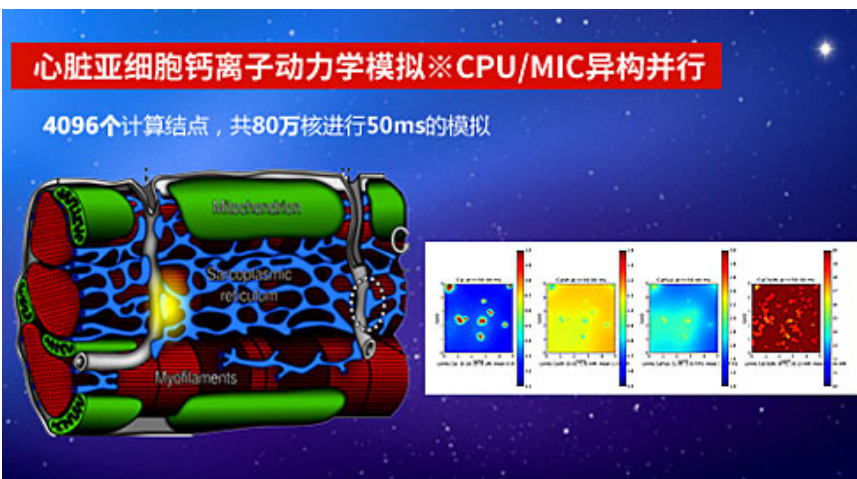
## 商用客机全机全参数气动优化设计

2.4万核计算  
计算6天  
完成自身计算约2年的工作量



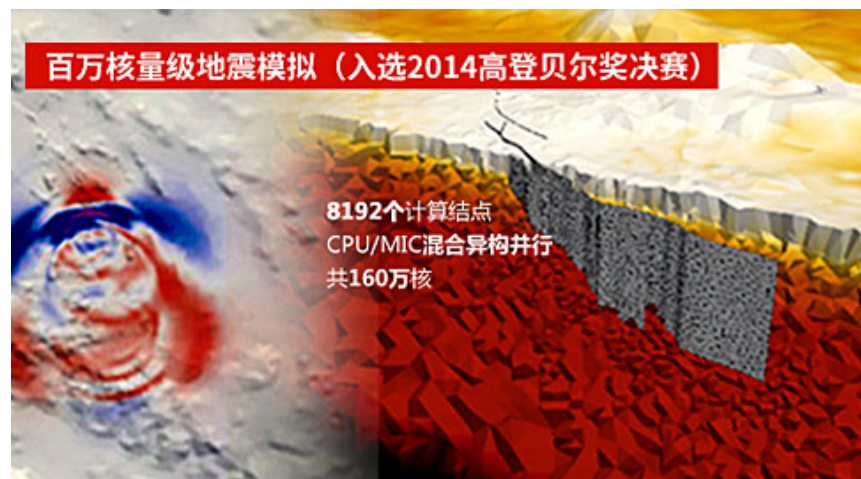
## 心脏亚细胞钙离子动力学模拟※CPU/MIC异构并行

4096个计算结点, 共80万核进行50ms的模拟



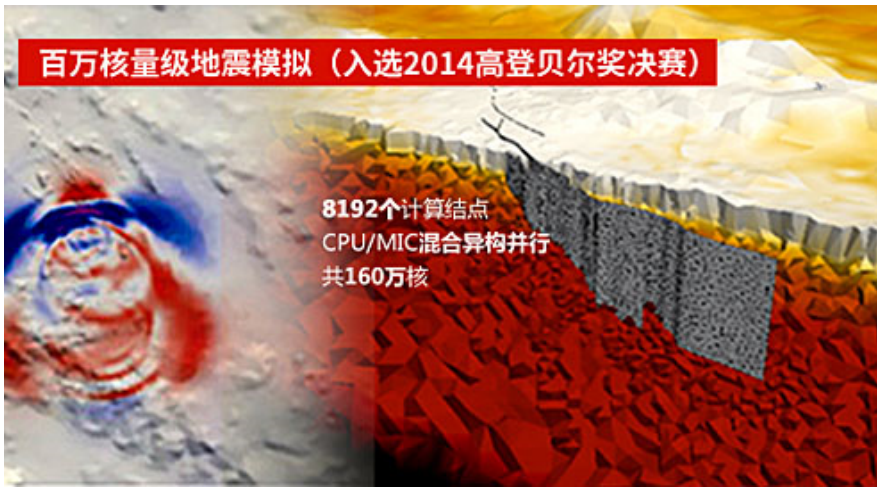
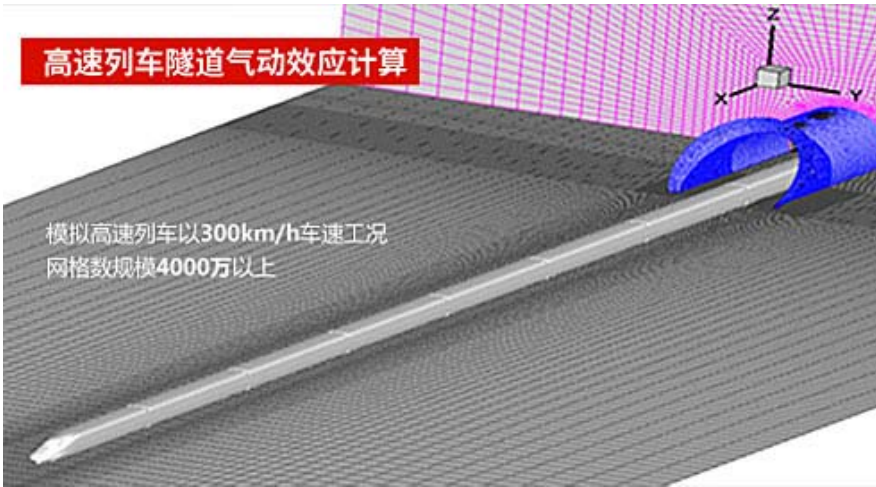
## 百万核量级地震模拟 (入选2014高登贝尔奖决赛)

8192个计算结点  
CPU/MIC混合异构并行  
共160万核





## 天河二号 (Tianhe-2/MilkyWay-2) · 应用



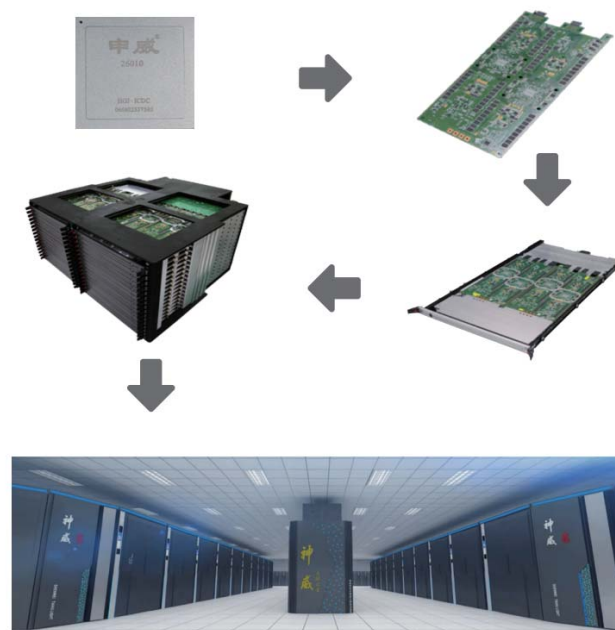
超级计算机

神威·太湖之光



# 神威·太湖之光（Sunway TaihuLight）· 建造情况

网站	<a href="http://www.nscctx.cn/">http://www.nscctx.cn/</a>
制造商	国家并行计算机工程技术研究中心（NRCPC）
位置	国家超算中心（无锡，江苏）NSCC-Wuxi
耗资	\$273 million
首次进入Top500	2016.6



## 神威·太湖之光 (Sunway TaihuLight) · 机器情况

Hardware		Software	
Architecture	Sunway MPP !	Operating System	Sunway RaiseOS
Processor	Sunway SW26010 260C 1.45GHz !	Parallel Programming	MPI, OpenMP, OpenACC !
Cores	10,649,600		
Memory	1,310,720 GB		
Interconnect	Sunway !		
		Power Consumption	
		Power	15,371.00 kW

