



中山大學  
SUN YAT-SEN UNIVERSITY



国家超级计算广州中心  
NATIONAL SUPERCOMPUTER CENTER IN GUANGZHOU

# 超算与大数据融合创新发展

杜云飞

国家超级计算广州中心 总工程师  
中山大学 数据科学与计算机学院 教授 博导  
[yunfei.du@nscg-zh.cn](mailto:yunfei.du@nscg-zh.cn)

[www.nscg-zh.cn](http://www.nscg-zh.cn)

# OUTLINE

## 主要内容

国家超级计算广州中心

NATIONAL SUPERCOMPUTER CENTER IN  
GUANGZHOU

---

### 1、超算的发展

---

### 2、超算与大数据的技术

---

### 3、超算与大数据的应用

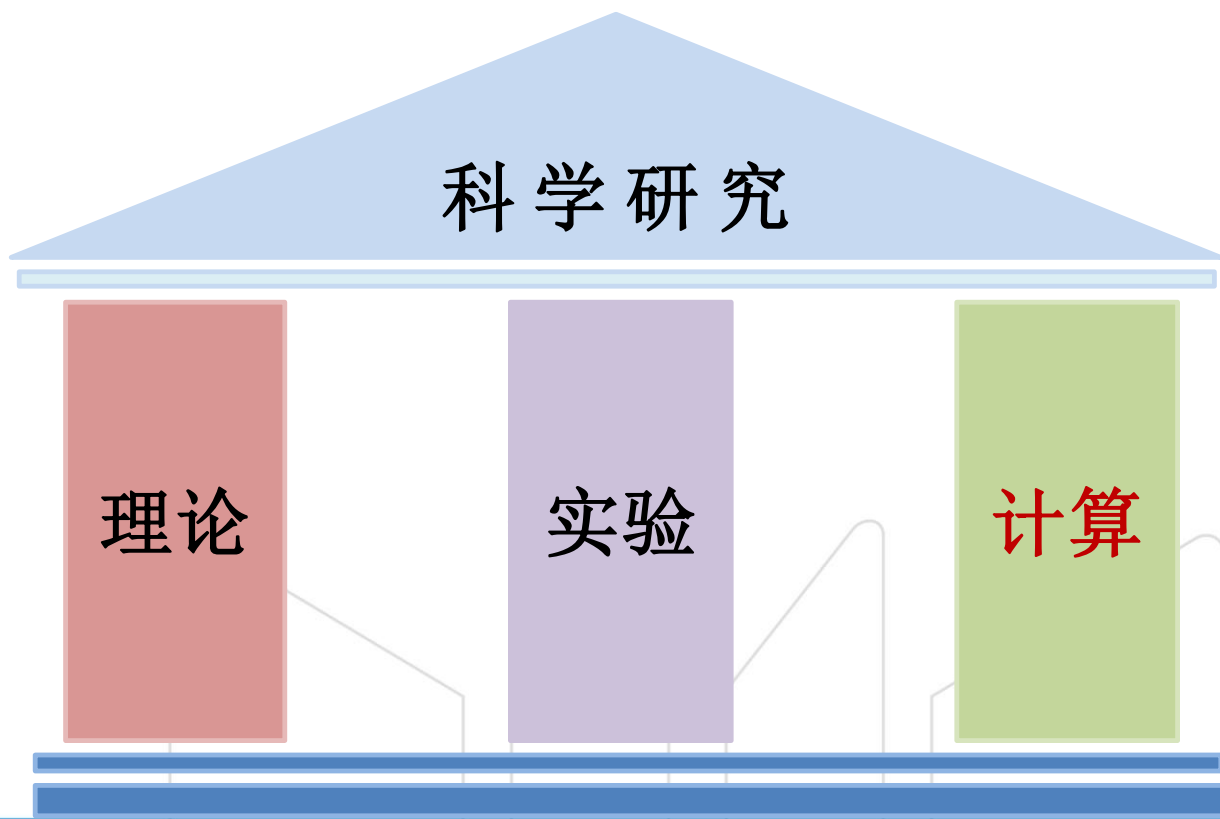
---

### 4、超算与大数据的未来

---

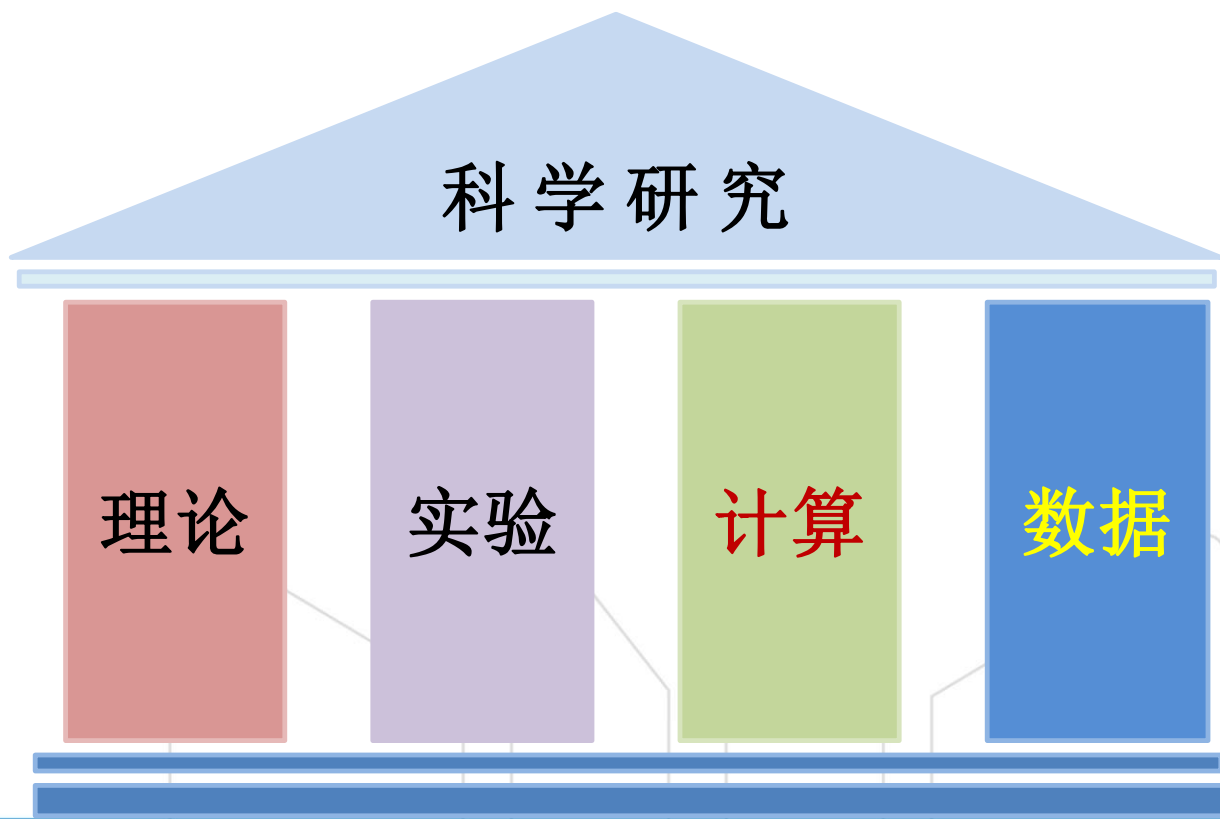
## ● 超级计算机对科学发现、技术创新、产业革命的重要作用

— 高性能计算：是科学研究的三大手段之一



## ● 超级计算机对科学发现、技术创新、产业革命的重要作用

- 高性能计算：是科学研究的三大手段之一
- 大数据处理：正成为科学研究的第四范式





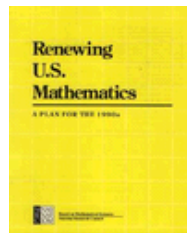
## ○ 计算技术的历史

### — 计算物理

- 美国1963年“计算物理方法”丛书，到1977年共出版17卷
- 统计物理、量子力学、流体力学、核粒子运动、核物理、天体物理、固体物理、等离子体物理、原子与分子散射、地表波、地球物理、射电天文、受控热核反应和大气环流等方面的物理问题，在计算机上进行计算所需要的计算方法以及反映当时水平的研究成果

### — 科学计算

- 1983年，由美国国防部、能源部、国家科学基金会及国家航天局等主持，拉克斯（P.Lax）为首的不同学科的专家委员会，强调“科学计算是关系到国家安全、经济发展和科技进步的关键性环节，是事关国家命脉的大事”。1990



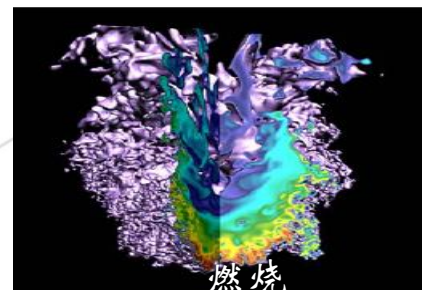
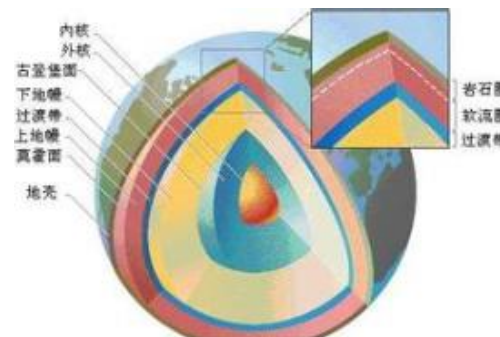
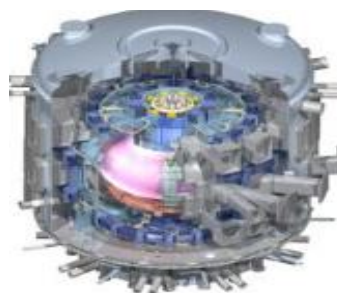
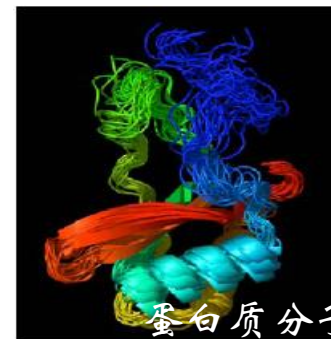
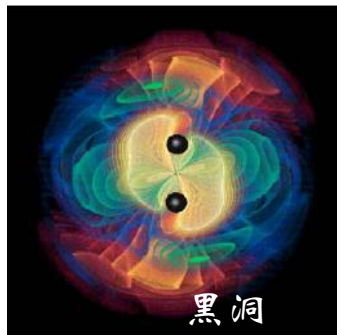
### — 战略计算

- 1995，加速战略创新计划（ASCI）
- 2002，高产出率计算计划（HPCS）

## 高性能计算可以帮助人们解决一系列重要问题

- Too big
- Too small
- Too fast
- Too slow
- Too expensive
- Too dangerous

For experiments





**Cray-1**  
1976 160Mflops



**Intel ASCI Red**  
1997年, 1万亿次



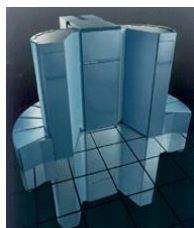
**IBM BlueGene/L**  
2004年, 70Tflops  
2007年, 596Tflops



**中国 NUDT TH-1A**  
2010. 4.7Pflops



**美国 Cray Titan**  
2012. 17.6Pflops



**Cray-YMP**  
1988年



**DT 天河2号**

2013. 33.86Pflops

**40年系统性能增长5.8亿倍**



**Cray T3D**  
1993, 19Gflops



**日本地球模拟器**  
2002年, 40Tflops



**Cray Jaguar**  
2009, 1.76Pflops

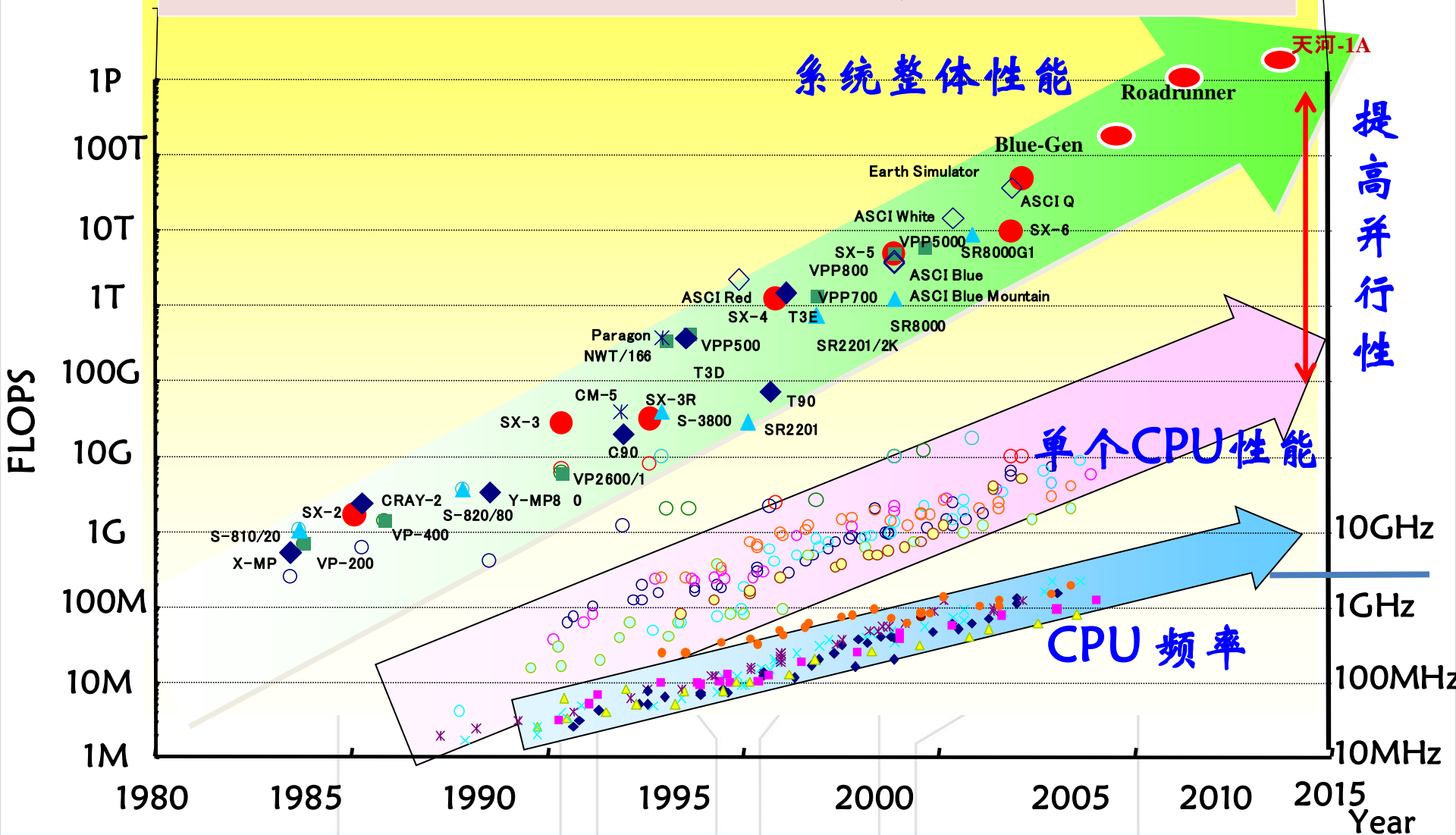


**美国 IBM Sequoia**  
2012. 16.32Pflops



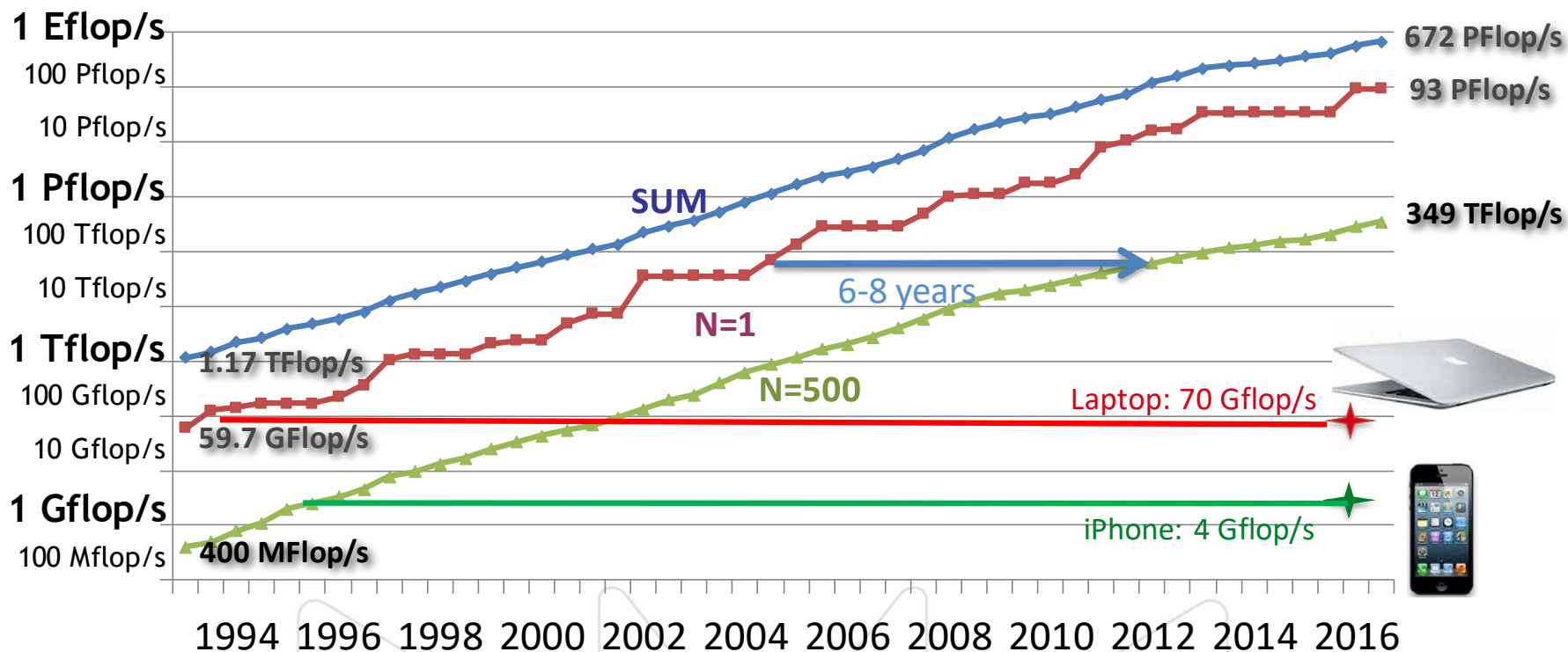
**中国 太湖之光**  
2016. 93Pflops

# “并行”是通往高性能计算的唯一途径

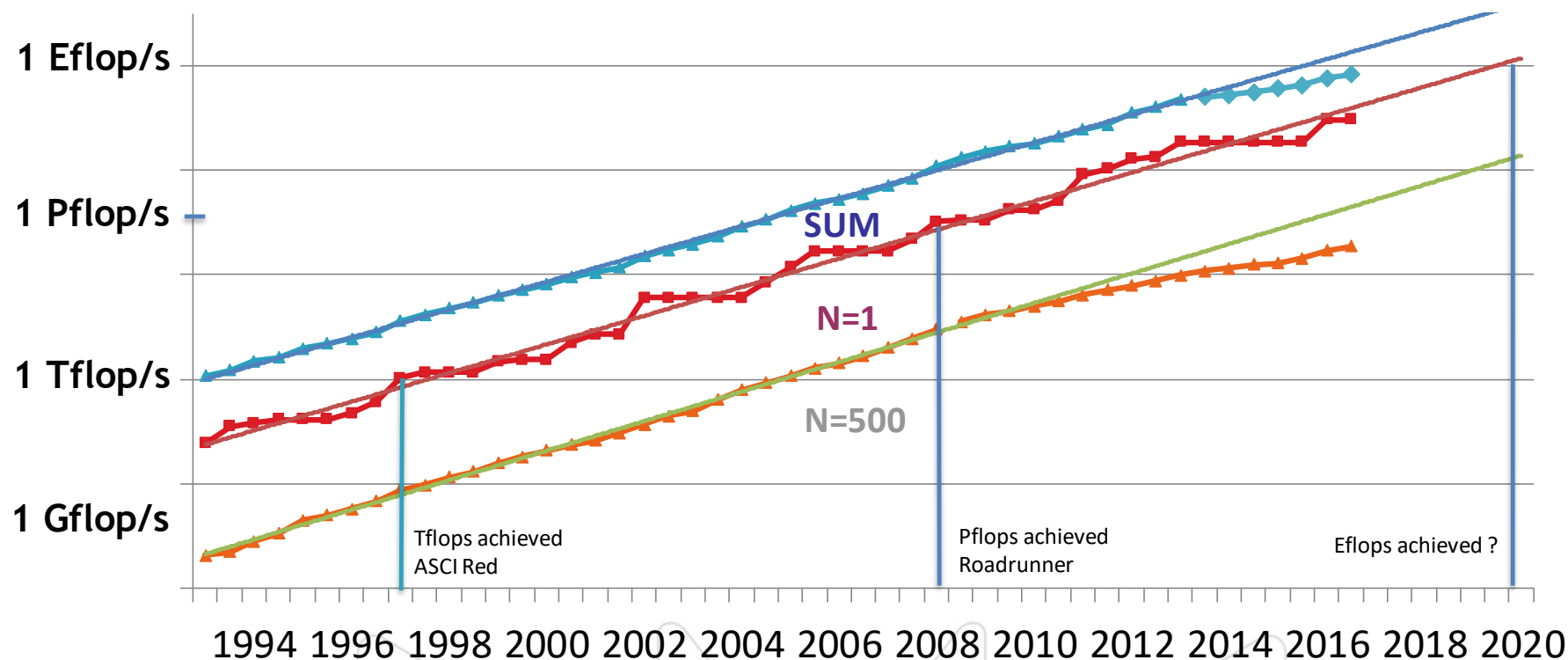




## Top500性能增长的趋势



## Top500性能增长的趋势



## 超级计算机的特征

某一时代性能最高的系统，服务于国家战略目标

○ 运算速度超级快

○ 存储容量超级大

○ 占地面积超级大

○ 能耗超级高

○ 造价超级贵

Meta → Gega → Tera → Peta → **Exa**  
 $10^6 \rightarrow 10^9 \rightarrow 10^{12} \rightarrow 10^{15} \rightarrow 10^{18}$

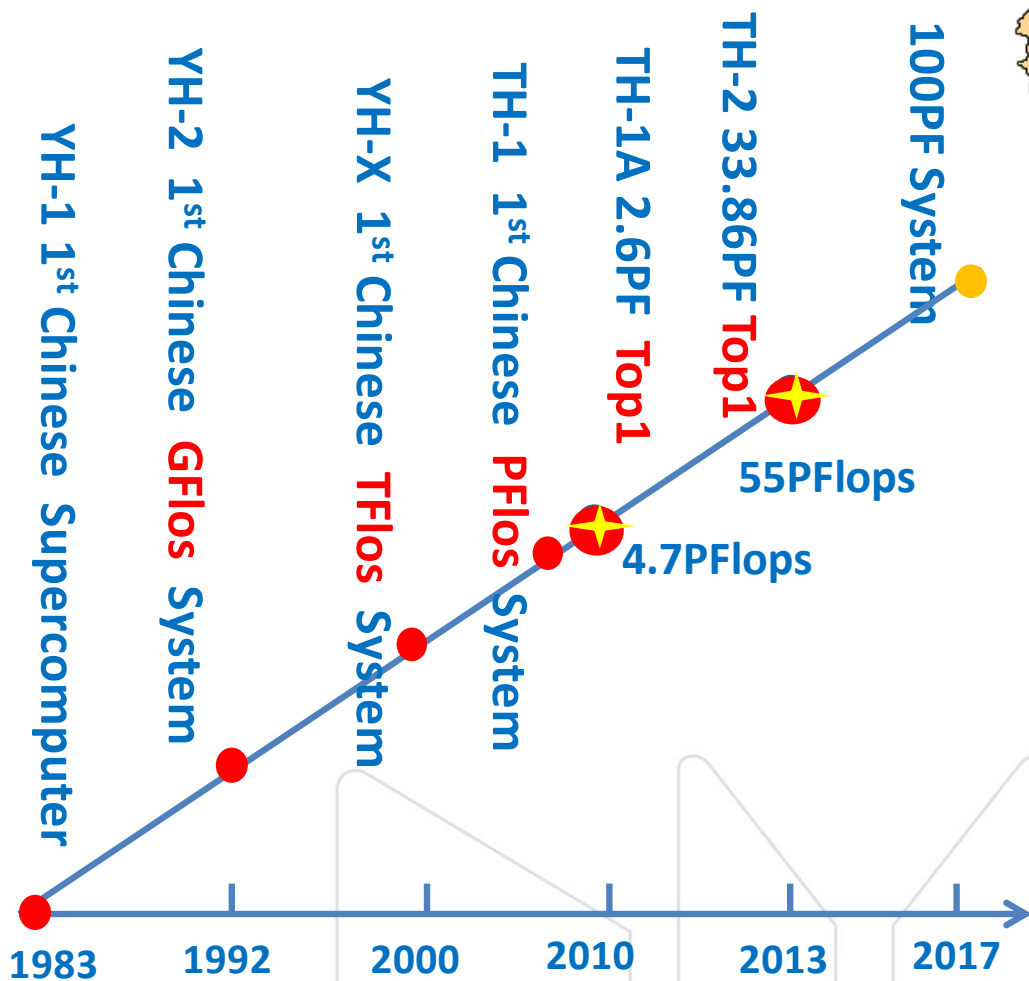
100-1000 m<sup>2</sup>

~10 MW

~Billion RMB

其关键技术迅速辐射到其它应用领域，推动  
国民经济建设、科学技术进步与人类社会发展

## 银河天河系列超级计算机 国防科大 长沙



NSCC-Tianjin, 2010



NSCC-Guangzhou, 2013



NSCC-Changsha, 2012



CCC-Lvliang, 2014



## ● 天河二号超级计算机系统

- 国家863计划和核高基重大专项的标志成果，2013年6月研制成功，位列世界超算Top500排行榜第一位

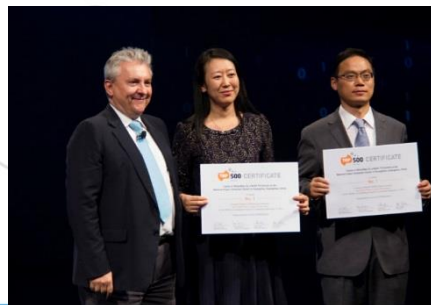


## ● 世界超级计算机Top500排行榜 “六连冠”

2013.6~2016.5

## ● 世界共轭梯度HPCG排行榜 “五连冠”

2014.6~至今



## 性能的不同层次

### ● 峰值性能 ( Peak performance )

### ● LINPACK 性能

– Avg. 80%

40+ years

### ● Gordon Bell

– ~30%

### ● 平均持续性能

– <5%

### ● HPCG

– 1% ~

No. HPL	Machine	Rmax/Rpeak (Pflops)	Linpack Ratio	HPCG (Tflops)	HPCG Ratio	Bytes/Flops	No. HPCG
1	神威太湖之光	93/125	74%	370	0.3%	0.014	3
2	天河2号	33.86/54.9	62%	580	1.1%	0.04	1
3	Titan (CrayXK7)	17.5/27.1	64%	320	1.2%	0.041	4
4	Sequoia (BG-Q)	17.1/20.1	85%	330	1.6%	0.08	5
5	K Computer	10.5/11.28	93%	550	4.9%	0.14	2

科学计算/大数据处理软件栈

HPC 应用

领域编程框架

并行程序运行支持系统

编译系统 并行文件系统 通信系统

操作系统

编程框架

结构网格

非结构网格

三维组合几何

有限元计算

由领域专家主导研发

由计算机专家主导研发

运行时系统

数据管理结构

执行模型

调度模型

容错功能

性能分析

性能	54.9PFlops / 33.86PFlops
系统	16000结点, 1.4PB内存
机柜	125+8+13+24=170 (720m <sup>2</sup> )
能耗	17.8 MW (1902MFlops/W)
制冷	密闭水风冷

## 天河二号系统



TH-2 系统

机柜  
(8 x 机柜)

TH-Net

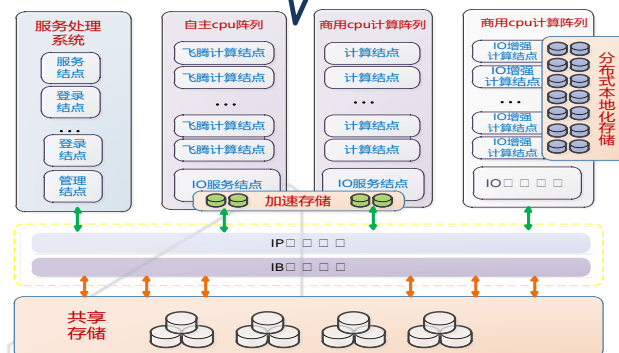
机柜  
(16 x cn)

计算主板

ION

CPM

APM

Phi  
#48000IVB  
#32000FT-1500  
#4096大规模混合层次式并行存储系统  
12.4PB

## 技术挑战 (PSPRD)

- 高性能 (Performance)
- 可扩展 (Scalability)
- 低能耗 (Power consumption)
- 可靠性 (Reliability)
- 大数据 (big Data)

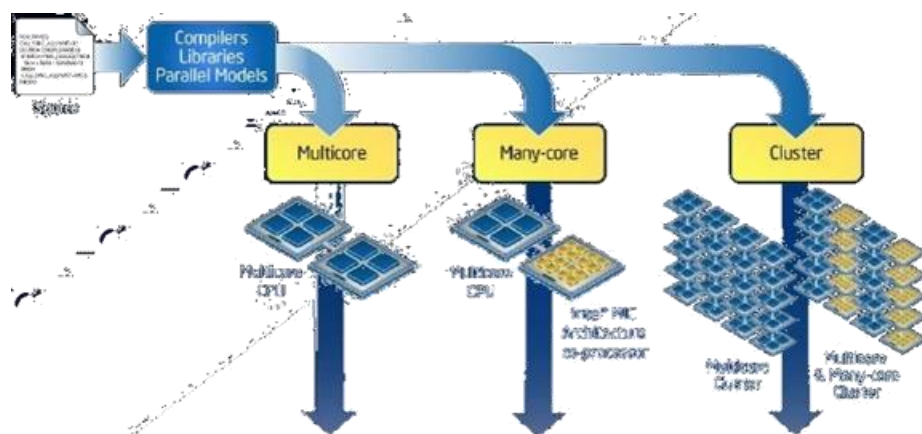
## 超算系统设计关键技术

- 体系结构
- 互连网络
- 存储架构
- 软件栈

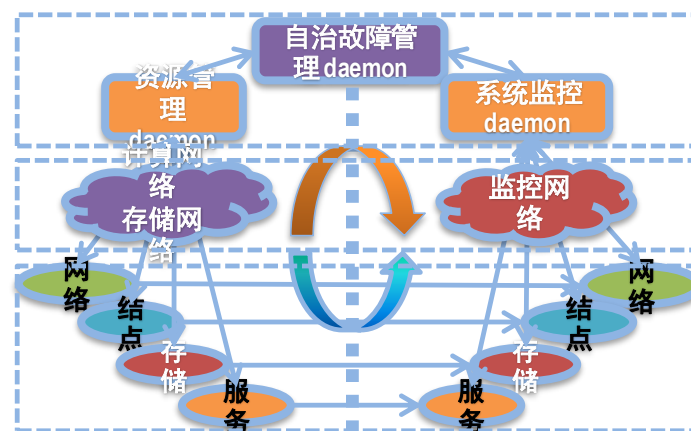
## “微异构”提升了应用的兼容性、适用性和易用性

- CPU端的程序只需重新编译，便可在MIC上运行
- MIC上可以运行操作系统，CPU、MIC均可以启动主函数，CPU与MIC也可以同时启动主函数
- 丰富的编程模式

Offload模式，Symmetric模式，MIC Native模式



## 天河二号系统

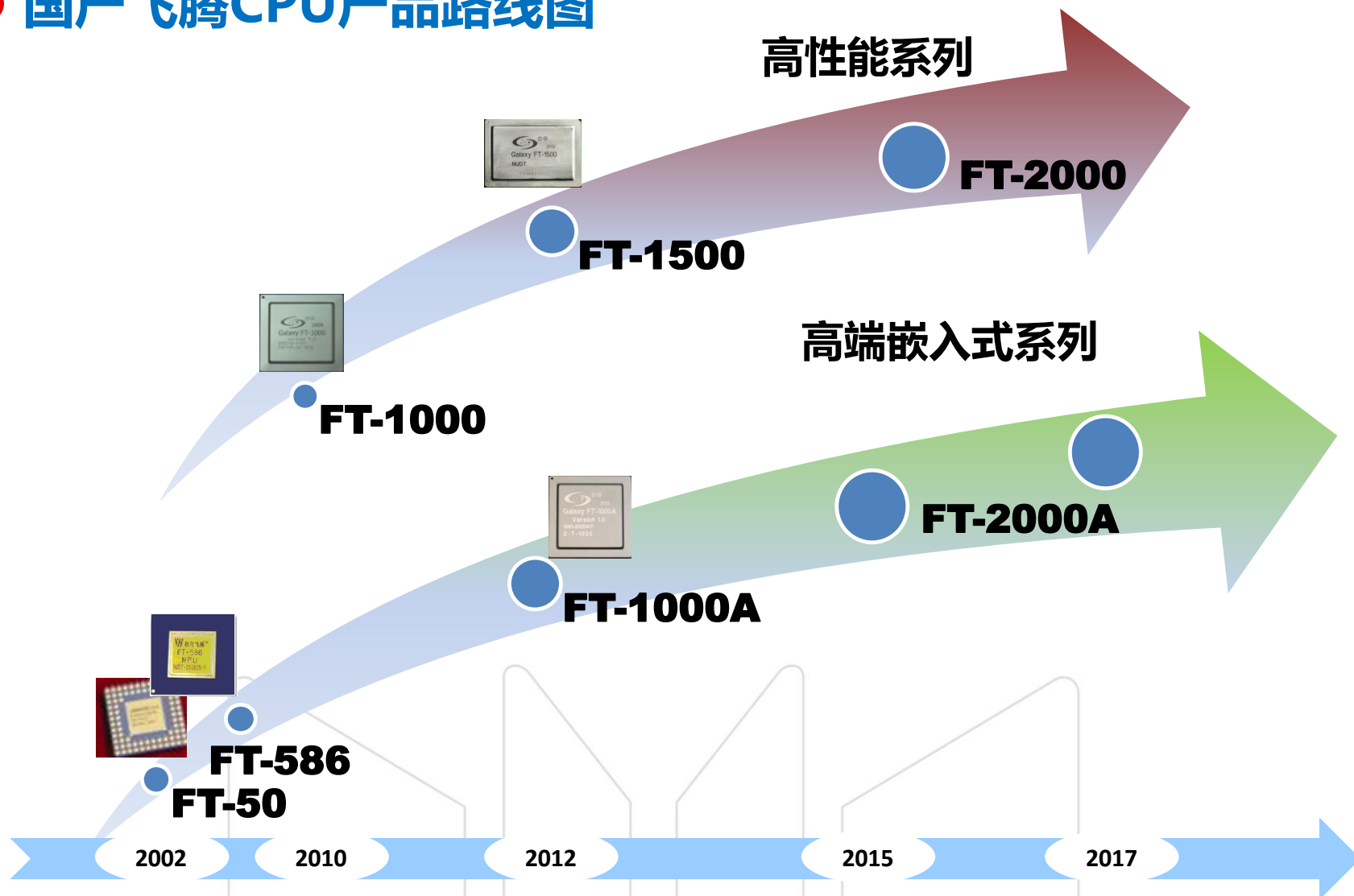


## 支持100PF系统的容错设计与故障管理

- 多层次容错设计，一体化故障管理
- 提高了系统的可用性



## ● 国产飞腾CPU产品路线图



## 飞腾CPU产品系列

### – 高性能CPU系列：

- FT1000、FT1500、FT2000
- 面向高性能计算、服务器应用需求，主要用于高性能计算机、数据中心、云计算和互联网服务器



### – 高端嵌入式与桌面CPU系列：

- FT50、FT586、FT1000A、FT2000A
- 面向各类信息化武器装备需求，主要用于桌面、终端、服务器和武器装备信息系统



## 麒麟操作系统

### 面向天河高性能计算机研发的专用操作系统

- 面向高性能计算需求
- 面向异构并行体系结构
- 面向大规模系统管理需求
- 面向差异化资源使用需求
- 面向高安全性使用需求

### 面向多核多线处理器的优化

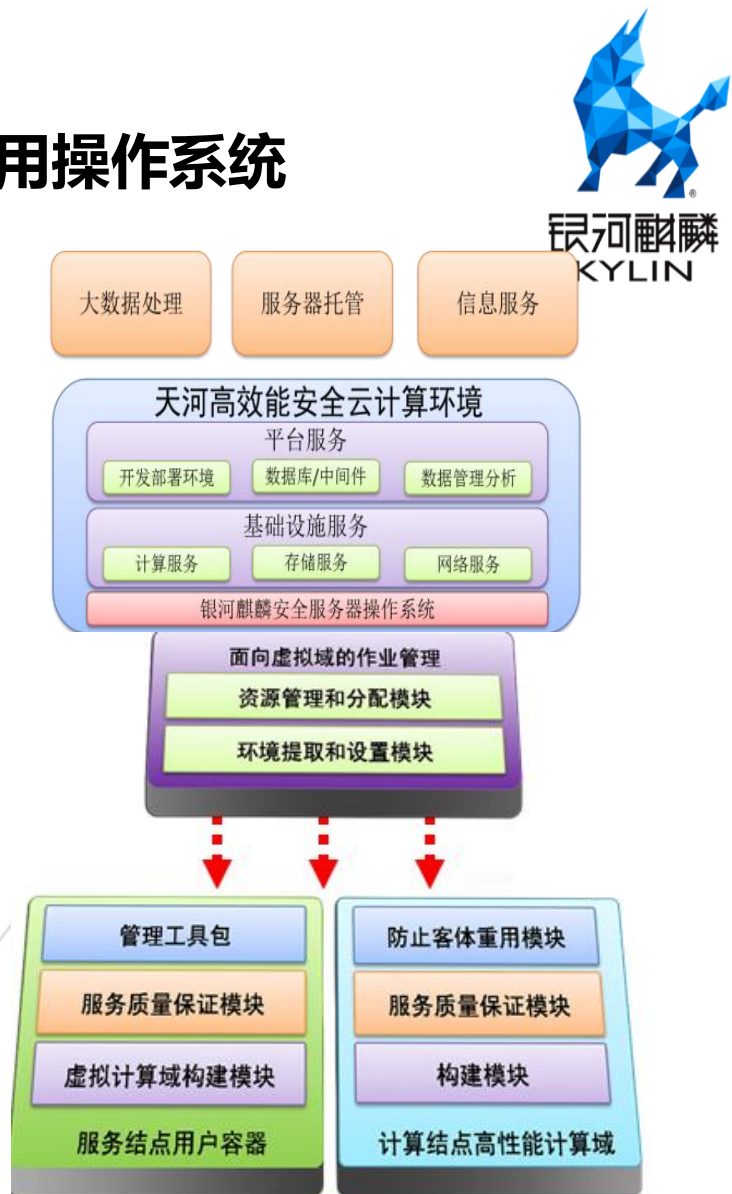
### 低开销的轻权运行环境

### 支持低功耗管理

### 支持高速通信接口

- 低开销用户级通信技术
- 主动消息形式的内核通信
- 零拷贝流水化TCP/IP报文传输

### 高性能，稳定可靠

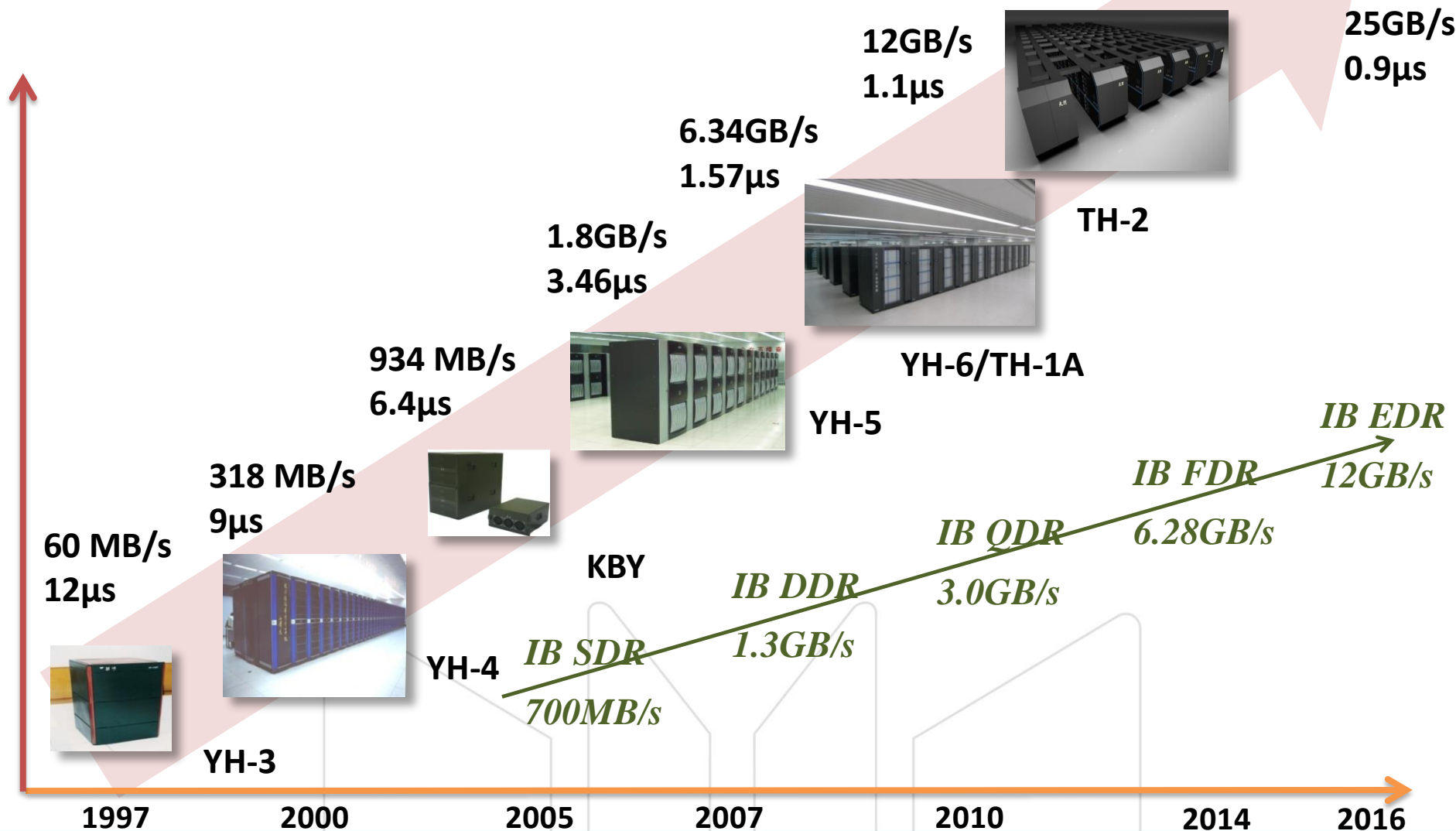




## ● 系统互连网络对高性能计算机的实用性能、可靠性、可扩展性具有至关重要的影响

- 多核微处理器、异构融合体系结构的发展，计算性能的增长速度已远远超越了互连通信性能的增长速度，互连通信的性能瓶颈更加突出；
- 国外HPC厂商不遗余力研发定制互连通信技术
  - Cary : SeaStar、SeaStar2+、YARC、Gemini
  - IBM : HiPS、SP Switch、SP Switch2、HPS
- 近15年TOP500排名第一的10个系统中，9个系统采用了定制互连通信技术

## 国内唯一坚持定制互连通信网络研发的单位

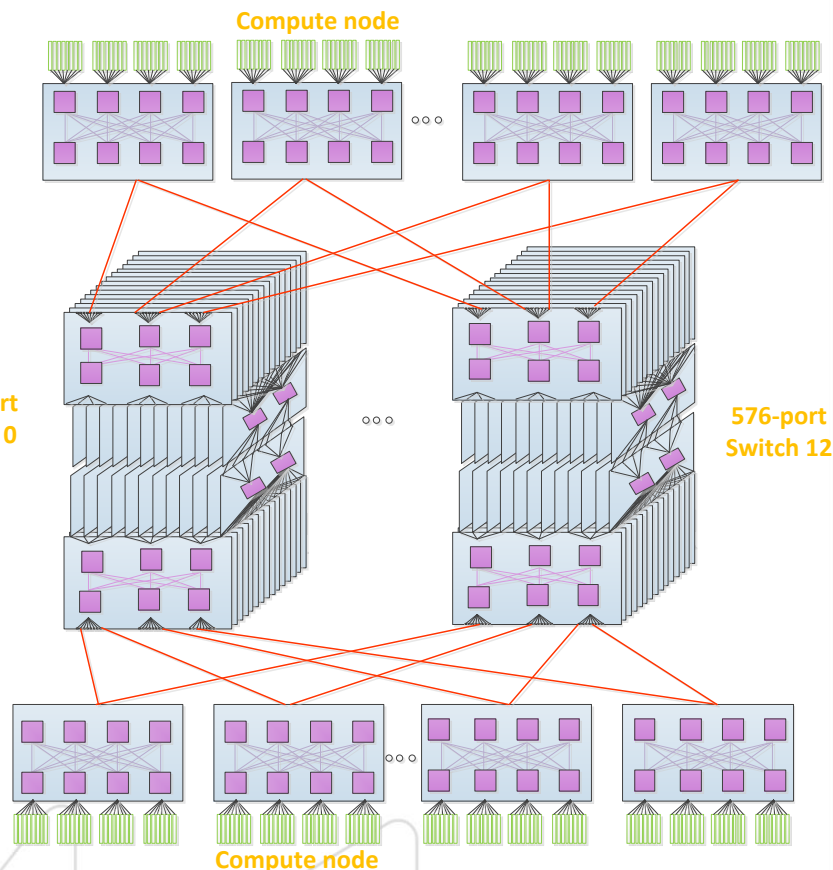


## 互连通信系统是决定系统规模能否平衡扩展的关键

- 计算结点性能越高对点对点通信的要求越高
- 系统规模越大，聚合通信的性能越成为制约因素

## 天河二号自主定制互连网络可高效支持100PF系统

- 自主网络协议
- 胖树拓扑结构
- 光电混合互连技术
- 自适应互联架构



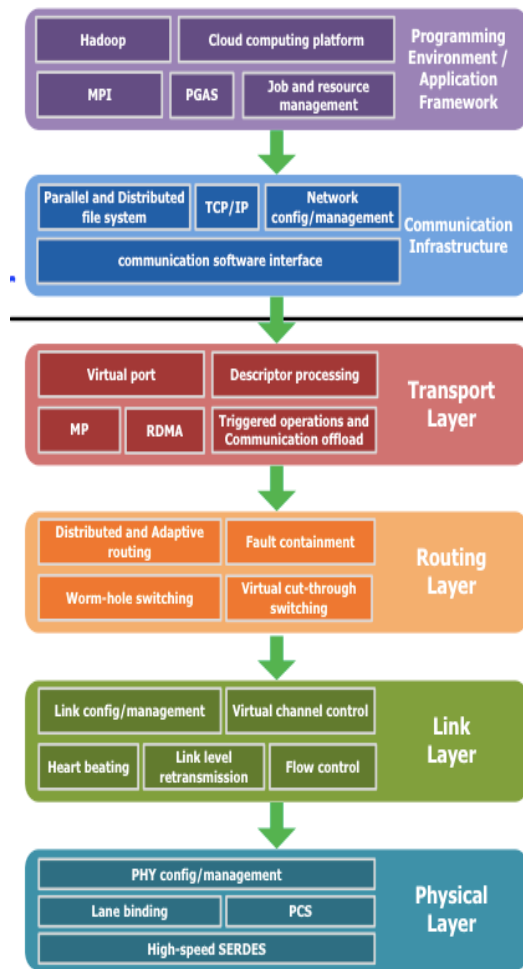
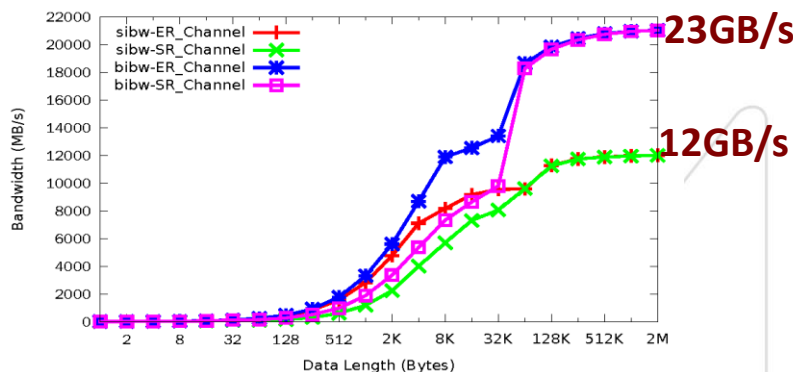
## TH-Express 2+ Interconnect