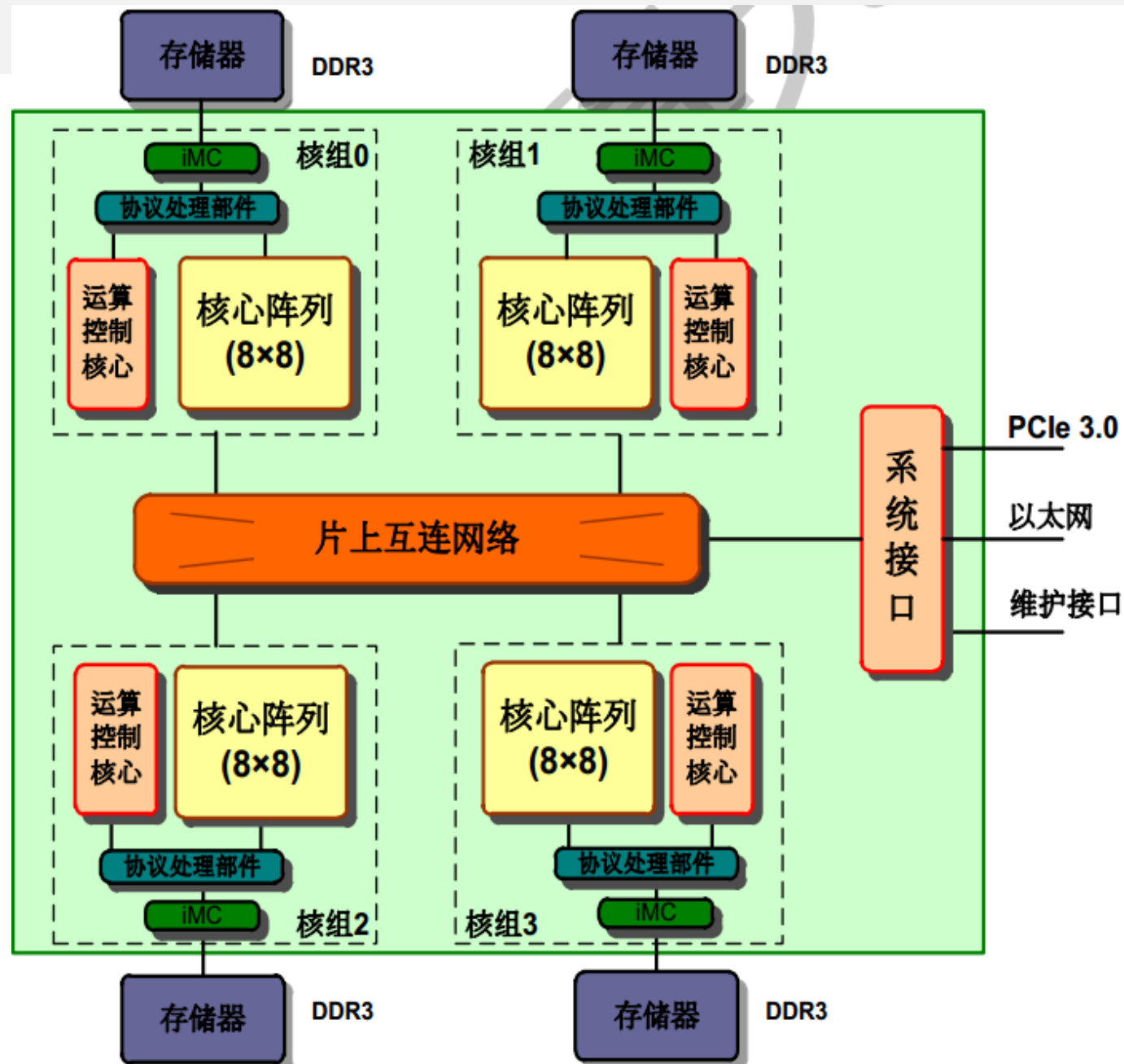
Performance	
Linpack Performance(Rmax)	93,014.6 TFlop/s
Theoretical Peak (Rpeak)	125,436 TFlop/s
Nmax	12,288,000