### ● Interconnect fabric

- Optic and electronic
- Topology: multi-dimension tree
- Scalability: extensible to 100 PFlops
- In/Out-Band network management

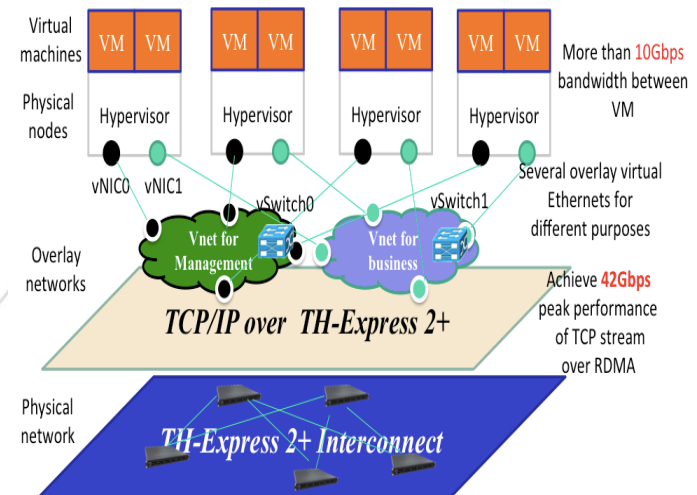
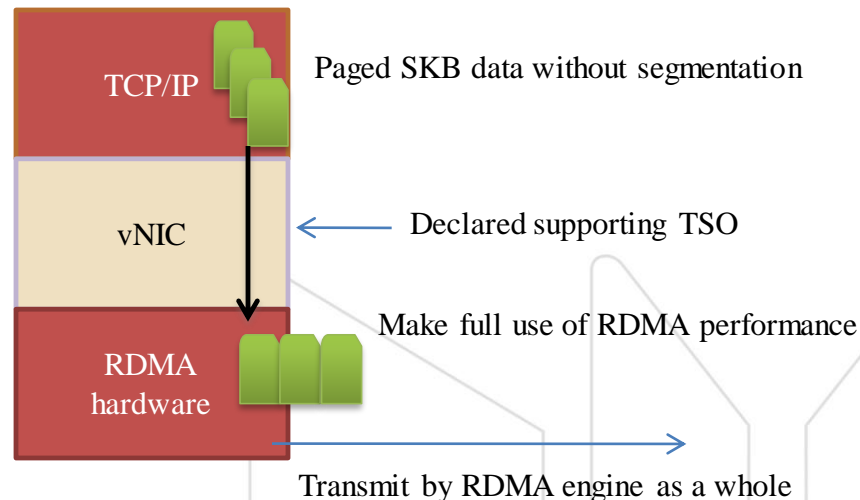
### ● User-level communication system

- Low latency user interface
- Non-blocking transfer and RDMA op
- Triggered communication offload operations to bypass system noise



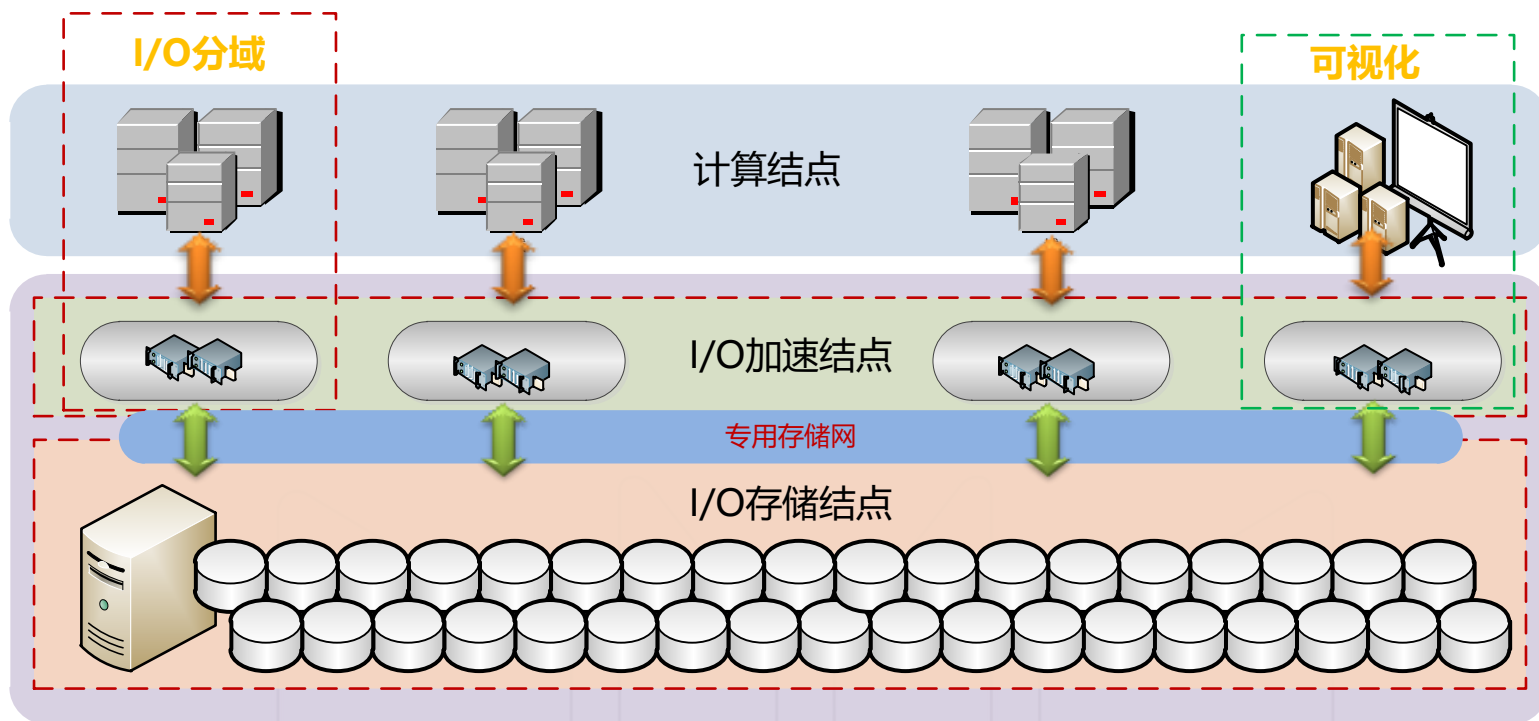
## TCP/IP over TH-Express 2+

- Fulfill the requirement of network application services and big data framework
- High bandwidth socket-based communication
  - Pipeline network packet transfer using RDMA
  - Interrupt consolidation and poll based packet processing
  - ARP broadcast elimination
  - Transmitting TSO (TCP segment offload) data as a whole



## ● 层次式加速存储架构

- 高速互连网络和存储网桥接，优化I/O路径
- 基于固态存储阵列的I/O加速层，分域并发加速
- 基于磁盘的全局存储层，大规模高可靠数据共享



## TH-2 Hybrid Storage

- Local Disk
- PCI-E SSD
- Disk Array

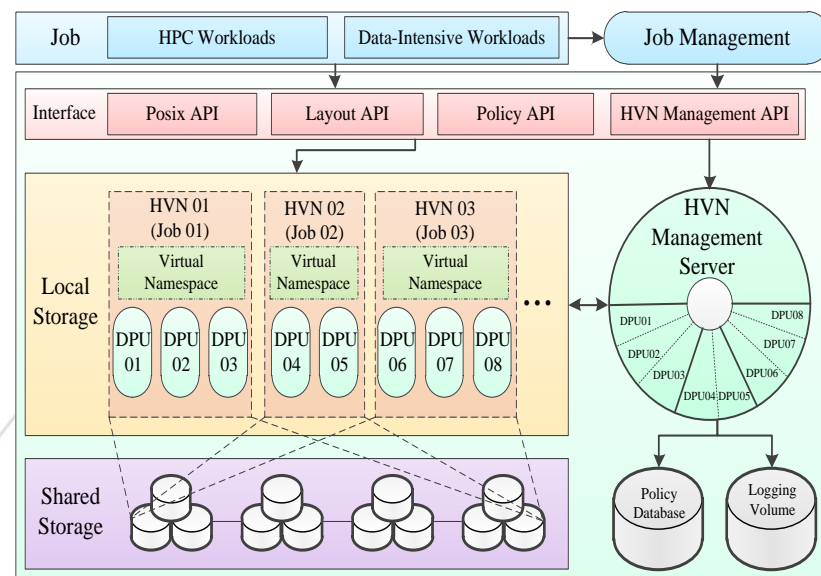
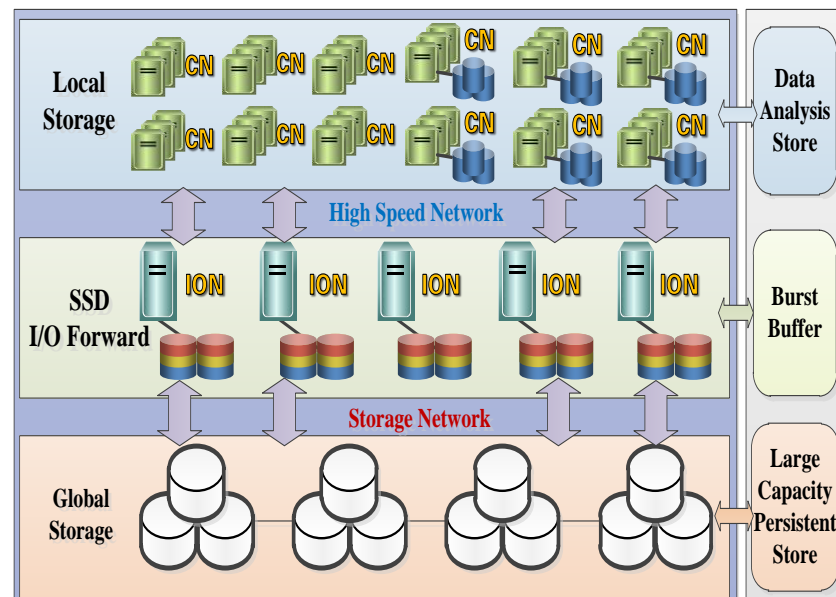
## Storage Hierarchy

- Burst BW: ~ 1TB/s
- Sustained BW: ~ 100GB/s

## Supporting in-situ data analysis

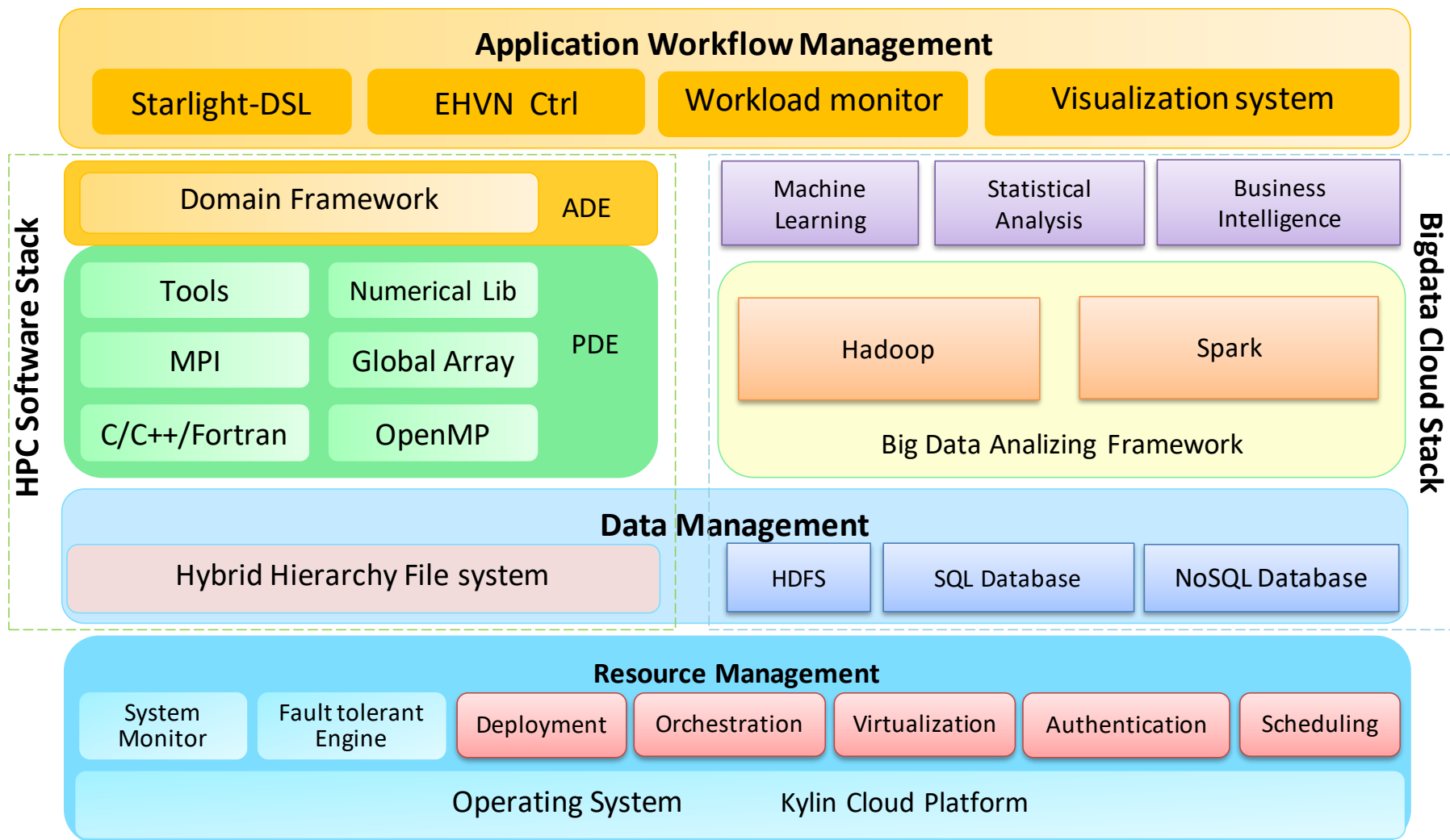
## H2FS: Hybrid Hierarchy File System

- DPU, A fundamental unit for data processing, coupling compute node with its local storage
- HVN, Hybrid, unified and Isolated dynamic namespace
- Layered and enriched metadata, I/O hints as high level metadata





## ● 天河星光 Starlight Software Stack—HPC & Bigdata convergence Stack





目标与定位：具有中国特色的世界一流超算中心

大科学

大工程

新产业

科学研究平台

技术创新平台

交流合作平台

学科建设平台

人才培养平台

○ 新型能源

○ 材料化工

○ 气候环境

○ 航空航天

○ 天文物理

○ 工程设计

○ 石油勘探

○ 生命科学

.....

○ 智慧城市

○ 金融分析

○ 互/物联网

○ 智能交通

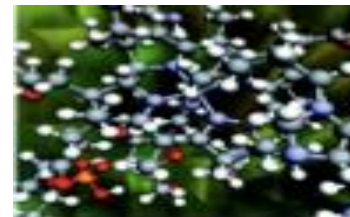
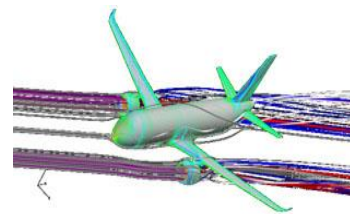
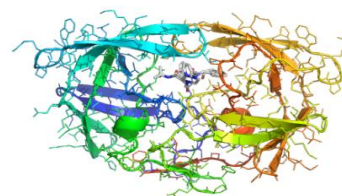
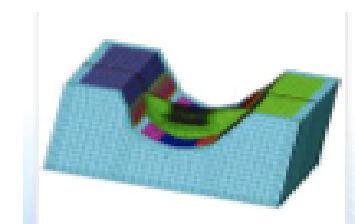
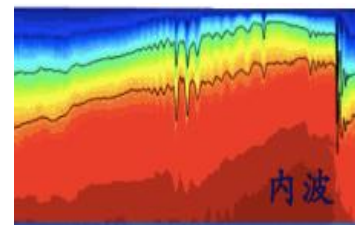
○ 动漫渲染

○ 人工智能

○ 测绘分析

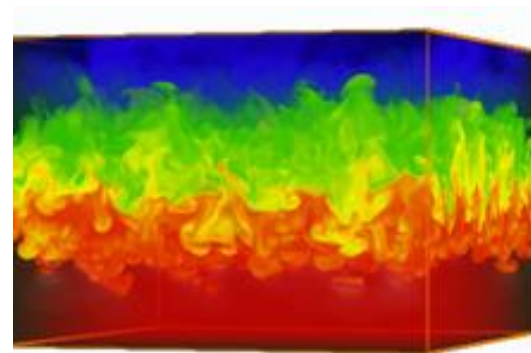
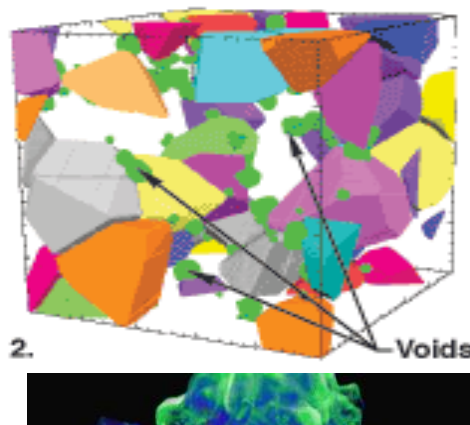
○ 精准医疗

.....



## 复杂物理问题

- 数值模拟核反应全过程
  - 爆轰流体力学、辐射流体力学和辐射与中子/ $\gamma$ 光子输运三类方程



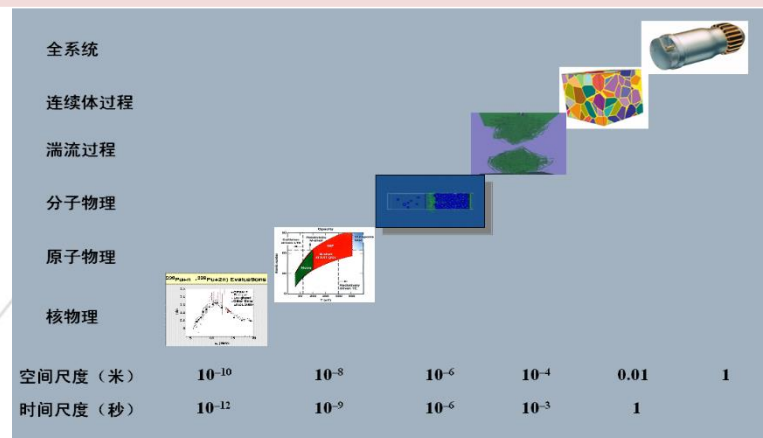
通过增加模拟的分辨率来提高精度

全局二维和局部三维计算需要每秒百万亿次至万万亿次计算能力

高逼真度、三维计算需要每秒百万万亿 ( $10^{18}$ ) 次计算能力

多物理过程紧耦合的非线性、非定常的大型系统

- 多物理过程紧耦合的非线性、非定常的大型系统

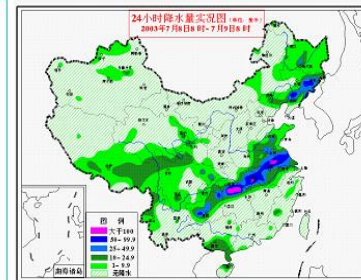
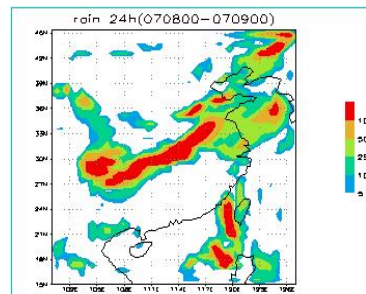
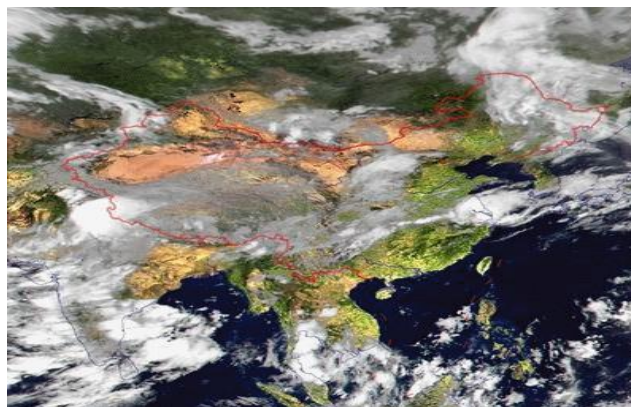


## 数值天气预报



1904年比杰克尼斯(Bjerknes)  
首次提出数值预报理论

地球表面发生的千变万化天气现象，  
都遵循基本的物理定理，可以通过求解  
这些方程，推算出未来时刻天气变化。  
这就是数值天气预报思想



$$\frac{d\vec{V}}{dt} = -\frac{1}{\rho} \cdot \nabla P - 2\vec{\Omega} \times \vec{V} - \vec{g} + \vec{F}_V$$

牛顿运动定律

$$C_p \frac{dT}{dt} - \alpha \frac{dP}{dt} = Q_T + F_T$$

热力学定理

$$\frac{1}{\rho} \frac{d\rho}{dt} = -\nabla \cdot \vec{V}$$

质量守恒定律

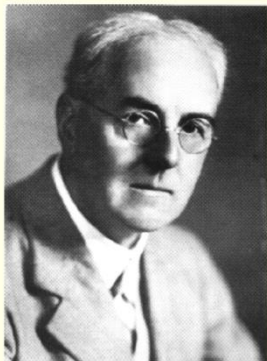
$$P = R\rho T$$

大气状态变化关系



## 数值天气预报

Lewis Fry Richardson, 1881–1953.



Bjerknes believed that for the solution of the forecasting problem **graphical or mixed graphical and numerical methods** were appropriate.

However, Richardson was bolder — or more foolhardy — than Bjerknes.

He attempted a **bulldozer approach**, calculating changes from the full partial differential equations.

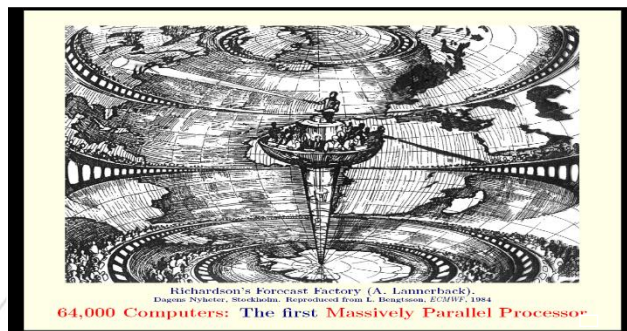
1916~1918年期间（第一次世界大战），理查逊博士应征到法国前线的救护车队当司机，他利用出车的空余时间，尝试实践Bjerknes的数值天气预报理论思想，用手工的计算工具，直接解大气方程，推算未来伦敦6小时的气压变化。

这是世界上首次“计算出未来的天气”。



理查逊工具：计算尺

第一次世界大战后，理查逊回到了伦敦，将他的“失败”工作进行总结，出版了书，供后人借鉴。同时，也设想了一种方法若业务应用的话，需要建设6万四千人的“计算工厂”，才能做全球预报



理查逊设计的“预报工厂”：把全球分为64000个点，3h做出6h天气预报

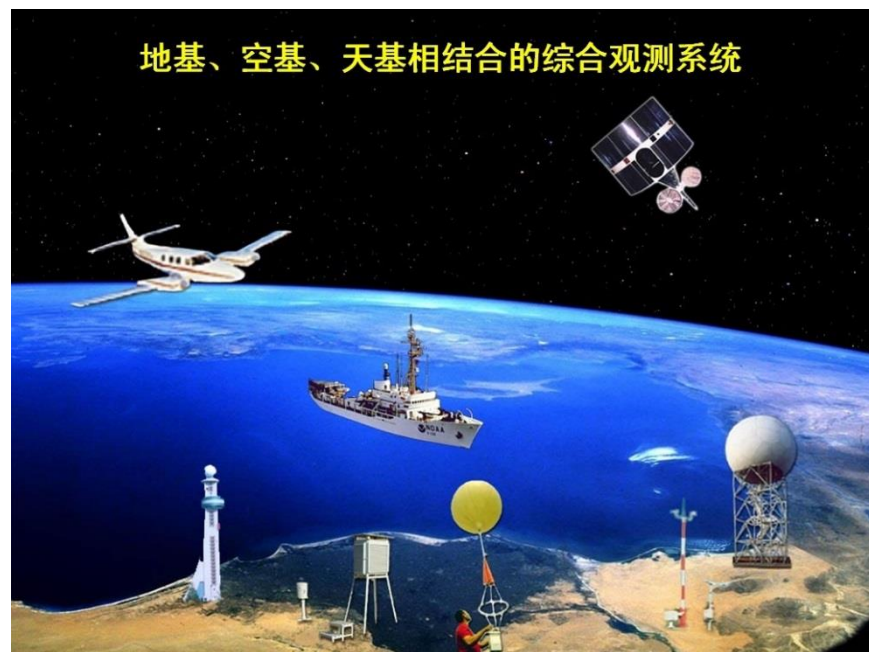
## 数值天气预报

### 预报模式的误差

一模式由不可分辨的大气运动过程造成，主要由模式有限的分辨率产生，由离散网格点代替连续时空的计算方法都会带来一定误差

- 预报模式应能分辨或准确模拟大气运动的各种过程，大气运动过程的尺度从 $10^{-6}$ 米到 $10^7$ 米，跨越13个量级

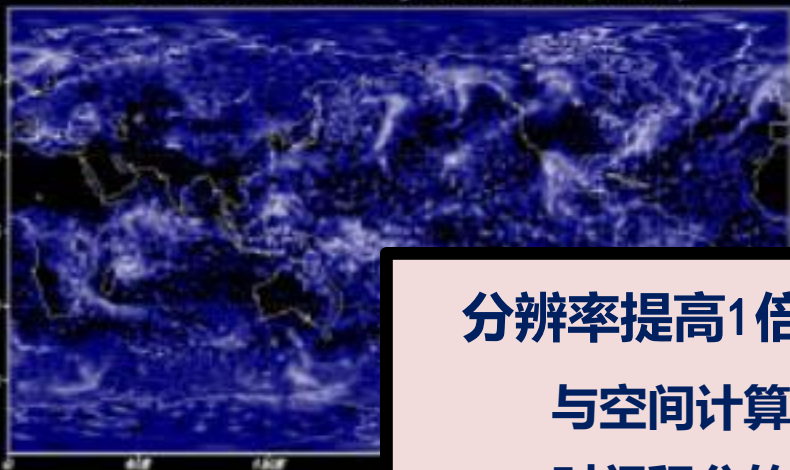
气象观测资料量巨大  
资料同化计算复杂





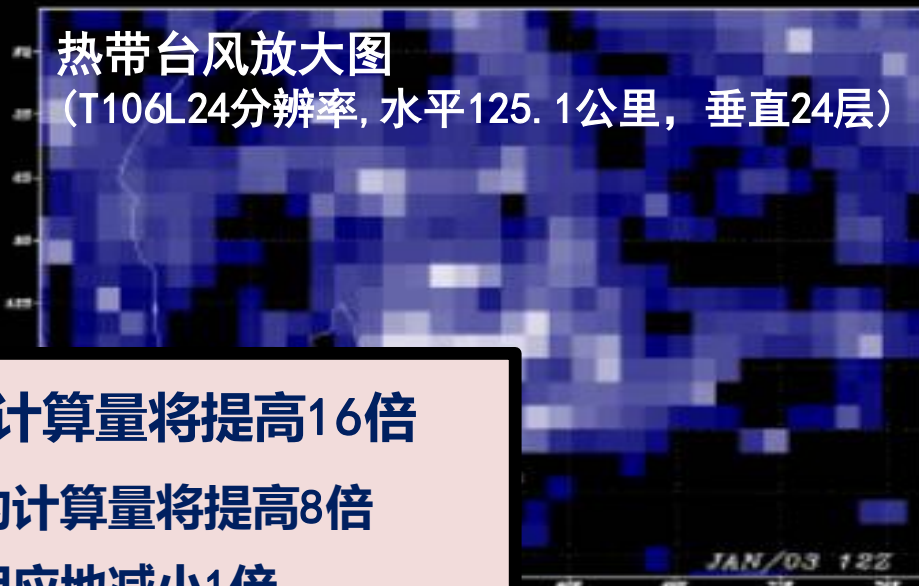
全球场

AFES T106L24 Precipitation(mm/hour)



热带台风放大图

(T106L24分辨率, 水平125.1公里, 垂直24层)



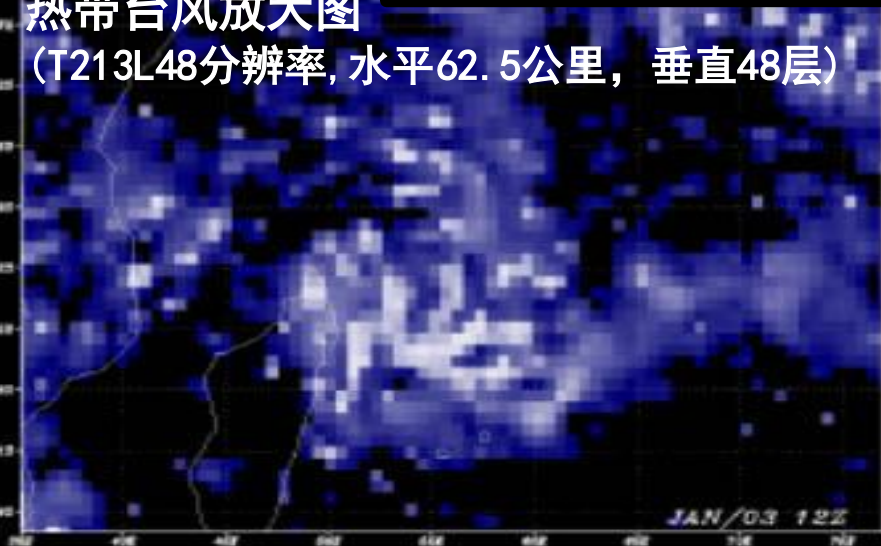
分辨率提高1倍, 总计算量将提高16倍

与空间计算相关的计算量将提高8倍

时间积分的步长相应地减小1倍

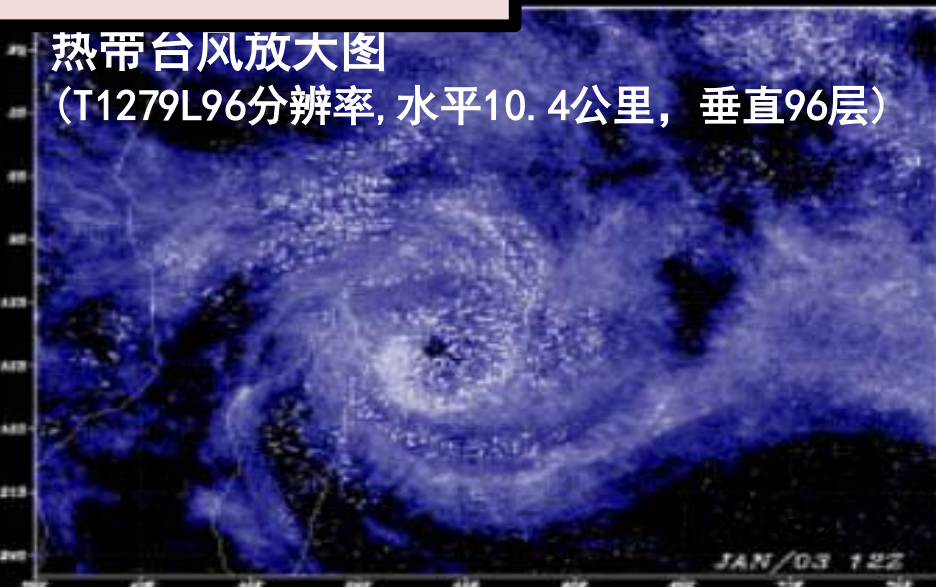
热带台风放大图

(T213L48分辨率, 水平62.5公里, 垂直48层)

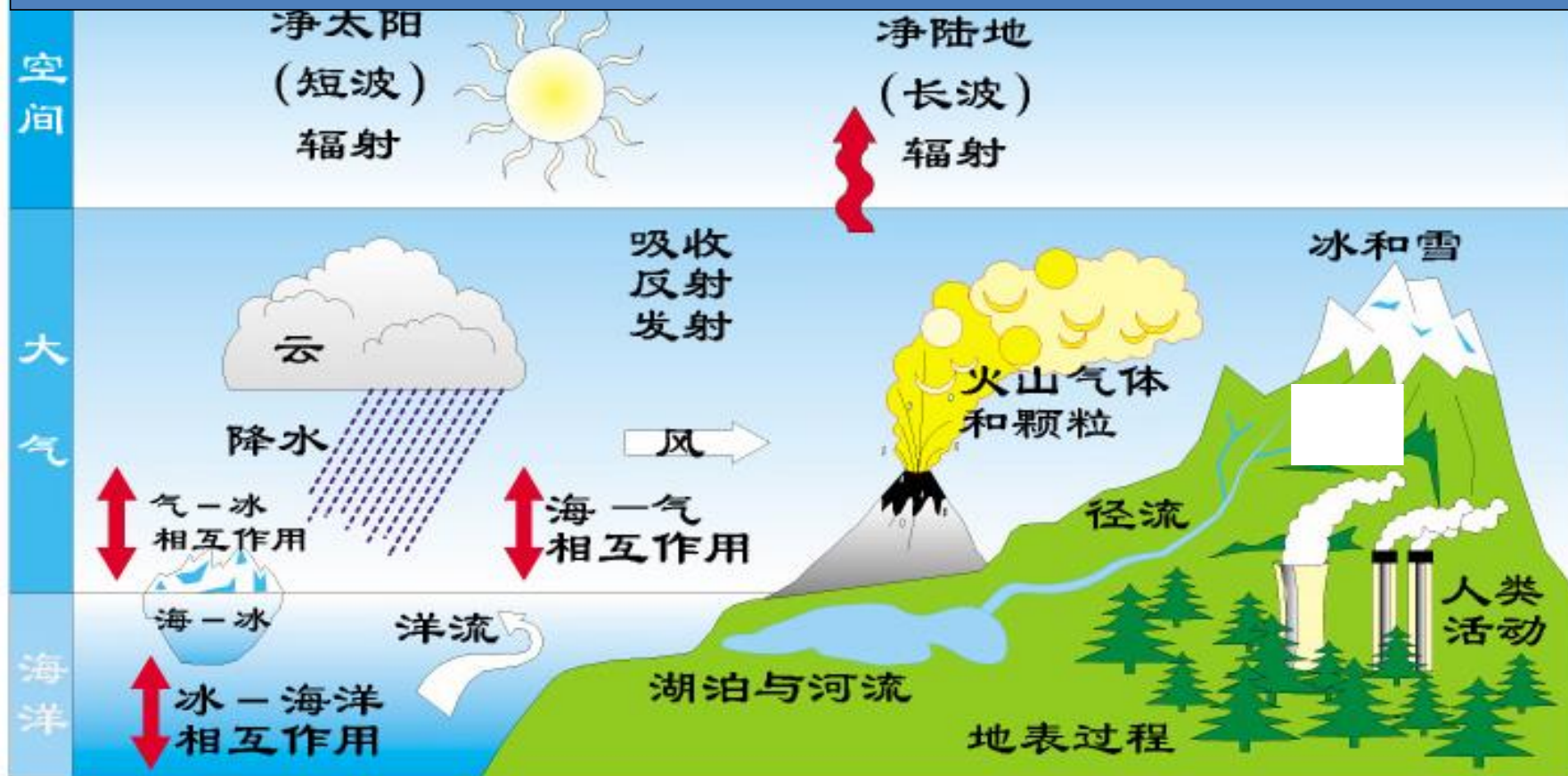


热带台风放大图

(T1279L96分辨率, 水平10.4公里, 垂直96层)



## 空间、大气和海洋无缝隙天气保障



建立天气、气候、地球系统一体化的全球/区域预报系统



## 数值天气预报对计算能力的需求

大气运动各种过程的尺度从 $10^{-6}\text{m}$ 到 $10^7\text{m}$ ，**跨越13个量级**

模拟尺度10,000km至10km天气过程，需处理跨越3个量级的过程，就全球云尺度极端天气预报来说，模式水平分辨率要达到1km

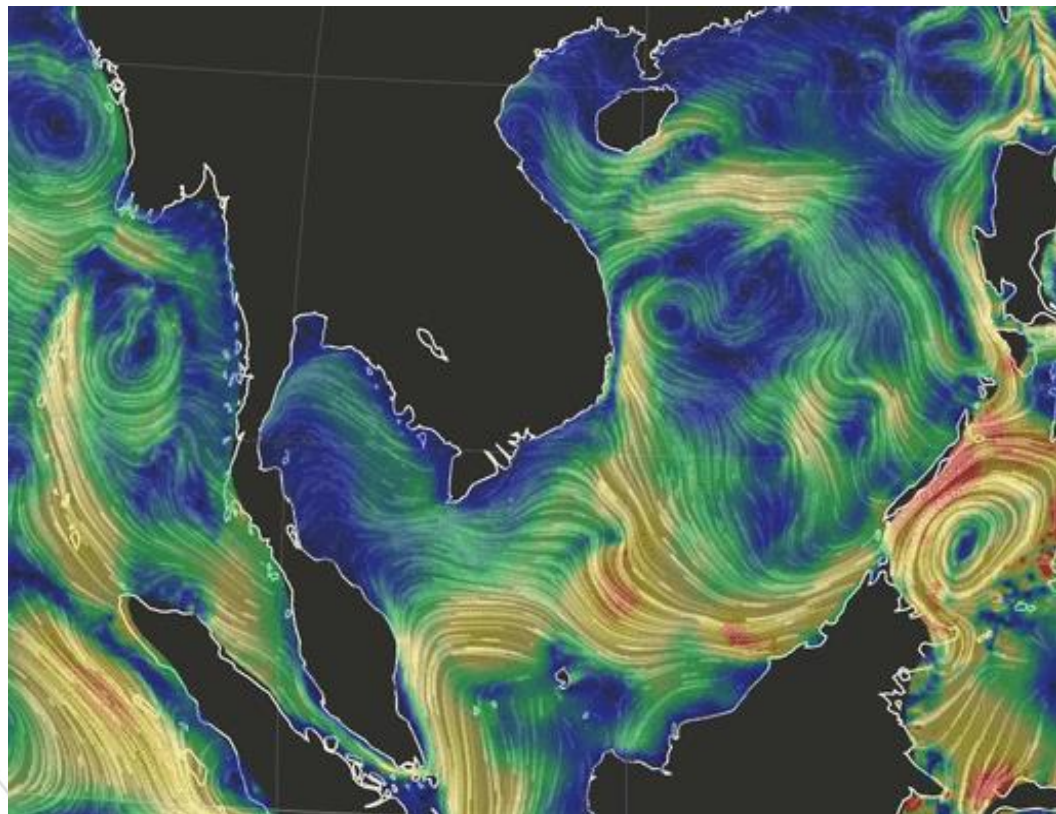
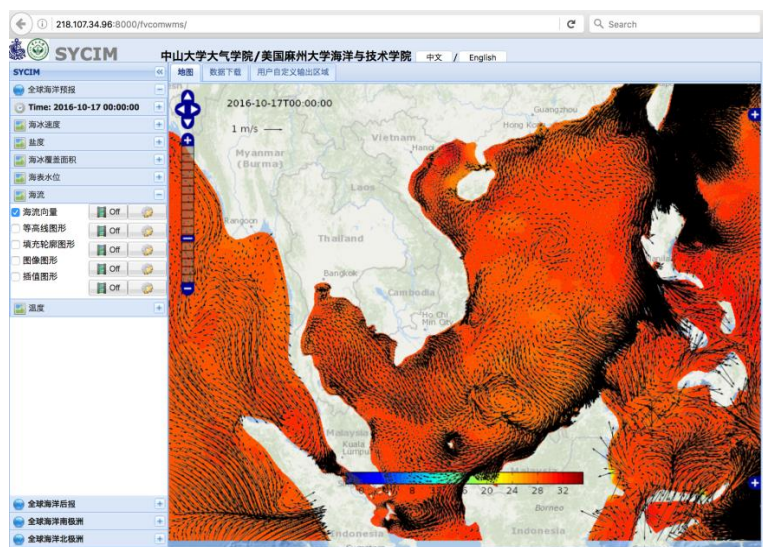
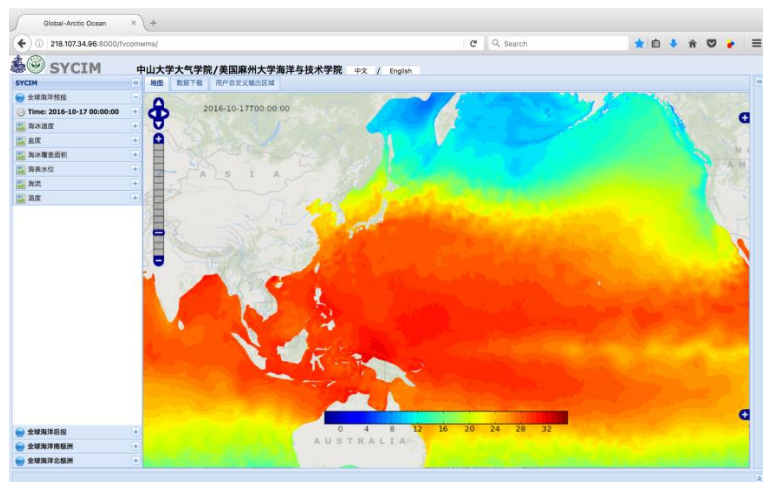
精确模拟所需网格数 $10^{13} \times 10^{13} \times 10^{13} = 10^{39}$ ，再考虑时间积分，总计算量是一个天文数字

**数值天气预报对超级计算机有无止境的能力需求**

## 天河2号应用：大气海洋环境

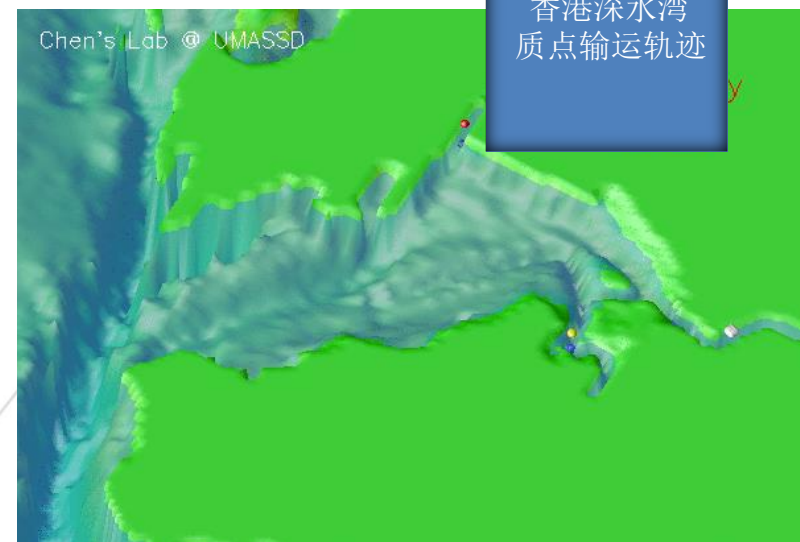
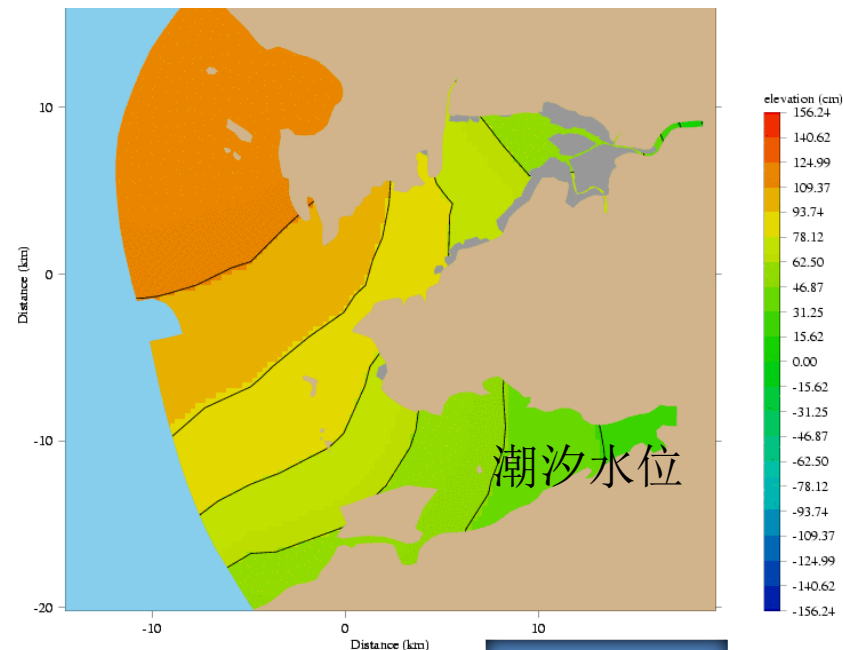
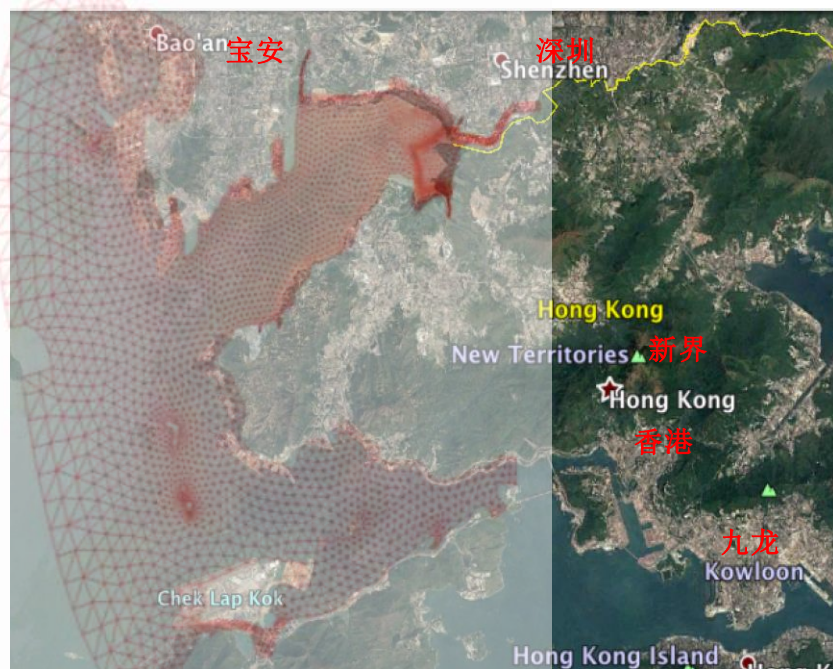
## 全球FVCOM预报模型