Upper bound on performance

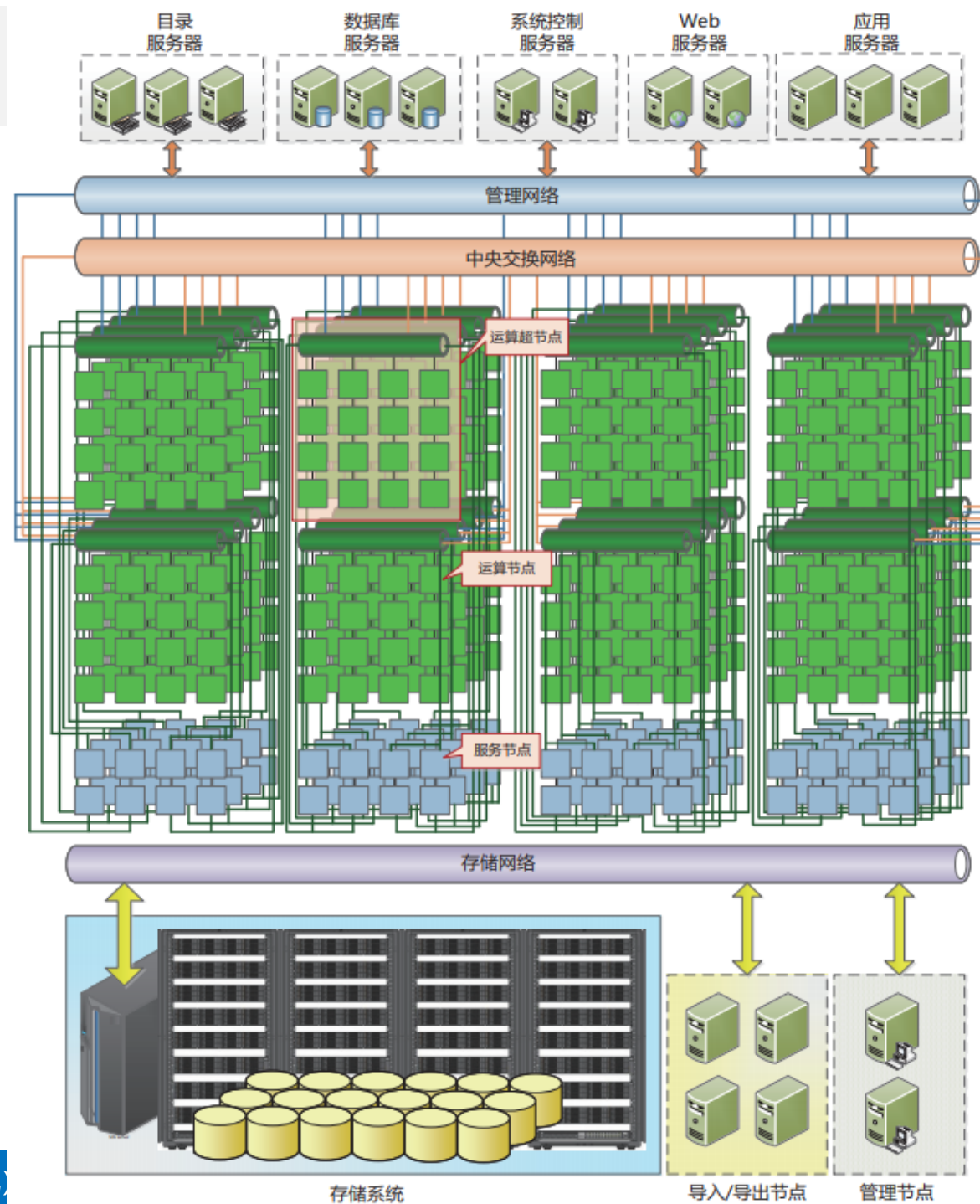
Measure of floating-point rate of execution