## 南海秋季表层环流



## 天河2号应用：大气海洋环境

## 香港深水湾潮汐模型





## 天河2号应用：大气海洋环境

### 新一代多尺度海洋模型

— 海洋动力学，台风预报，污染物赤潮预警

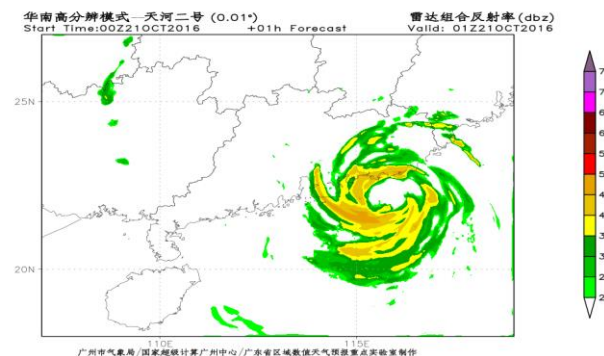
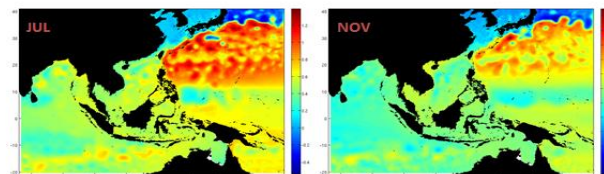
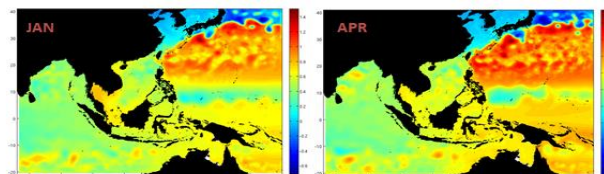
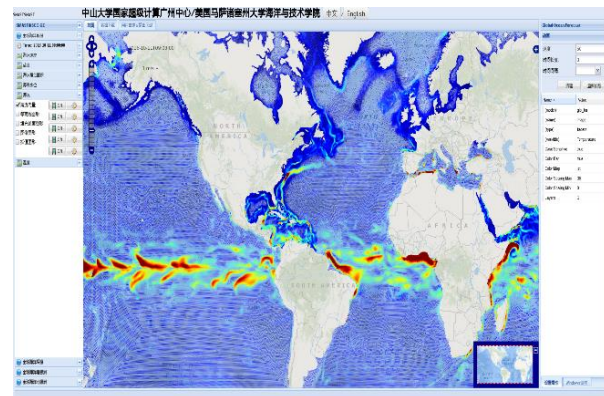
— 25分钟完成全球未来10天海洋模拟，首次耦合南北极海冰及海岸因素，全面揭示全球气候变化带来的环境影响

### 区域海洋模式

— 50年高分辨率的暖池区海洋环流模式，印太暖池-南海地区中深层环流结构模拟，消除2000米以下深度的模拟“盲区”

### 广州市气象业务预报系统

— 率先实现气象预报在超算中心的业务



广州中气基局/国家超级计算广州中心/广东省区域数值天气预报重点实验室制作

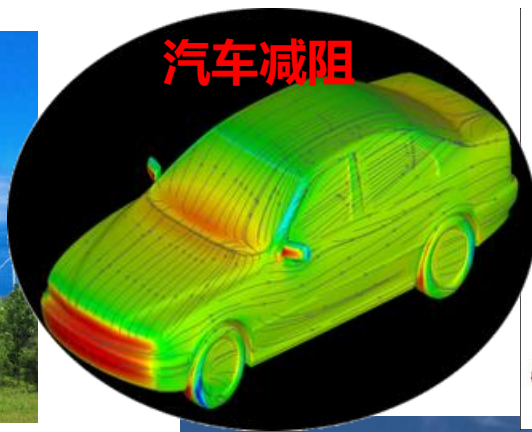
## 计算空气动力学

- 空气运动以及与其它物体相对运动时的现象和作用规律
- 航空、航天、航海、交通、建筑、桥梁、能源、环保等领域

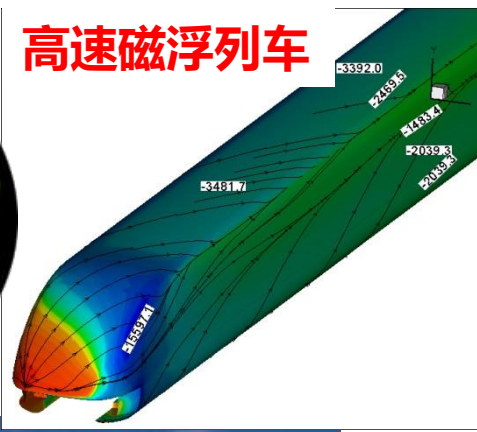
风力机设计



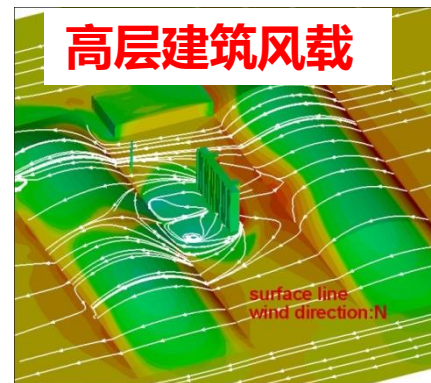
汽车减阻



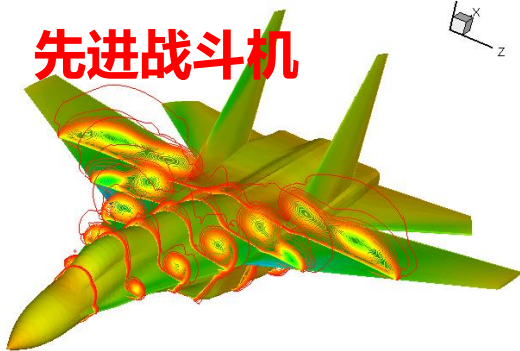
高速磁浮列车



高层建筑风载



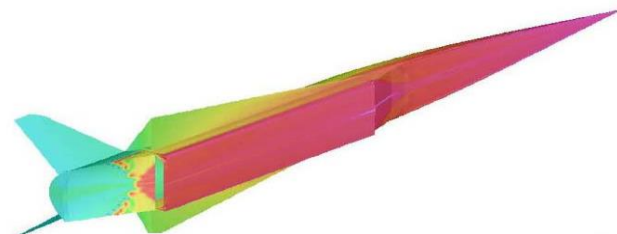
先进战斗机



大飞机工程



高超声速科技工程



## 空气动力学的三大研究手段

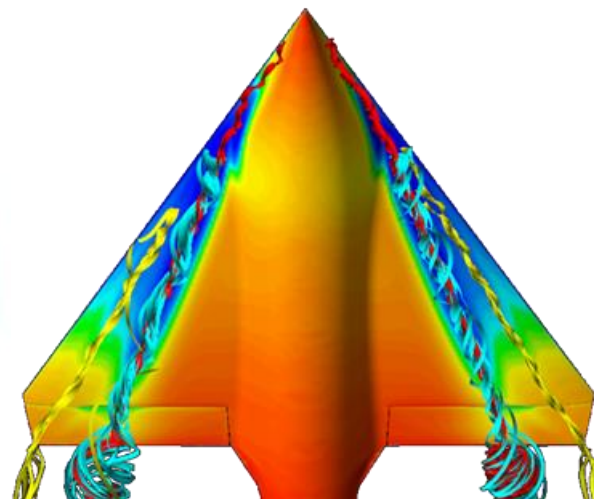
- 风洞试验、模型自由飞和数值模拟
- 由于模型自由飞的高昂成本和高风险，主要研究手段为风洞试验和数值模拟



风洞试验



模型自由飞



数值模拟

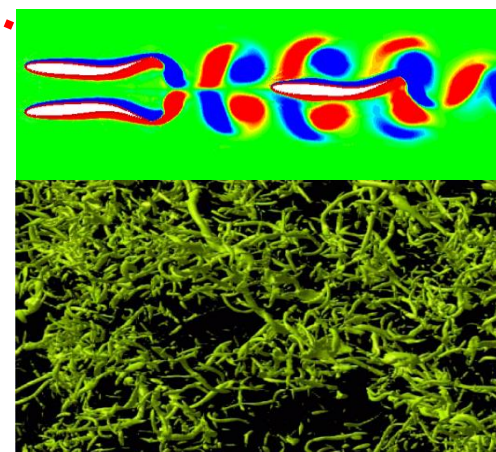
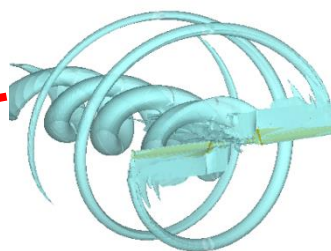
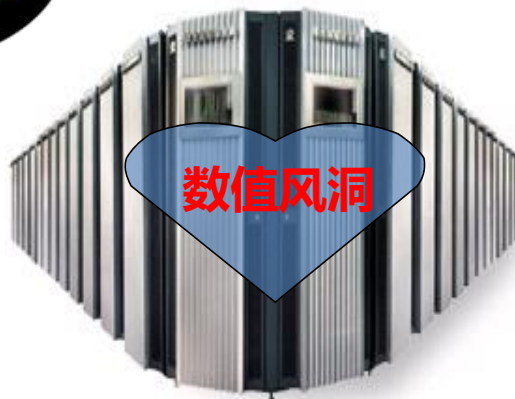
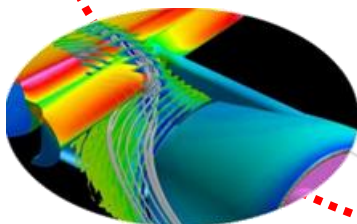
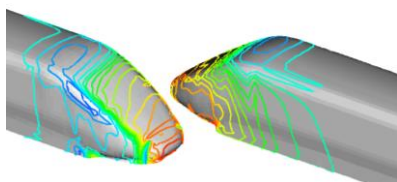
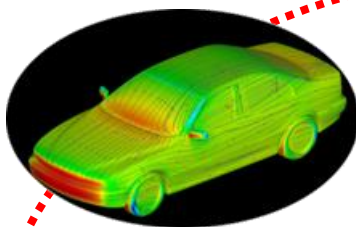
风洞试验约占空气动力学研究设计总工作量的70%以上



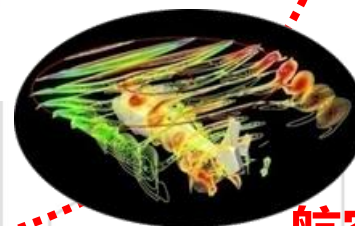
- 数值风洞：将CFD软件系统与大型计算机系统有机结合，能像风洞一样批量生产满足型号工程设计精度和效率要求的飞行器气动数据的基础设施



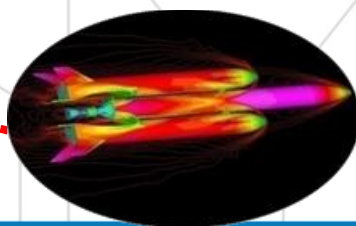
工业空气  
动力学



流动机理研究

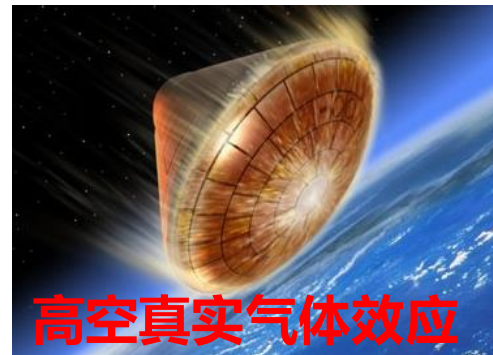
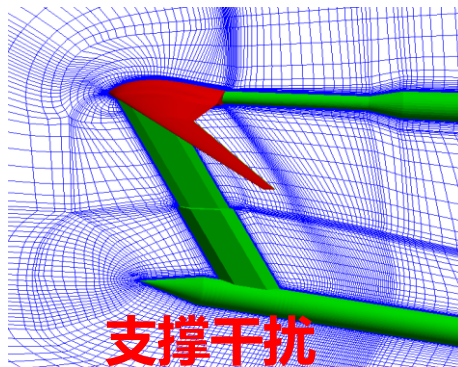


航空航天技术

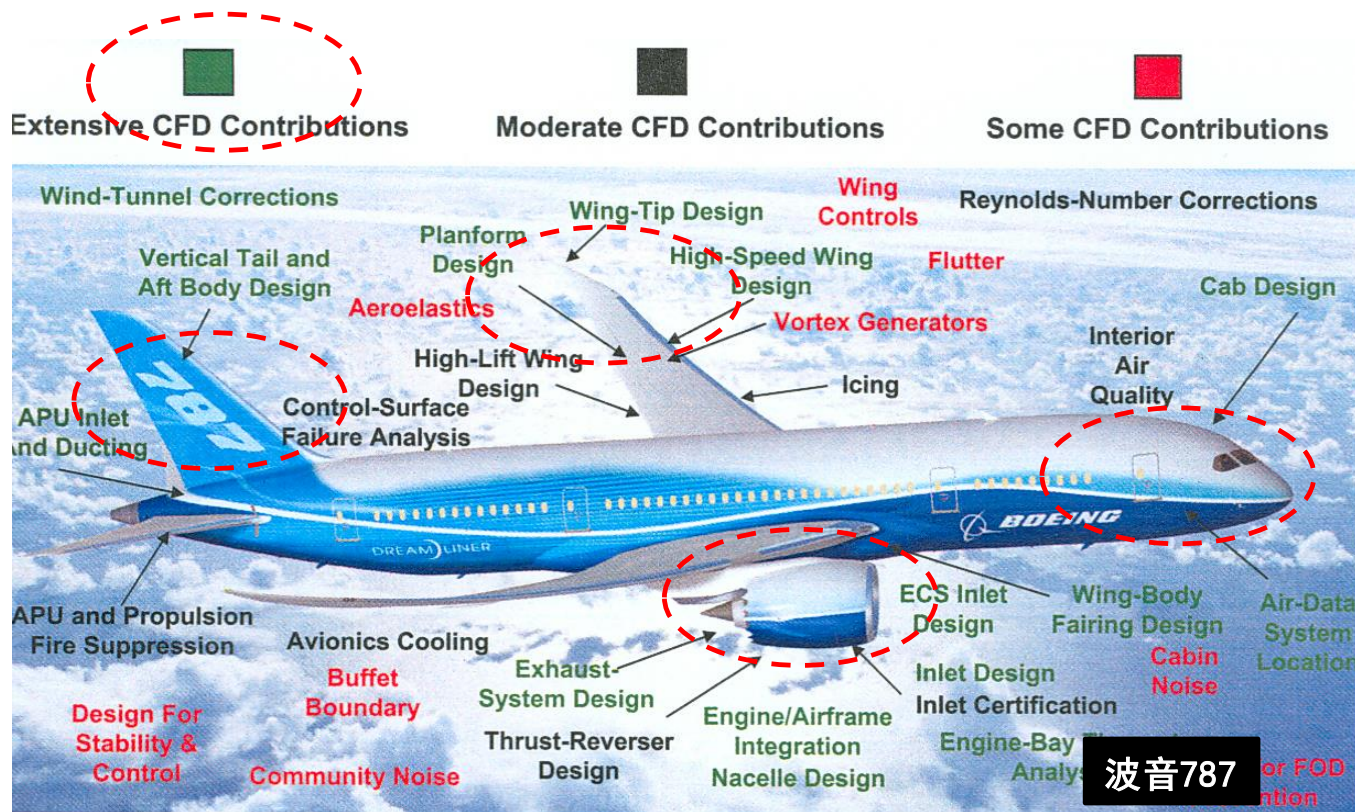


## 数值风洞

- 节约大量的试验费用
- 无洞壁、支架和测量干扰，相似性参数能全部满足
- 可实现许多风洞试验无法实现的空气动力学模拟，使空气动力学特性设计更为科学化。
- 化学反应、真实气体效应问题
- 刻画流动细节；强非定常问题







时至今日,在美国飞机研制领域, **CFD已占到气动设计工作量的70%左右**, 而风洞试验只需要做几个模型,试验工作量也下降到大约只占气动设计工作量的30%。

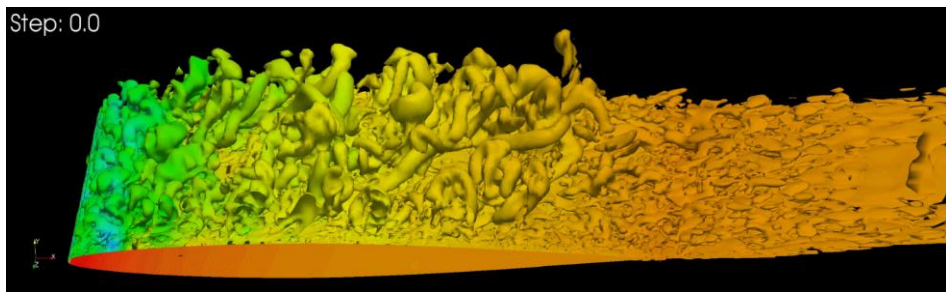
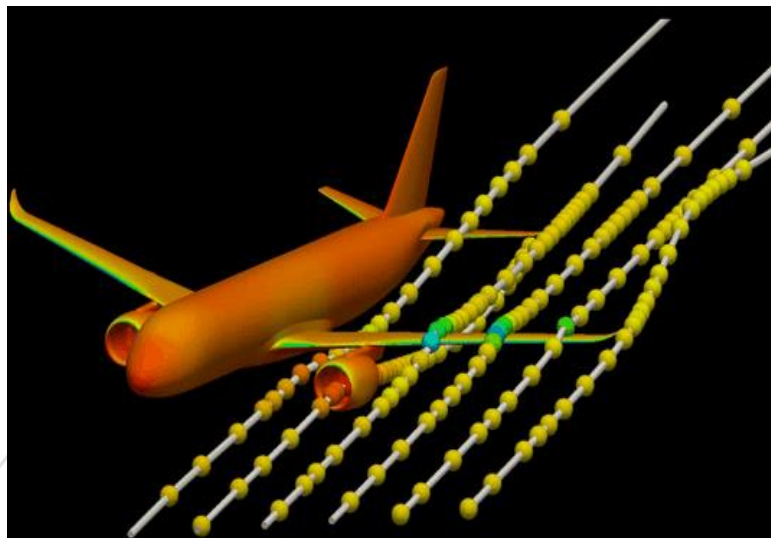
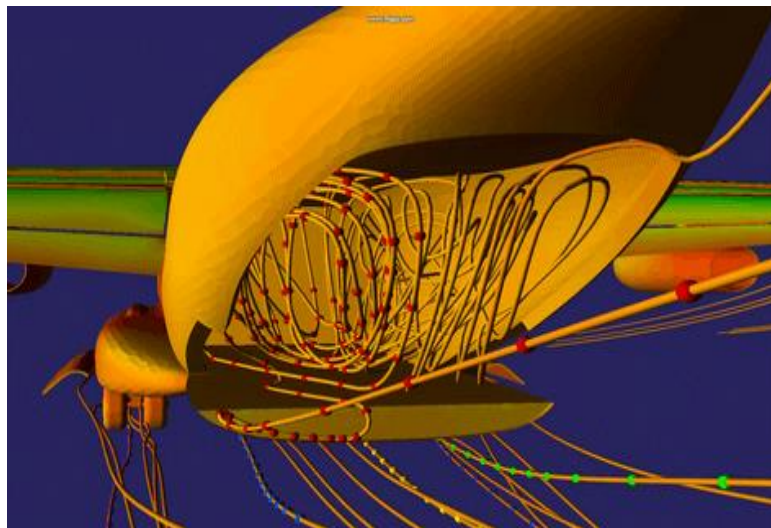
除此之外,通过以CFD为基础的多学科优化方法,还可以大大提高飞机的总体性能,这更是其它手段所不能替代的

- 在实际的空气动力学设计中，计算机只有在10到15分钟内实现模拟，它才能当作一种工程手段
  - 如果要在15分钟内对飞机的机翼进行大涡流模拟，需要千万亿次（P级）计算能力
  - 如果要在相同的时间内对整架飞机进行大涡流模拟，则需要百亿亿次（E级）计算能力
- 多学科优化设计
  - ◆ 气动/结构、气动/结构/控制、气动力/气动热/隐身等
  - ◆ 学科间的耦合增加了计算量，速度要求可能高几个数量级

**计算空气动力学对超级计算机有无止境的能力需求**

## 天河2号应用：航空航天

- 国产大型运输机
- 国产大型客机
- 航空发动机



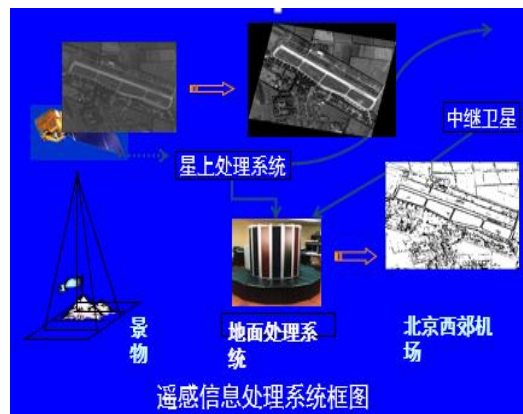


## ● 地理信息处理

- 卫星图象处理、航拍图象处理、各种传感器图象处理等

## ● 数字地球

- 把地球的地形地貌信息都收集起来，按照地理坐标，建立完整的信息模型
- 为快速、完整、形象地了解地球上各种宏观、微观的情况提供了可能，用途十分广泛



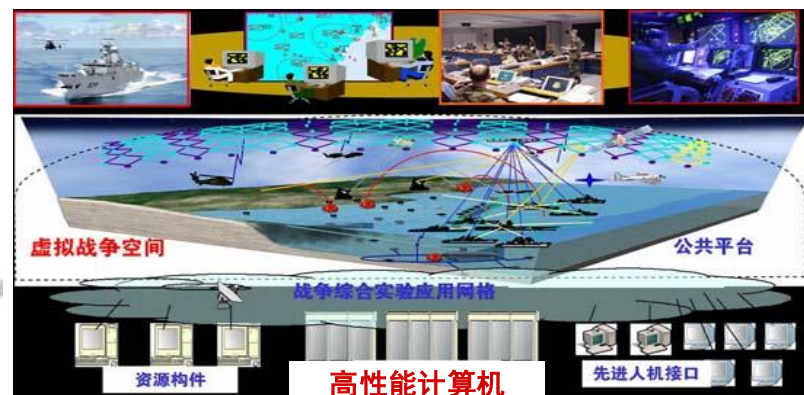
## 仿真

— 大样本、超实时

● 社会

● 舆情

● 演习





## 天河2号应用：天文地理

### 中微子对宇宙大尺度结构演化的影响

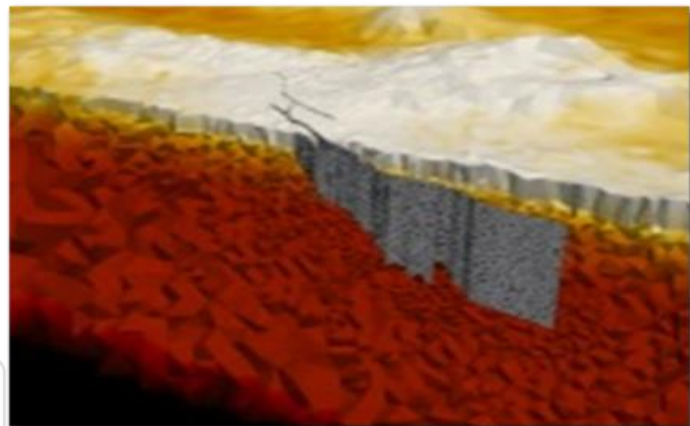
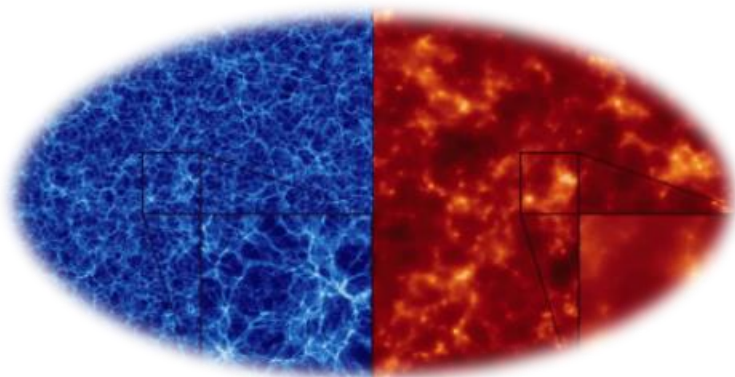
- 160万CPU和MIC核，成功进行了1100亿粒子的模拟

✓ 国际上目前规模最大的中微子宇宙尺度演化模拟 **Nature Astronomy** 发表

### 大规模地震模拟

- 大规模地震模拟应用Seissol，入围2014年国际高性能计算应用领域的最高奖 - GordonBell奖提名

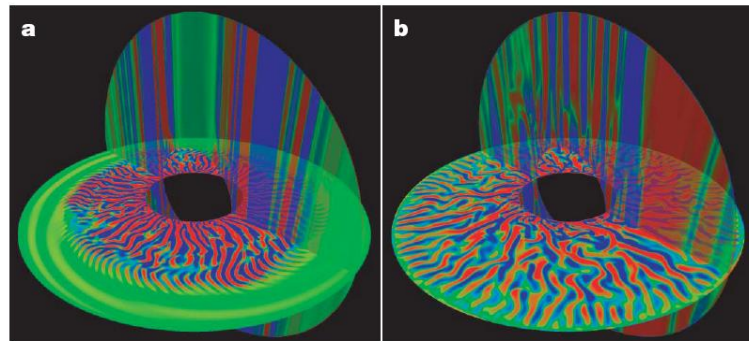
✓ 这是中国高性能计算系统上的应用首次入围该奖项



## 天河2号应用：天文地理

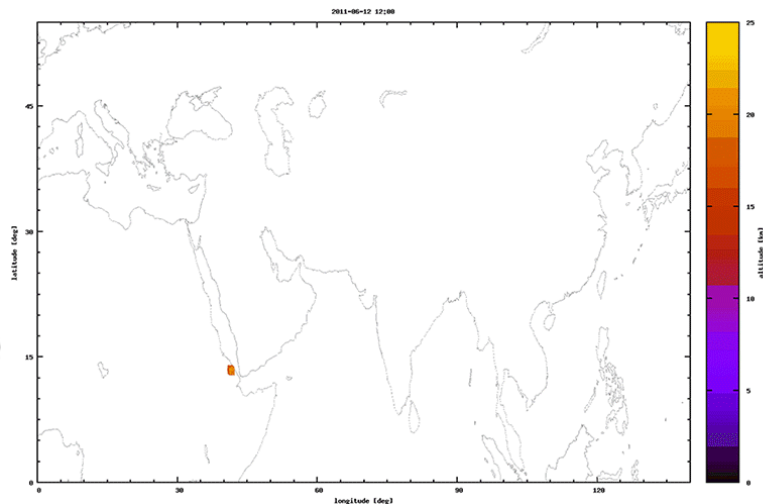
### 地核地幔的流体力学模拟

- 对判断星球能否成为可居住行星有重要指导意义
- CPU+MIC矩阵算法比petsc效率提高4倍，与单核串行的运行结果相差在 $10^{-10}$ 量级内，最大程度利用了天河二号计算能力



### 火山爆发污染物传输动态模拟

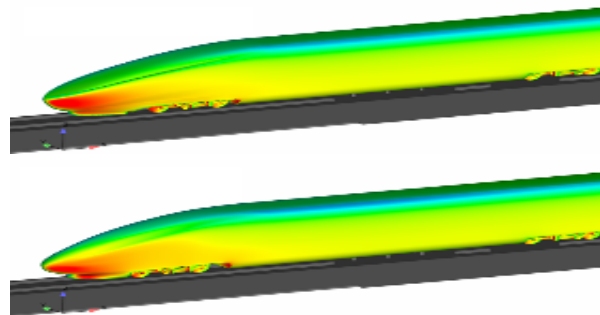
- 结果与AIRS, GOME-2和IASI等卫星观测数据吻合
- 成功捕捉火山爆发期间二氧化硫污染物在我国境内传输的动态信息



## 天河2号应用：工业设计制造

### ● 南车高速列车

- 外型设计、高体爆波设计



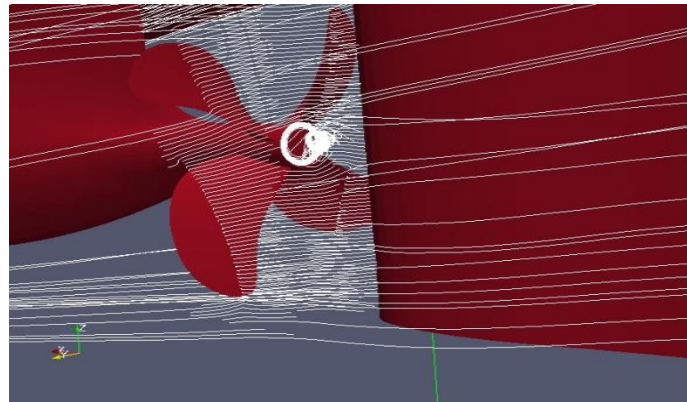
### ● 广汽集团车身侧面碰撞模拟

- 精度达到85%以上



### ● 广船万吨排水量船舶数值水槽

- 自主推进，缩短设计周期，降低研发成本

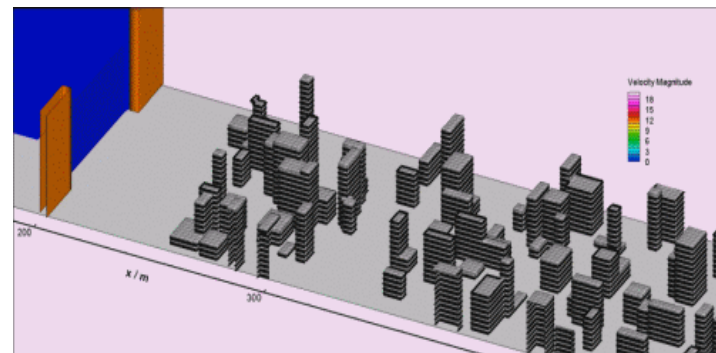




## 天河2号应用：工业设计制造

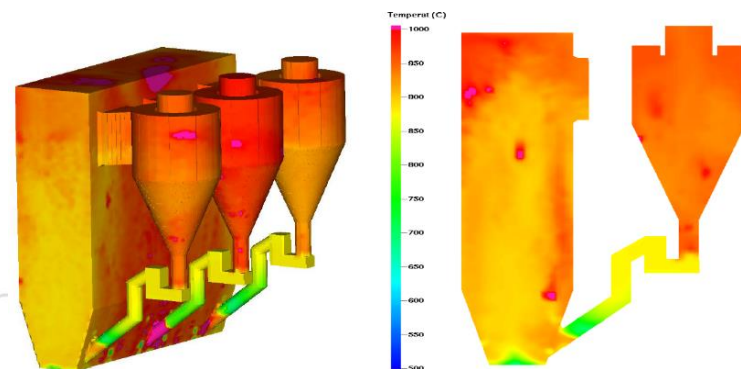
### 三维溃坝水流精细仿真模拟

- 突破混凝土坝溃决机制离散元模拟在实际大型水利工程应用的效率瓶颈
- $1\text{ m}^2$ 的高分辨率网格覆盖约11平方公里城区，5分钟内完成洪水传播全过程模拟，计算效率国际领先



### 大型循环流化床锅炉全热态数值模拟

- 对大型锅炉进行“CT扫描”、对锅炉设计、提高燃烧效率，控制有害气体排放有重要指导意义
- 国际首例基于Barrcuda软件的300MW循环流化床数值模拟



## 天河2号应用：新能源新材料

### 风电设计全生命周期

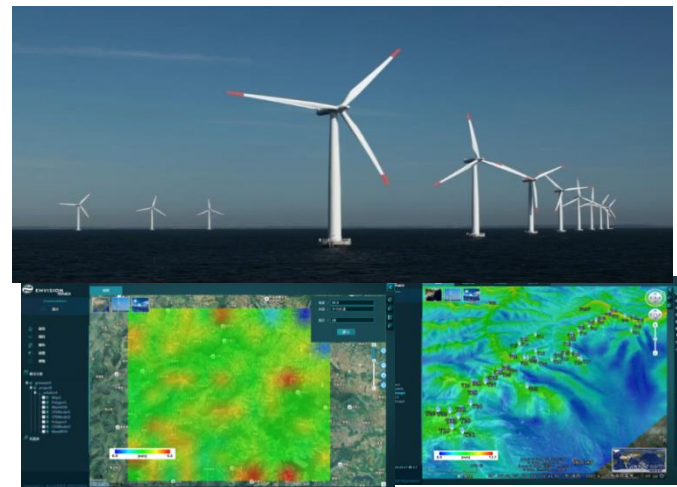
- 风机设计-全球风资源预测-风力发电机组选址优化-风电场项目评估
- 智能、精准、稳定、高效
- 0.5 小时完成一个风电场分析

### 纳米材料研究

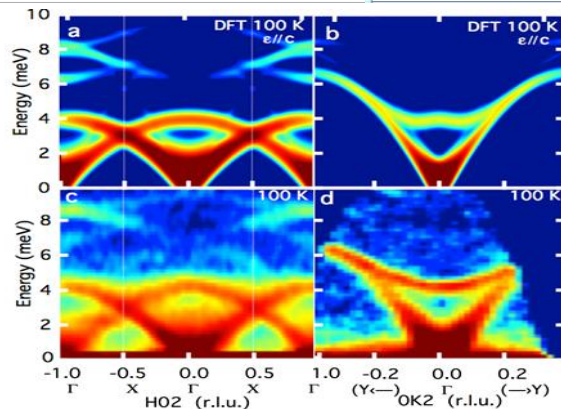
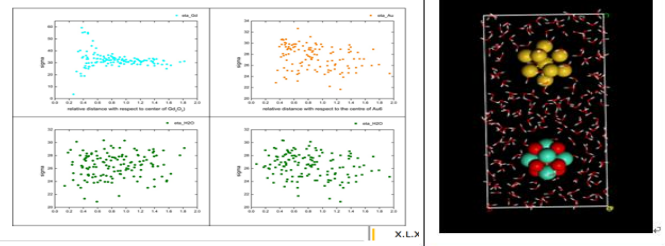
- 基于介孔硅的核磁荧光双模态靶向纳米探针机理研究及制备
- 用于癌症检测的光 - 磁双模态纳米探针

### 先进功能材料与能源材料研究

- 首次获得金属绝缘体相变材料VO<sub>2</sub>完整的温度相关的声子色散关系
- 与散裂中子源、同步辐射源等国家重大科技基础设施运行单位深入合作，加速科研成果产出



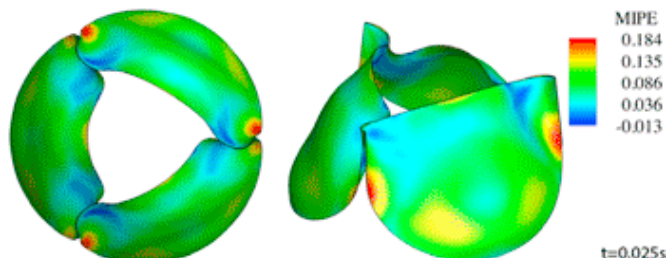
chemical shift calculations



## 天河2号应用：生物医药健康

### 1. 组织器官的模拟

- 腹主动脉瘤
- 心脏瓣膜跟血液的相互作用



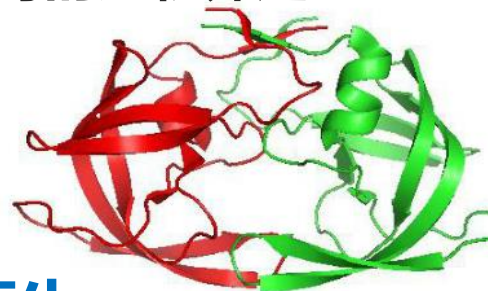
### 3. 人类基因组相关研究

- 基因组装
- 全基因组重测序及相关应用
- 广泛应用于癌症治疗、产前检查等



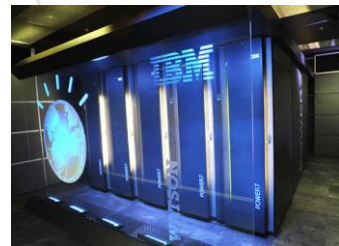
### 2. 药物研发

- 使用分子对接技术发现HIV-1蛋白酶抑制剂
- 抗埃博拉病毒的药物筛选



### 4. 人工智能医生

- 病历、影像、文献分析
- 协助医生提供个性化的疾病治疗方案



## 生物数据爆炸式增长

### 欧洲生物信息学研究所 (EBI)

- 核酸序列数据库：37 亿条记录，7700 亿碱基，共20PB



### 美国国家生物技术信息中心 (NCBI)

- 40个数据库，总共包含约69 亿条记录



### 国际肿瘤基因组TCGA

- 33种癌症数据，仅组学数据即达到2.5PB



### 日本DNA 数据库 (DDBJ)

- 截止2016.3超过2007亿碱基数据



### 深圳国家基因库 (CNGB)

- 规划有超过500PB存储、百万物种和种质数据

### 广东省历年医疗病历摘要记录以及医学影像数据

- 已有数十TB~PB量级，且持续增长

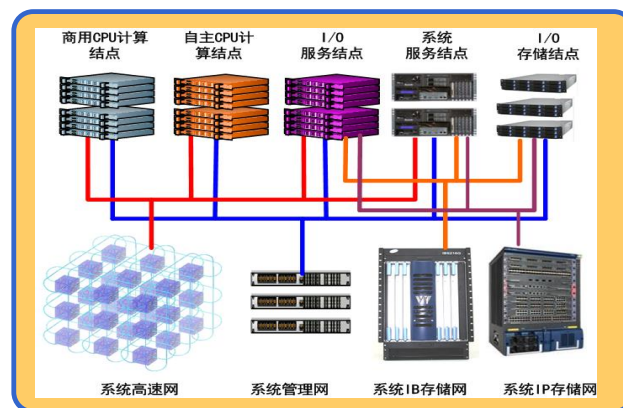




## 平台架构？领域专家

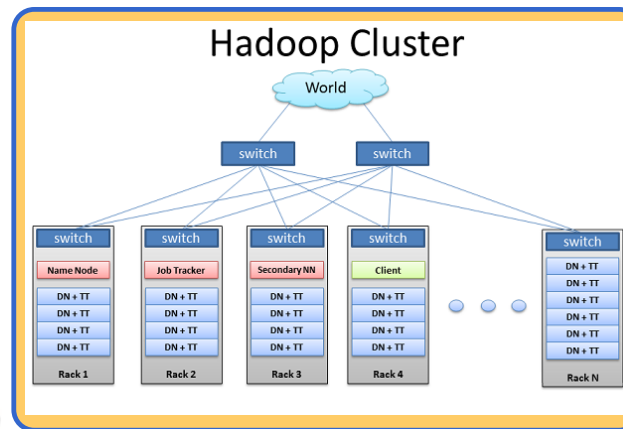
### 高性能计算平台 (MPI+OpenMP)

- 针对科学与工程迭代计算
- 大规模并行，计算性能高
- 不适合生物专用算法特征



### 大数据分析平台 (Map-Reduce)

- 针对海量数据批处理
- 高通量计算，吞吐率高
- 处理模式单一，难以应对复杂多样的  
生物学大数据应用



生物学健康大数据应用的生命周期处理亟需能够  
支持多种处理模式的异构融合平台



## 数据管理？开放共享

## ◆ AutoDOCK

- Lustre、GPFS

## ◆ Genome Analysis Toolkit

- HDFS、GFS

## ◆ Cloud BioLinux

- S3、Ceph

药物设计筛选

组学数据挖掘



云端数据处理



## 生物医学健康大数据

医学影像存储

个人健康档案

诊疗记录存储



## ◆ 美国Explorys公司的临床医疗决策数

## ◆ Oracle Healthcare Transaction Base

## ◆ 海量小文件

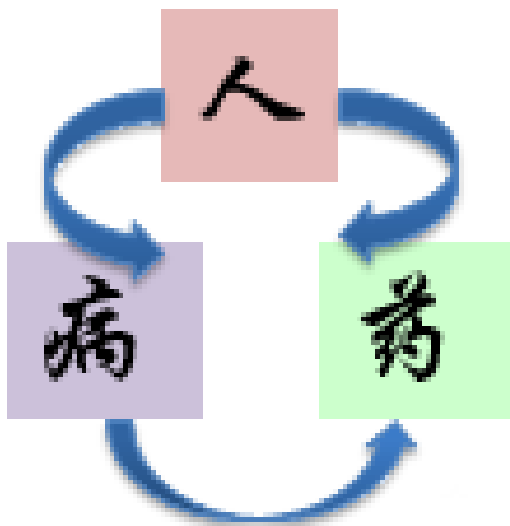
ack、TFS

亟需将多源异质海量数据整合存储并统一管理

## 天河2号应用：生物医药健康

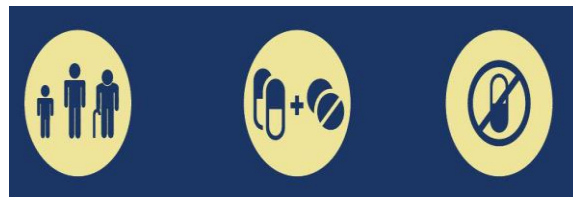
### 精准医疗技术

- 基因测序技术的飞速发展
- 超算技术、大数据存储分析技术的广泛使用
- 医学前沿技术



### 短期目标：癌症治疗

- 靶向抗肿瘤药物的试验
- 多种药物的联合疗法的有效性
- 抗药性产生的原因和克服方法



### 长期目标：健康管理

- 建立百万人级别的数据库
  - 基因数据, 生物样本
  - 电子健康记录, 生活饮食信息
- 基于这些数据开展研究
  - 药物基因组学
  - 疾病预防
  - 基因编辑

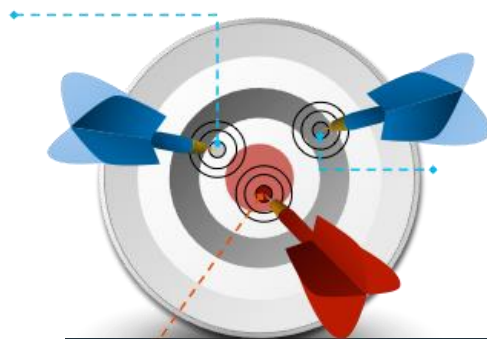


## 天河2号应用：金融风险分析

### ● 千万级投资组合的多维度市场风险计量

- 天河2号加速基于蒙特卡罗模拟的VaR计算过程
- 计算节点规模扩展209倍，程序整体加速151.69倍，扩展效率达到72.58%
- 利用数千个天河2号的异构计算结点实现千万级投资组合的多维度市场风险分钟级准实时计量

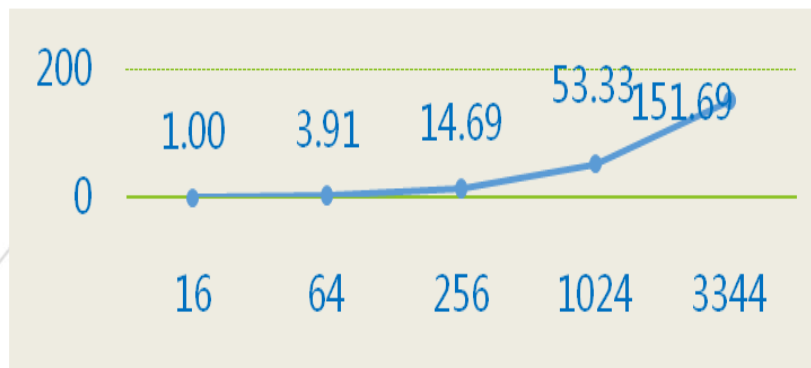
### ● 自动化交易



高性能回溯测试  
并行计算模块



高性能压力测试并行计算模块



计算规模：3344 个计节点，总核数65.2 万  
(80256 颗CPU核心，10032 块MIC卡)

**● 广州市无线电监测中心**

## — 170个监测站点

## 一 每天200G~流量

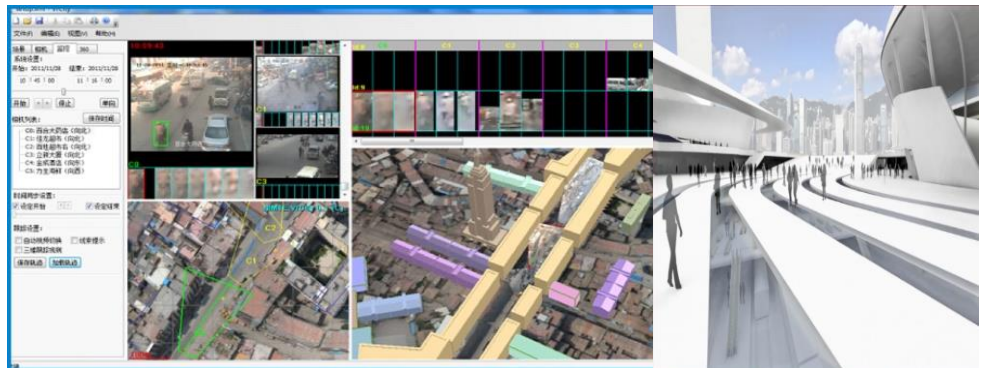
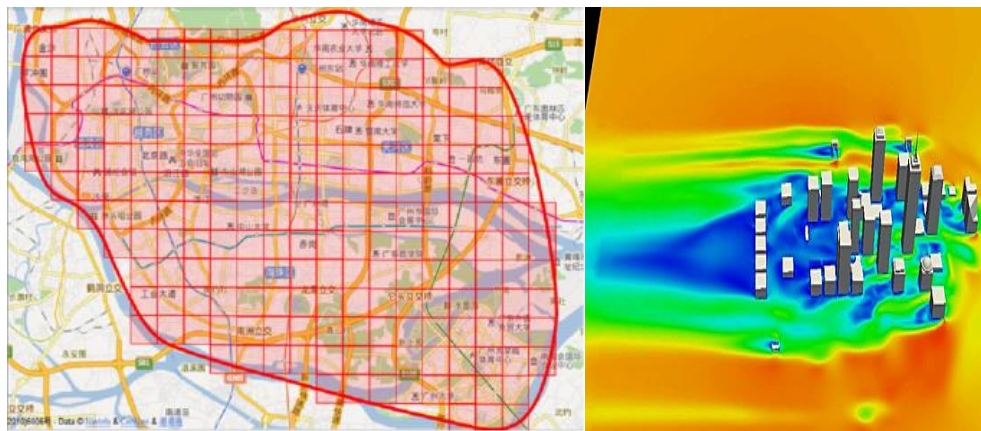
## ● 广州市道路卡口监控

## — 2026条道路，每天3T的存储

## 广州市地铁交通数据

## 一 日均客流600多万

## ● 教育信息公共服务平台





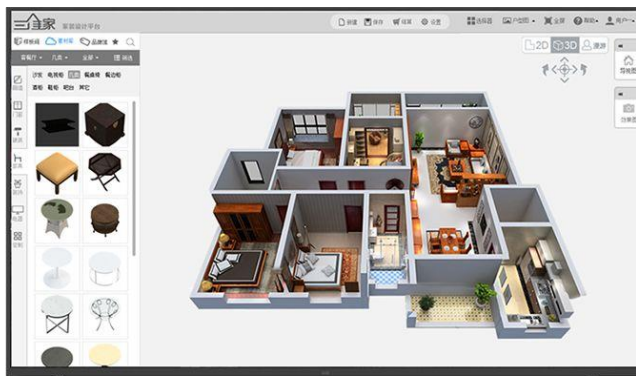
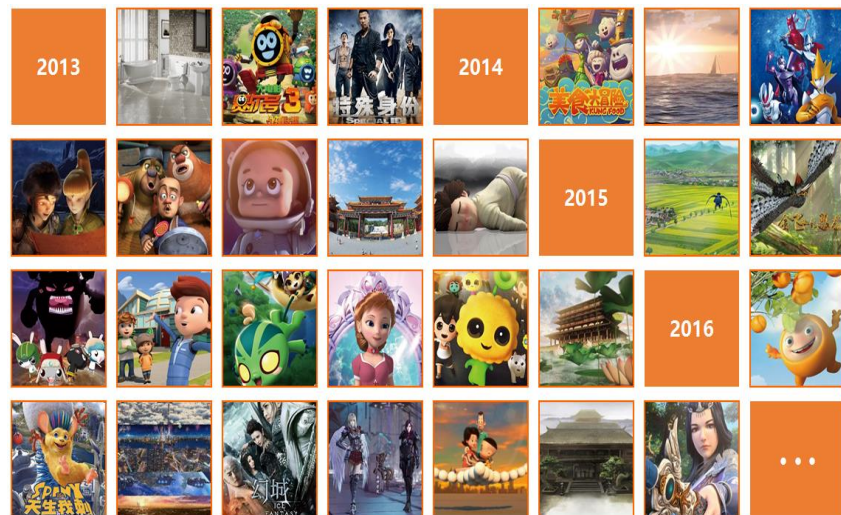
## 天河2号应用：文化创意设计