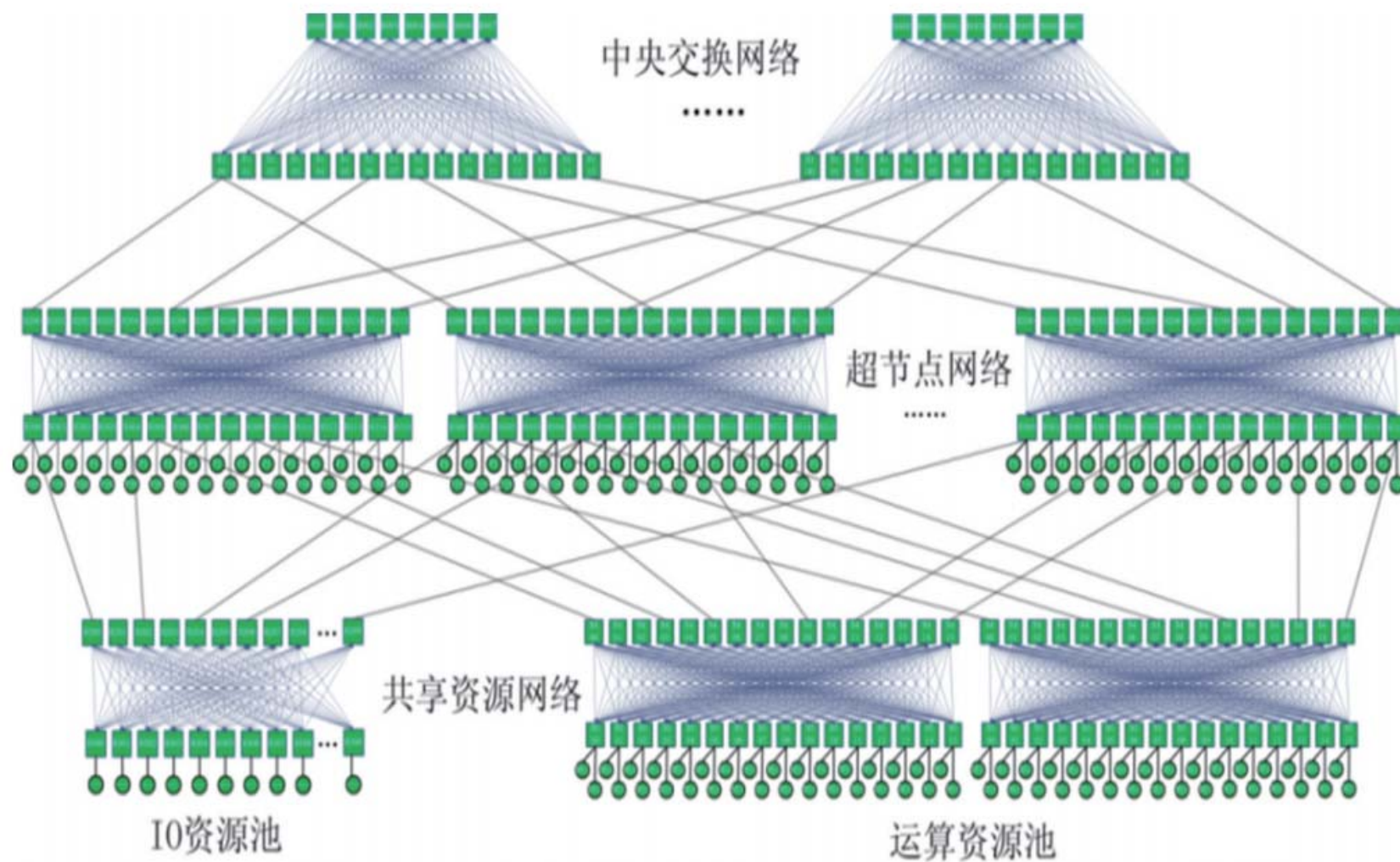
G T P E Z Y

“申威 26010”  
众核处理器



# 整体设计





互连方式



## 软件系统

### 并行应用

#### 并行开发环境

- 集成开发环境
- 并行调试
- 性能监测

#### 并行语言及编译环境

- OpenACC
- MPI
- OpenMP

#### 国产众核CPU基础软件

##### 基础编译系统

- C / C++, Fortran
- SIMD 扩展接口
- 异构代码生成

##### 基础函数库

- C 库
- 加速线程库
- 数学库

##### 自动向量化系统

- C / C++, Fortran