### 专业渲染云

- 影视、动漫制作
- 互动/游戏
- 建筑/视觉

### 3D家居设计云平台

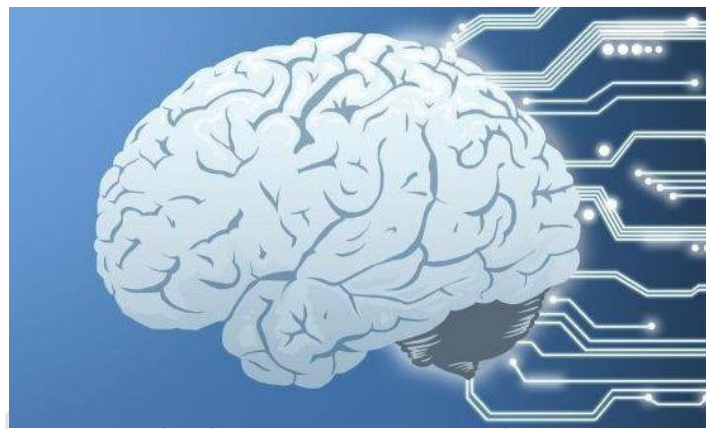
- 家居设计→3D渲染→VR体验
- 3分钟得到全屋720°全景图
- 国内外25万终端用户，每月215万张设计图



## “人狗大战” 引爆深度学习

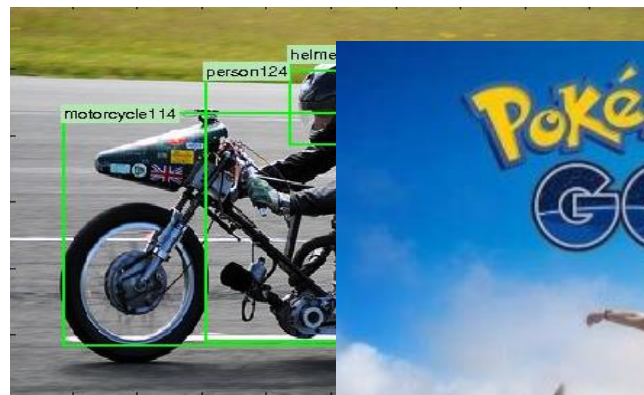
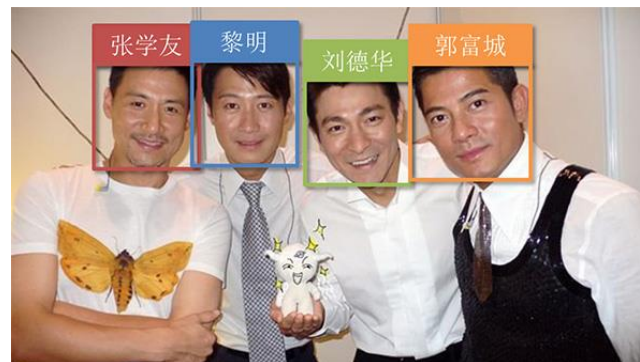
### 深度学习：AlphaGo VS 九段棋手李世石 = 4 : 1

- 人工智能的核心技术
- 模仿人脑的机制来解释数据，例如图像，声音和文本。
- 具有高精确度、自动学习、持续学习能力等优势



## 深度学习应用场景无处不在

- 人脸识别
- 图像搜索
- 广告预测
- 语音识别
- 文字识别
- 商品分类
- 视频分析
- 车牌识别
- 医学诊断





## 深度学习三大需求

- 巨大的数据样本
  - 图像（亿级）. 语音（万小时）
- 先进算法、模型、软件
  - 神经网络算法：DNN\CNN\RNN
  - 深度学习软: Caffe\TensorFlow\MXNet
- 高性能计算
  - AlphaGo
    - 1920个CPU的集群
    - 280个GPU的集群

## HPC与大数据结合推动深度学习的发展

## 深度学习带来HPC新模式：HPDA

- HPC + Big Data + Analysis = HPDA





- AlphaGo Lee
- AlphaGo Master
- TPU 40T
- TPU2 180TF
- TPU舱室 64个TPU2 11.5PF

- Tensor Processing Unit

- 加速推理，加速训练



模型和算法的进化  
+ 计算能力的提升

## 天河2号应用：深度学习

## 深度学习应用行业越来越广

- 医疗健康
- 金融行业
- 互联网
- 安防行业
- 公安行业
- 石油勘探
- 智能家居
- 数据分析
- 工业制造

- 无人机
- 机器人
- 自动驾驶



## 1. 面向学术前沿基础研究

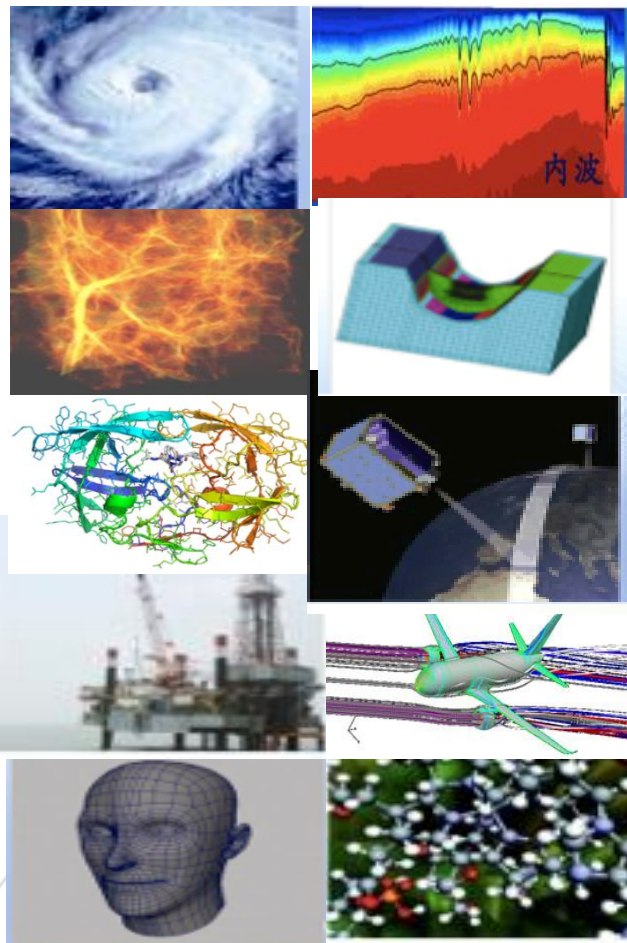
— 大科学重大创新成果

## 2. 面向国家重大发展战略

— 大工程重大应用突破

## 3. 面向地方经济社会发展

— 产业转型升级新动力



## ● 依托超算中心天河二号支撑大数据产业

### — 电子政务

- 搭建可扩展云平台，与国内顶尖企业合作，提早部署，专注架构、软件、网络、安全，具备上线试运行的能力



### — 智慧医疗

- 围绕精准医学中心的建设目标，基于基因、药物、病历影像等搭建超算和数据处理平台，已建成基因分析和药物筛选平台



### — 智慧交通

- 结合交通和公安大数据，构建数据存储、数据分析、数据共享平台，支持广泛的交通管理、城市道路规划、安防警务等多种应用



### — 智能制造

- 基于人工智能、深度学习、大数据实时分析成为支持工业、企业、媒体等多领域的人机交互的平台



### — 金融大数据

- 个人征信、企业征信、金融产品、金融服务的大数据分析处理与挖掘使用，提高金融服务水平、防控金融风险、维护金融秩序



### — 北斗导航信息处理系统

- 充分挖掘卫星导航数据的价值，支持交通、地理、地质、水文、防灾减灾、农林牧业、城市规划等广泛的应用





## 大数据时代的挑战

8.2EB data accumulated in Chinese internet industry by 2015

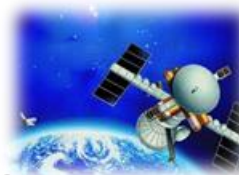
Internet



IoT



200PB data produced by Smart City projects in China



Astronomy

16TB data each day produced by a telescope in space

2PB data captured by 50,000 security cameras in Guangzhou



Traffic

More than 17PB biological information data accumulated in BGI of China



Medicine

More than 10 PB data of global weather report increasing by 50% each year



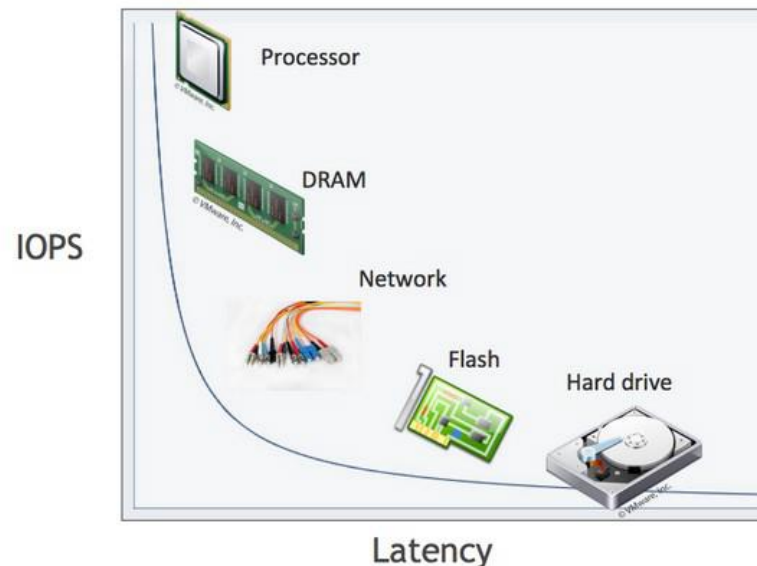
Meteorology

**Big Data**

## 大数据时代的挑战

### 数据访问时间

- CPU Registers:  $1 \times 10^{-10}$  sec
- L1 cache:  $10-50 \times 10^{-10}$  sec
- L2 cache:  $100-500 \times 10^{-10}$  sec
- Main RAM:  $1-10 \times 10^{-9}$  sec
- Network:  $10 \times 10^{-6}$  sec
- Flash Memory:  $200 \times 10^{-6}$  sec
- Disk:  $5 \times 10^{-3}$  sec



If Memory = Minute  
Network = Weeks  
Flash = Months  
Disk = Decades



## ○ E级系统的挑战

Systems	2009	2013	2017	2020
System peak	2 Peta	50 Peta	125 Peta	1 Exa
System memory	0.3 PB	1.6 PB	1.3 PB	10 PB
Node performance	125 GF	200 GF	3TF	1-10 TF
Node memory BW	25 GB/s	40 GB/s	14GB/s	200-400 GB/s
Node concurrency	12	32	O(100)	O(1000)
Interconnect BW	1.5 GB/s	22 GB/s	12GB/s	50 GB/s
System size(nodes)	18,700	30,000	40,000	O(million)
<b>Total concurrency</b>	<b>225,000</b>	<b>3,200,000</b>	<b>O(10,000,000)</b>	<b>O(billion)</b>
Storage	15 PB	30 PB	20PB	300 PB
IO	0.2 TB/s	2 TB/s	0.3TB/s	20 TB/s
<b>MTTI</b>	<b>1 day</b>	<b>4 h 48 min</b>	<b>unknown</b>	<b>29 min</b>
Power	6 MW	~10MW	~15MW	~20MW

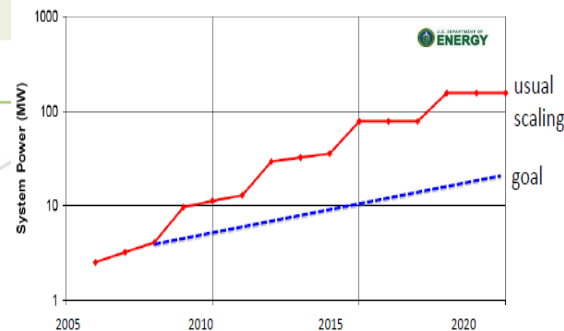
存储墙

通信墙

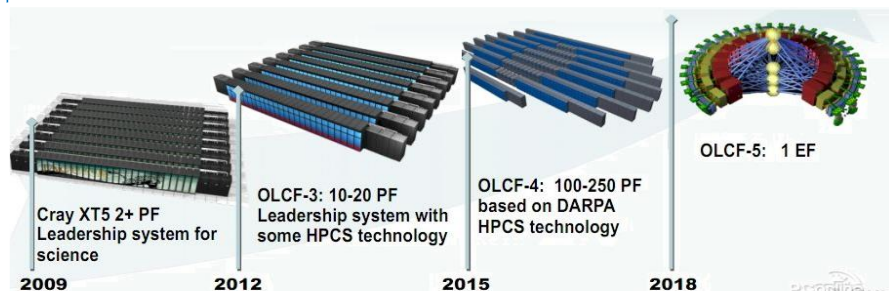
可靠性墙

能耗墙

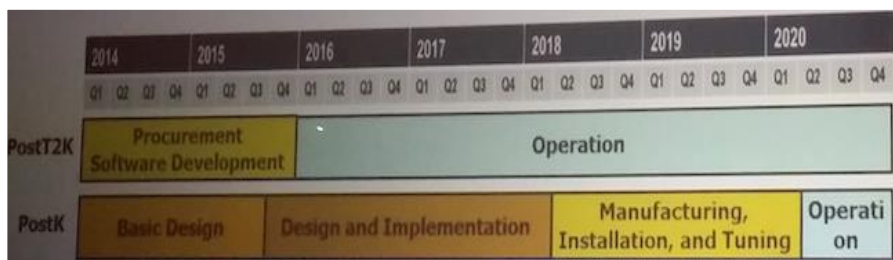
编程墙



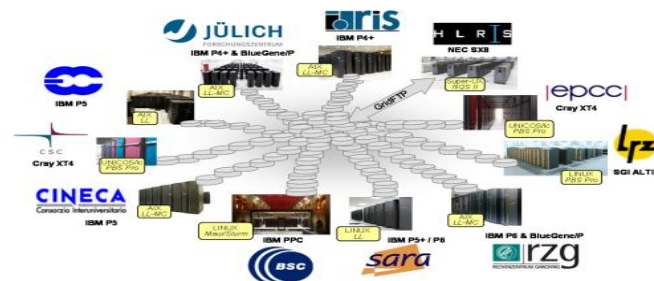
美国将E级计算列为21世纪美国最主要的技术挑战，实施了UHPC、ASCR、IESP、Coral，NSCI等计划



日本HPCI计划，近期推出Post T2K和Post K计划，2020年推出E级系统



泛欧成员机构发起了DEISA 计划，G8科学研究理事会也提出了E级计算计划 EESI, PRACE, Hc



俄罗斯：“2012-2020俄罗斯E级技术发展”计划，2020年前实现E级计算机  
印度：计划加大投入旨在建成世界最快的高性能计算机

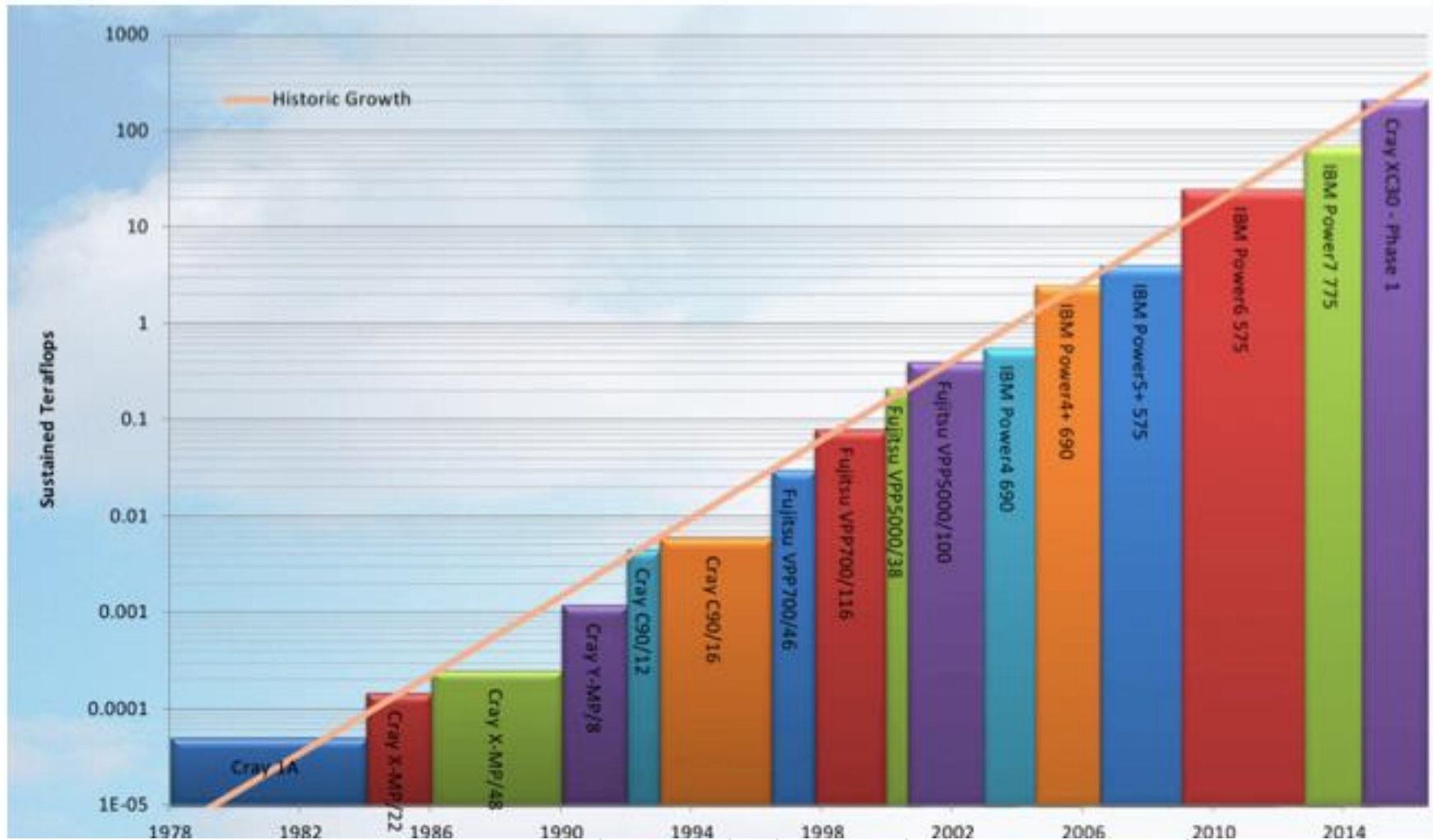
中国：十三五国家重大研发计划  
高性能计算专项

3台pre-E级系统，1台E级系统  
数值反应堆，数值发动机，数值风洞  
数字地球，数字人体.....





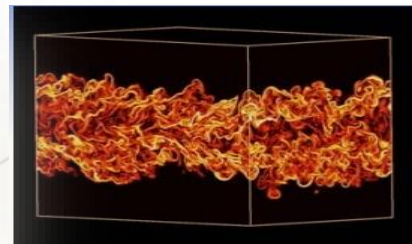
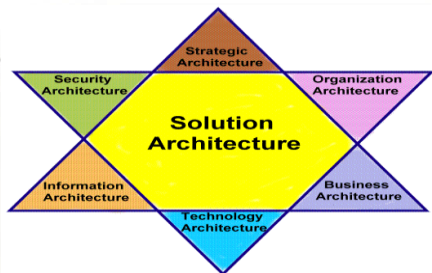
“only” 100-fold performance improvement for climate codes



## ● 软件技术是发挥大规模计算系统能力的保障

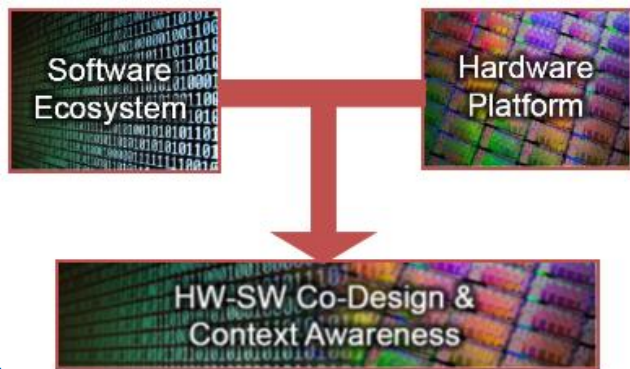
### — 下一代系统中领域软件所面临的机遇

- 数值模拟已成为科技发展的重要推动力量——研究的进展取决与模拟的规模和速度
- 超级计算机运算能力越来越强大，持续发展——需要大规模的、面向领域的专业软件



## ○ E级计算的挑战 $10^{18}$

构建平衡、动态、自适应的E级计算系统



### ○ 复杂性

- 充分挖掘体系结构能力

### ○ 可扩展性

- 松散同步与通信结构

### ○ 集合编程模型

- MPI+OpenMP+SIMD

### ○ HPC Vs. Big Data

- 混合层次存储系统与数据管理

### ○ 容错与自恢复

- 轻载策略与机制

### ○ 正确性

- 模型抽象、离散

### ○ 可扩展性

- 数据量与计算量

### ○ 容错、自恢复

- 容错模型

### ○ 可验证性

体系结构感知的软件与应用



## 高性能计算与大数据处理融合发展

1. 软件开发性能分析计算/数据处理融合应用支撑平台
2. 大气海洋环境应用平台
3. CFD与工业制造应用平台
4. 新能源新材料应用平台
5. 天文物理计算大数据应用平台
6. 生物医药健康大数据应用平台
7. 视频交通大数据应用平台
8. 数字媒体大数据应用平台
9. 深度学习与认知智能应用平台
10. 高性能计算与大数据处理教育平台

### 围绕超算中心建立

#### 国家高性能应用软件中心+大数据创新工程中心

- 工具软件、平台软件、应用软件开发与集成
- 聚集交叉学科人才，构建可持续发展生态环境

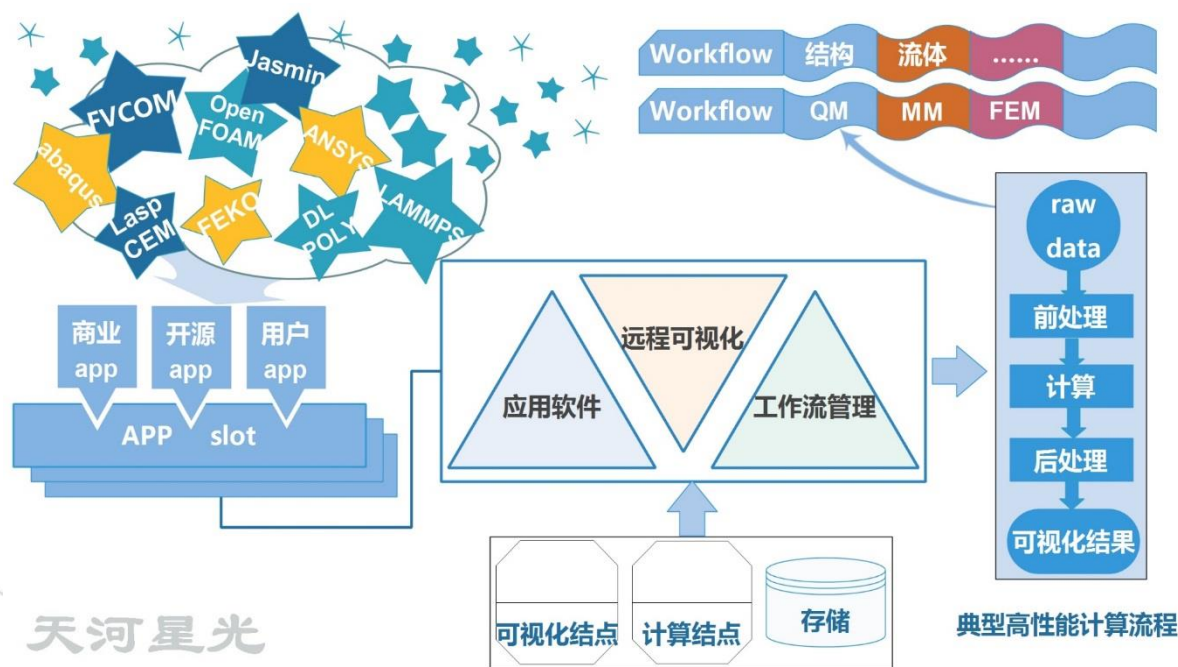
国产高性能计算系统的科学评价和应用引导

## 方便易用的超算系统使用界面

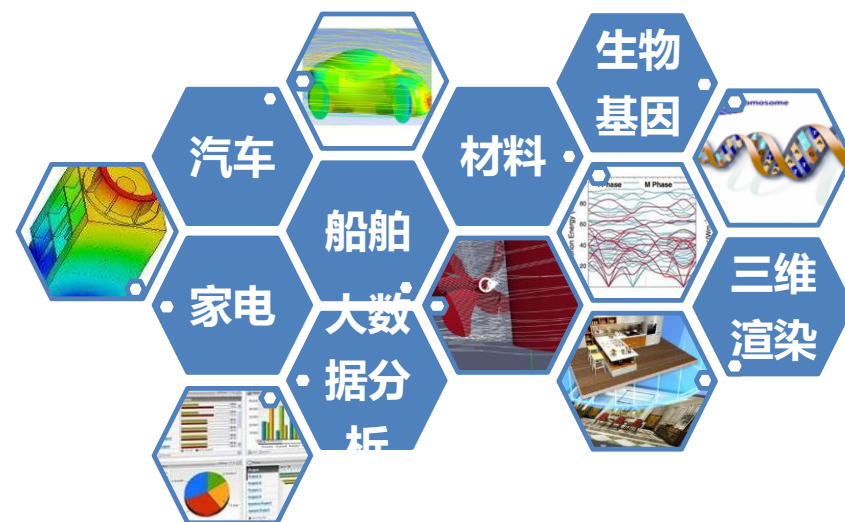
- HPC + Bigdata
- HPC on Cloud

## 天河星光云超算桌面

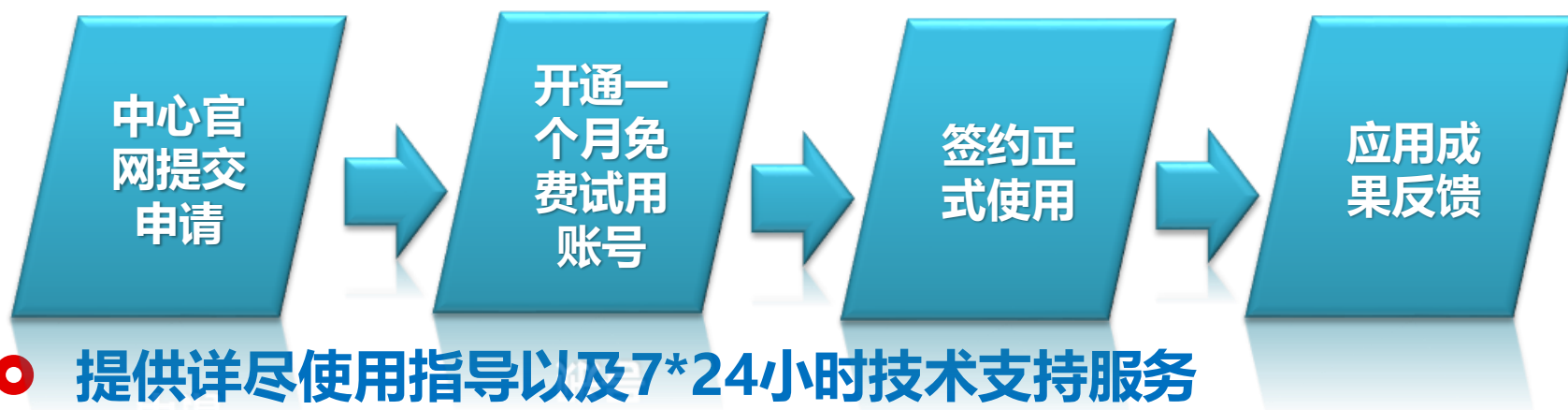
- 集成丰富的应用
- 提供远程可视化
- 高效管理工作流
- 提高超算易用性



- 为高性能计算、云平台&大数据提供强大计算与存储资源
- 打造独立、安全、资源弹性可扩展的超算服务平台
- 高达8G网络出口带宽，萝岗分中心（天河软件园）、中山分中心（中山工研院）、南沙分中心（南沙工研院、港科大霍英东研究院）还可提供网络专线直达我中心机房



## ● 高效、便捷的超算用户服务



## ● 提供详尽使用指导以及7\*24小时技术支持服务

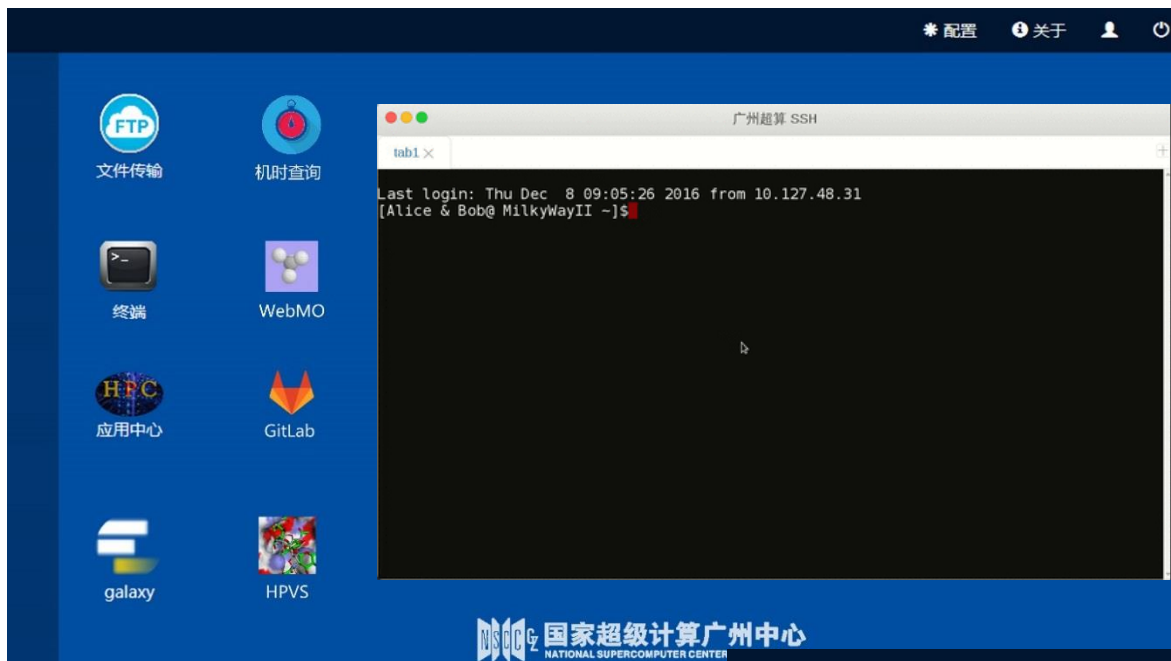
- 客服电话：020-37106021
- 客服邮箱：service.nscg@nscg-gz.cn





# 广州市发布“超算服务券”鼓励企业使用超算！

- 为了推动超算应用与产业融合发展，2016年11月广州市科创委发布了超算服务券实施细则（穗科创字〔2016〕275号），鼓励广大企业使用超算
  - 支持对象：广州市的科技型企业和众创空间
  - 支持内容：设立超算服务券，可后补助企业与超算签约合同金额的**40%**，单个申请人每年申请获取超算服务券金额最高500万元
  - 实施方式：全年开放在线申请（广州市科技计划管理系统 <http://wsbs.gzsi.gov.cn>），每年集中统一兑现



## 脚本式模式

欢迎使用ABAQUS求解器.

作业信息

作业名称:

项目名称:

选择应用

求解器版本:

输入文件[\*inp]:

额外选项:

输入文件与参数

作业参数

总核数:






























资源规模

提交作业

自定义作业名称

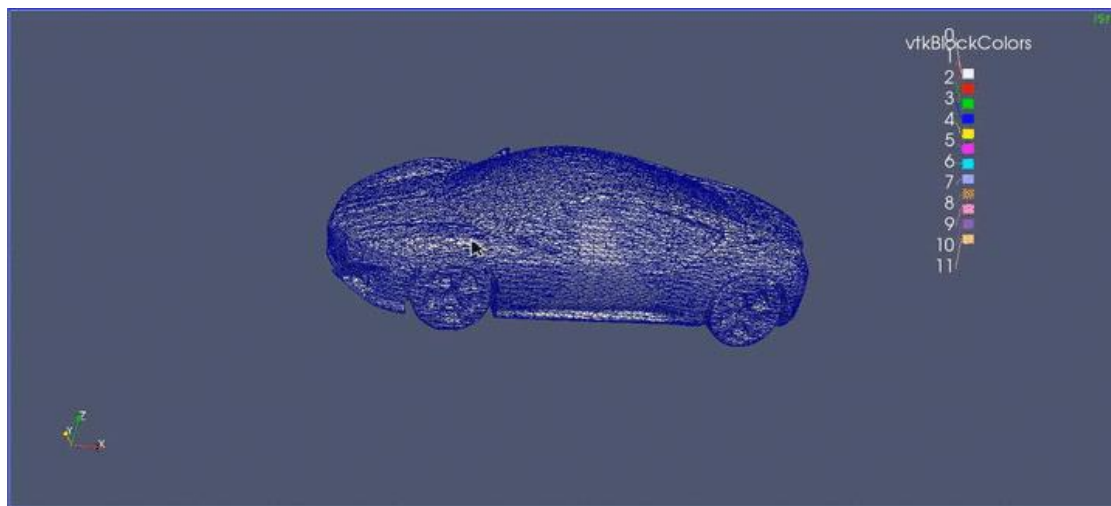
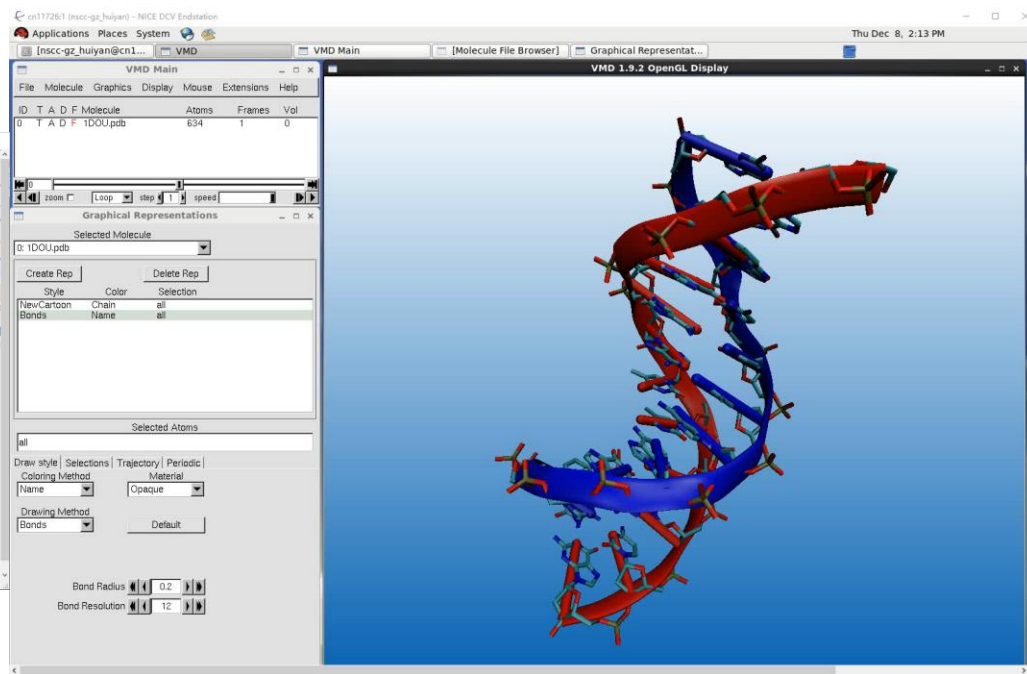
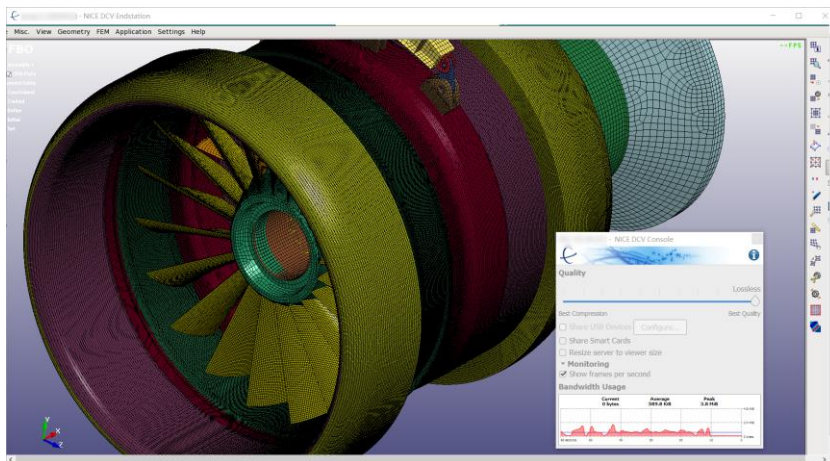
选取求解器

## 交互式模式

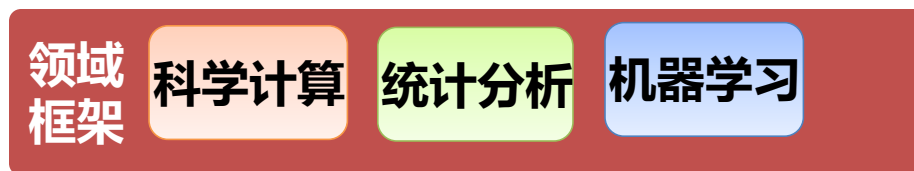
Menu	Sessions
▶ Data	Refresh   
▶ Monitor	Sort by: Started on 
▼ Application ...	
 ABAQUS ...	 
 Ensight	 
 ENVI IDL	☆  Hypermesh 
 Feko GUI	☆  Feko GUI 
 Fluent GUI	 
 Hypermesh	 
 Paraview	☆  Fluent GUI 
 Starccm+ ...	☆  Paraview 
 Visit	



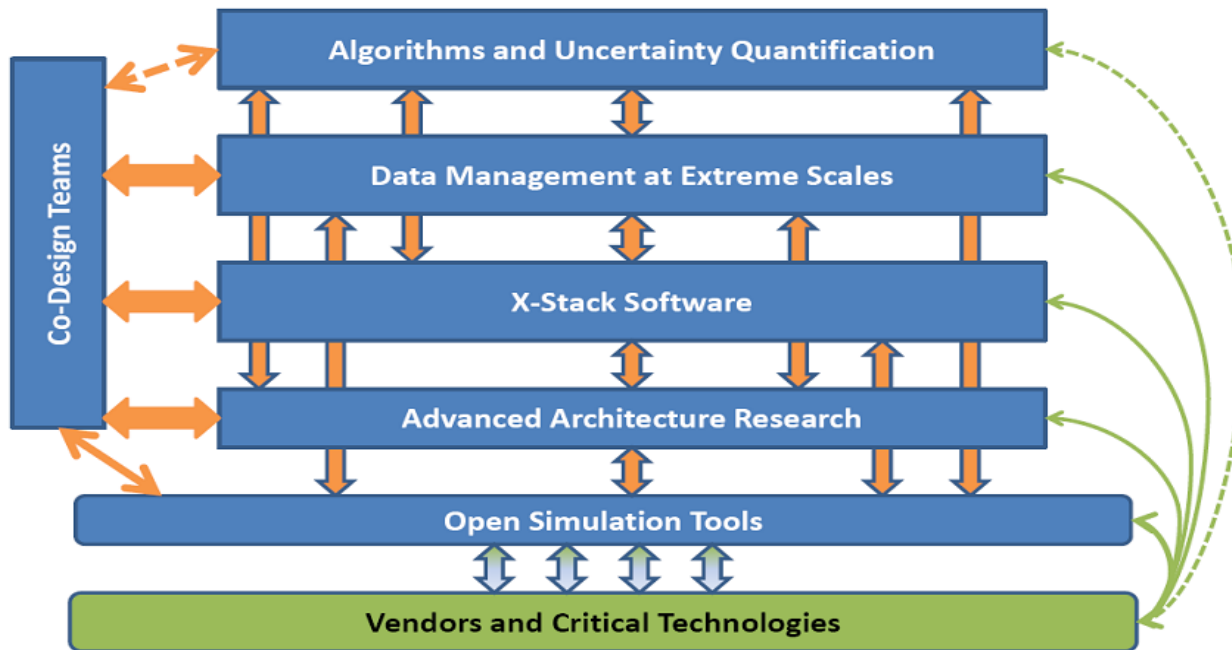
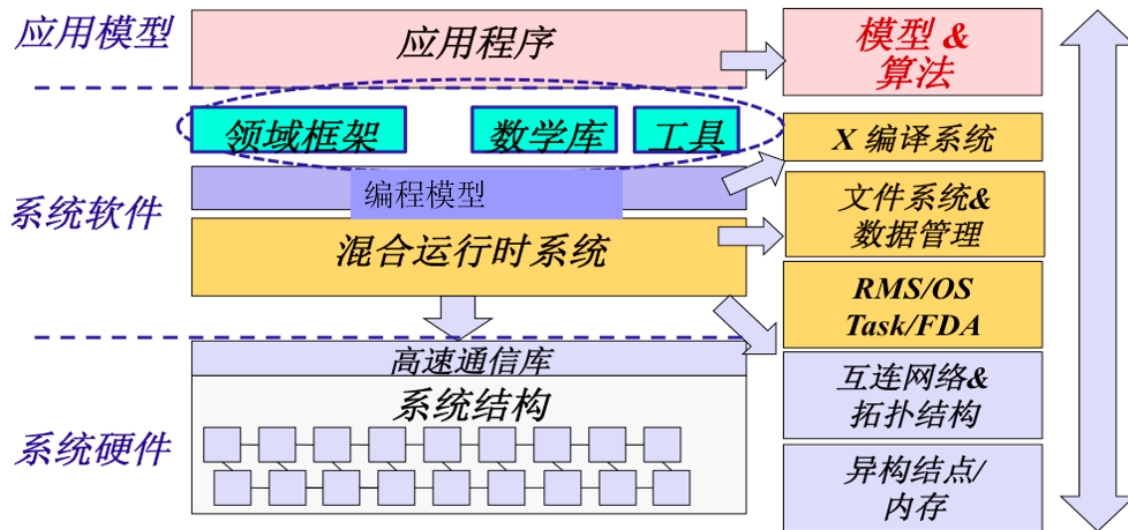
## 在线可视化



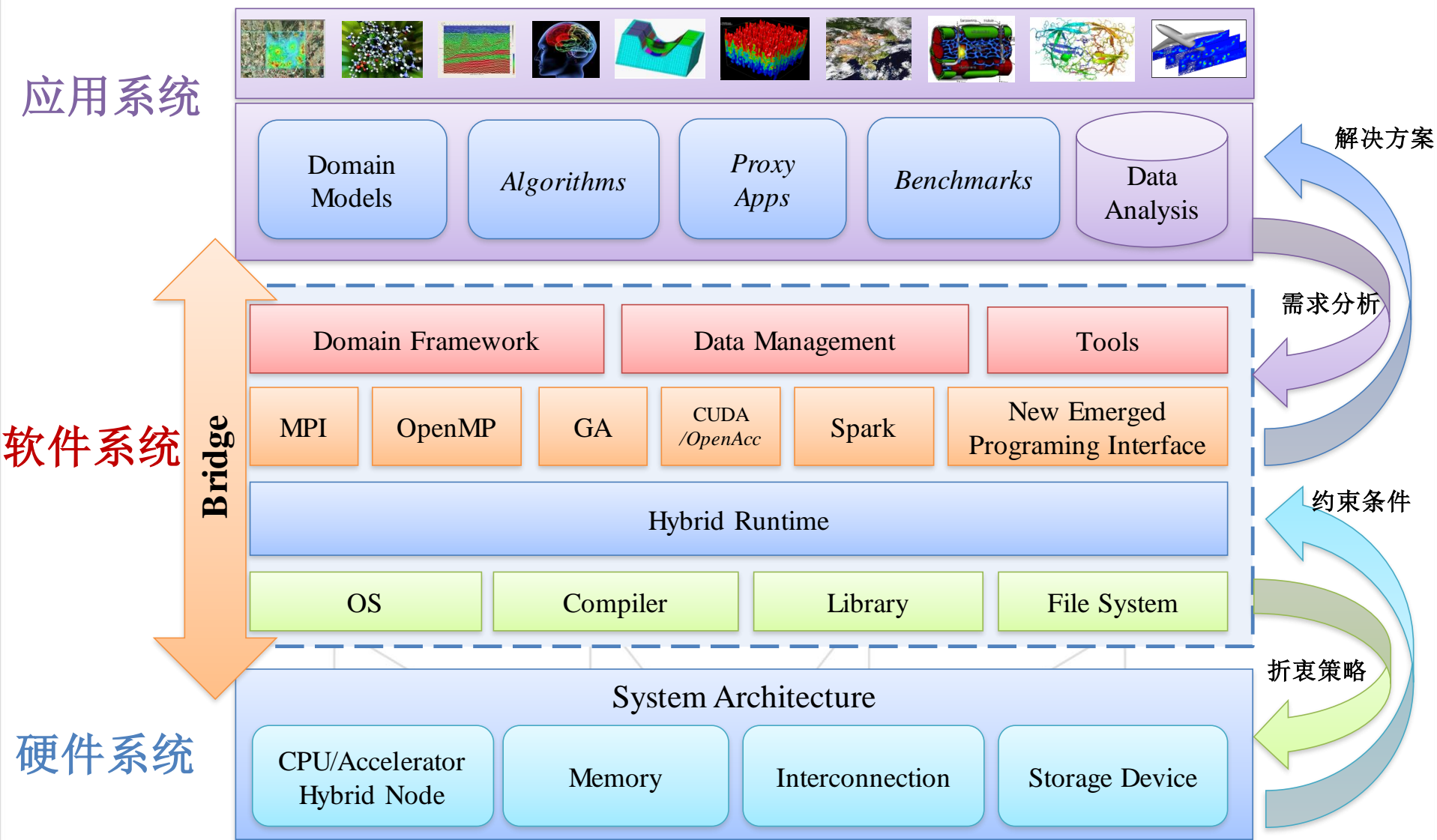
## ■ 大数据分析软件栈



# E级计算软硬一体化协同设计



## 应用与系统的协同设计

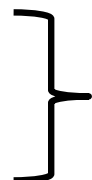




## 产学研用协同创新，实现两个转变

### ● 从以系统研制为主的发展模式，转变到

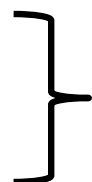
- 使能技术
- 系统技术
- 应用技术



相结合的一体化技术发展模式

### ● 从以科技开发为主的发展模式，转变到

- 科技开发
- 人才培养
- 学科建设

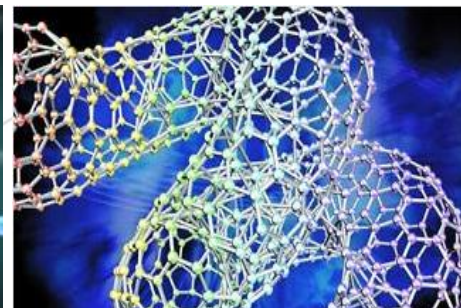
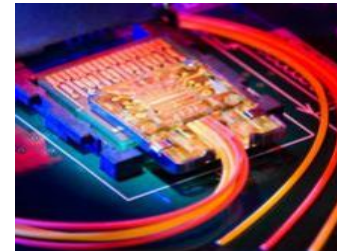
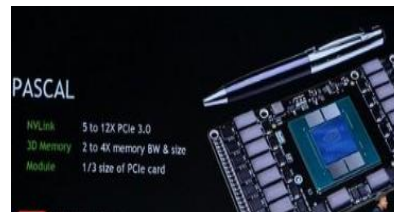


相结合的一体化能力发展模式



## 拥抱新技术，迎接新挑战

- 多核/众核处理器
- 新型存储介质
- 光互连网络技术
- 领域应用模型
- 软件定义系统
- 量子计算
- 纳米计算
- 生物计算



## 高性能计算与大数据处理融合发展

- 超算体系结构和存储架构协同设计
- 应用软件与系统软硬件协同设计
- 应用模型的多模式流程协同设计
- 超算中心局域与广域环境协同设计



## 大数据离不开超算

- 信息技术影响力
- 科学的影响力
- 工程的影响力
- 产业的影响力
- 文化的影响力
- 致力于超算系统结构技术
- 致力于超算应用模型方法
- 致力于超算应用软件平台
- 致力于超算交叉人才培养
- 致力于构建超算生态环境



# 谢 谢

开放、合作、创新、发展