- 循环级向量化
- 向量代码优化

#### 并行操作系统环境

- 作业管理
- 容错管理
- 资源管理
- 系统开工
- 功耗管理
- 安全管理
- 网络管理

#### 高性能存储管理系统

- SWGFS并行文件系统
- LWFS 轻量级文件系统
- 存储管理平台

# 神威·太湖之光 (Sunway TaihuLight) · in the Top500

Performance  
&  
Rank

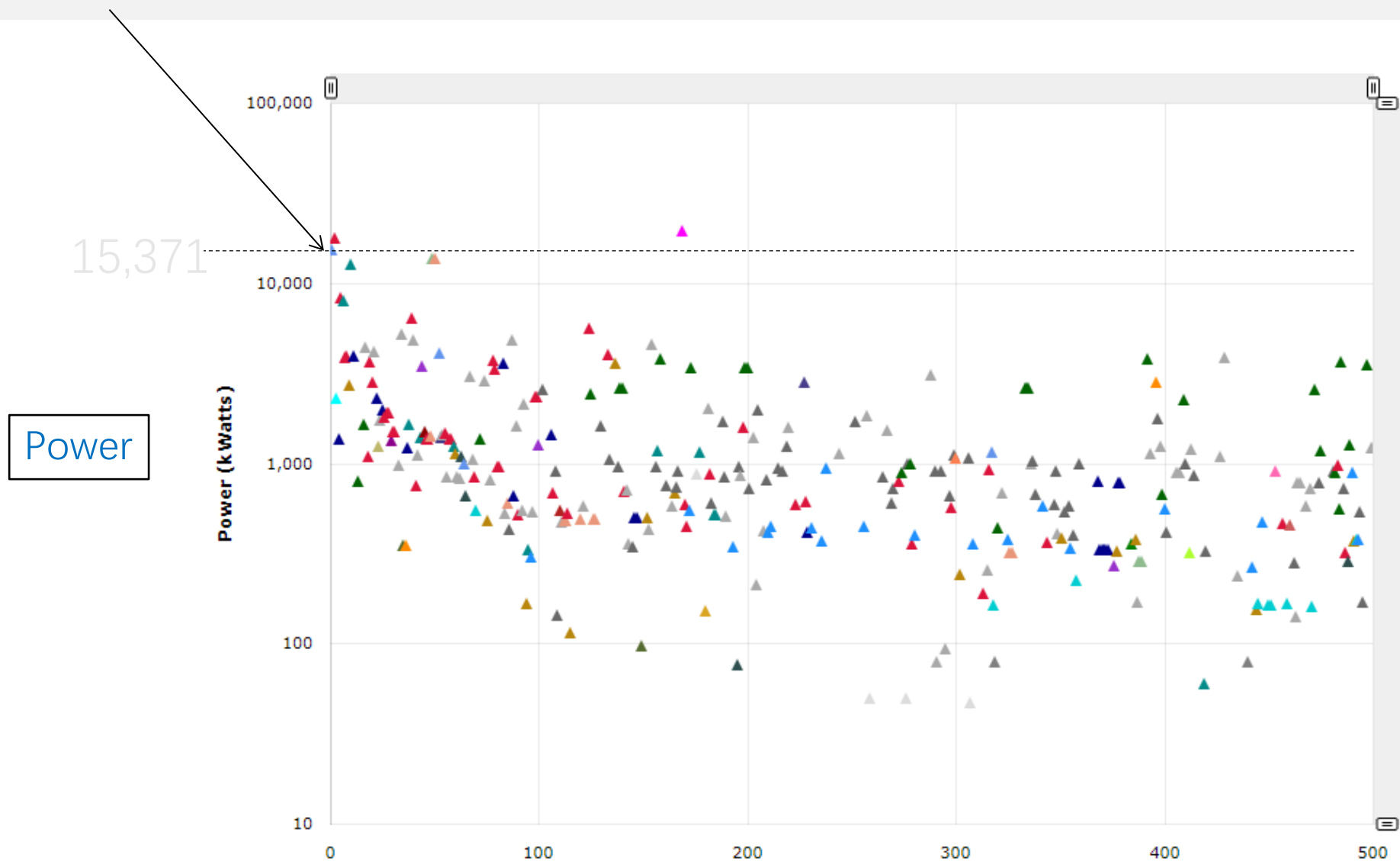
#1

2016.6~2017.11

G T P E Z Y

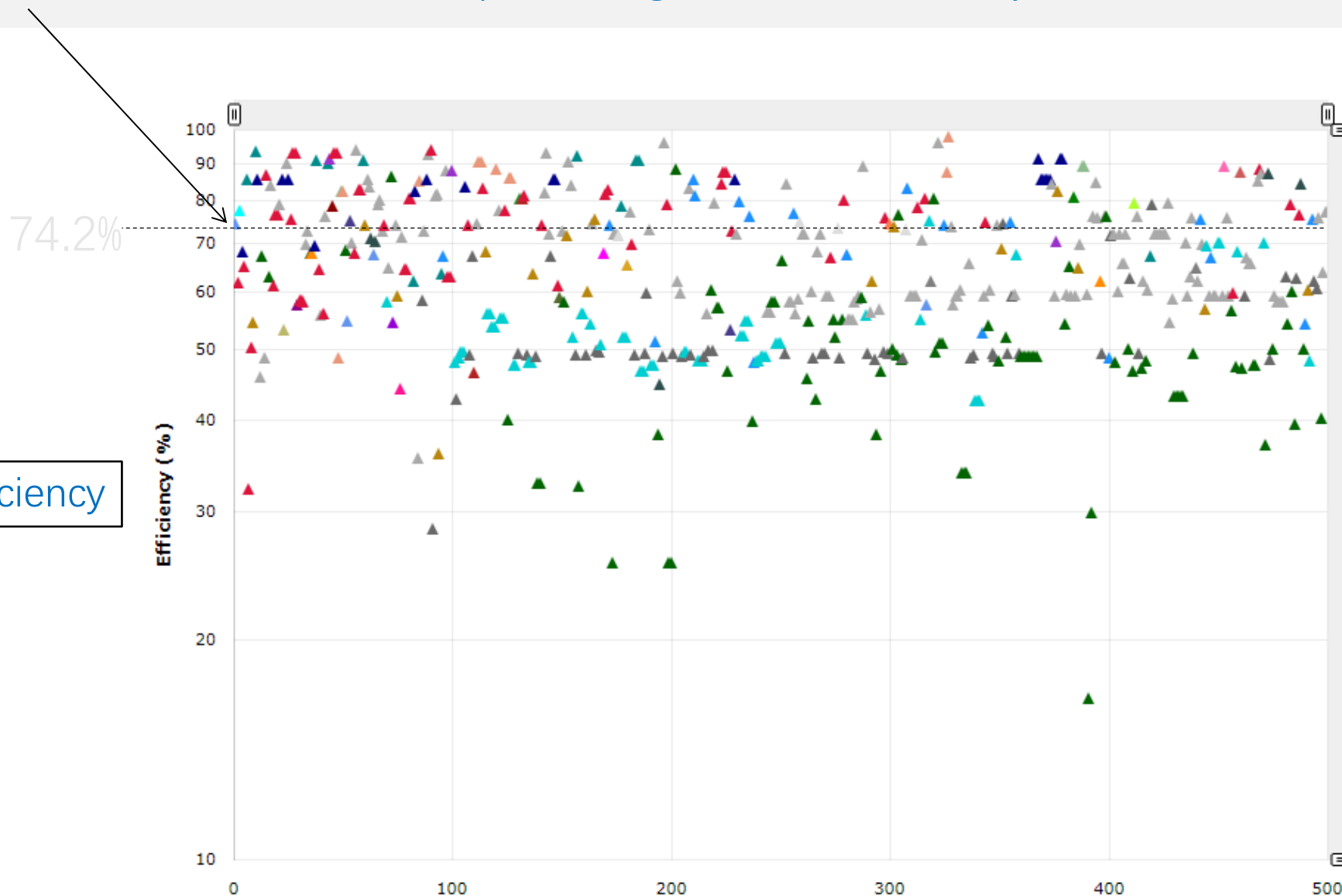


# 神威·太湖之光 (Sunway TaihuLight) · in the Top500

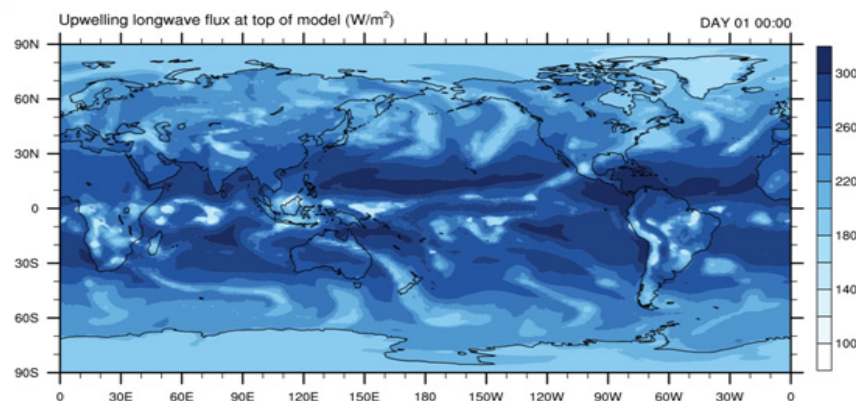




# 神威·太湖之光 (Sunway TaihuLight) · in the Top500



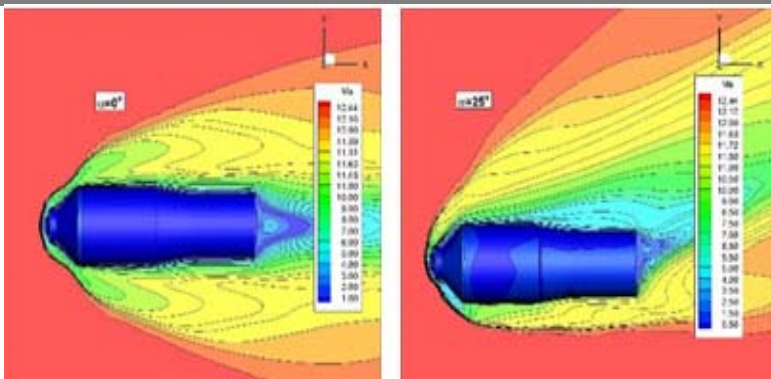
# 神威·太湖之光 (Sunway TaihuLight) · 应用



基于国产平台的国产地球系统  
模式



真实感动漫渲染系统

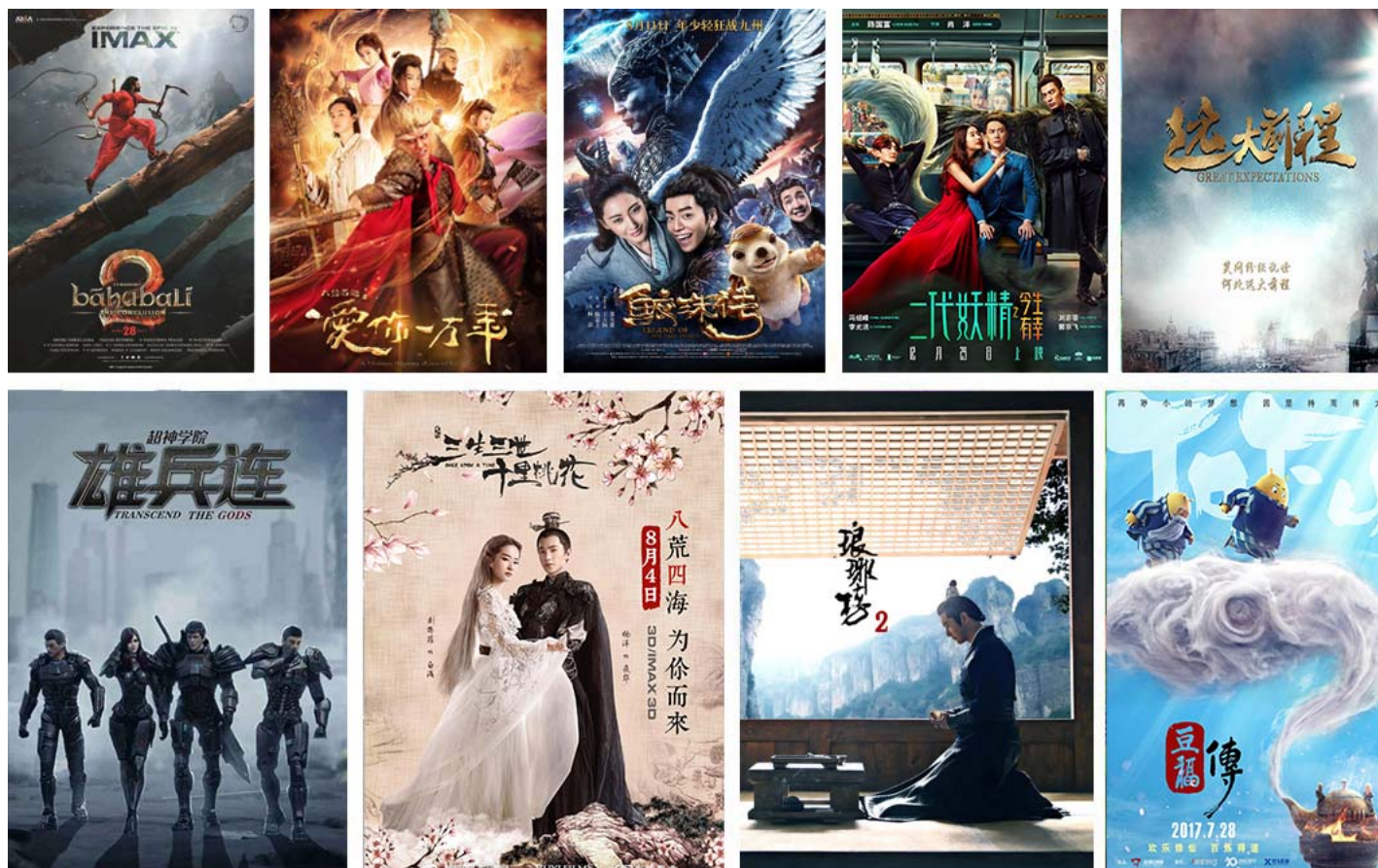


航天飞行器统一算法数值



岛礁建设浮式平台

# 神威·太湖之光 (Sunway TaihuLight) · 应用



## 渲染服务影视案例

超大规模云渲染



## 参考资料

<https://www.top50>

<http://www.nscg-gz.cn>

<http://www.nscgwx.cn>

神威·太湖之光 宣传手册 

<https://en.wikipedia.org/wiki/Tianhe-2>

<https://en.wikipedia.org/wiki/Sunway>

Visit to NUDT, Jack 

<http://tech.hexun.com/2013-07-17/156240746.html>