

# 高性能计算程序设计 基础

任课教师：黄聃（Huang, Dan）

# 教师信息

- 主讲教师： 黄聃
- 研究 方向：
  - HPC系统软件、内存计算框架、科学数据管理等
- 课程网站 <https://easyhpc.net/course/121/> （邀请码 1988）
- 联系方式
  - Email: [huangd79@mail.sysu.edu.cn](mailto:huangd79@mail.sysu.edu.cn)
  - QQ群： 772079175
- 助教：
  - 李冬升（课程、实验）， 张文毅（课程）
  - 梁嘉迪（实验）

# 课程简介

- 课程对象：没有接触过高性能计算的学生
- 课程目的：讲解高性能计算的基本软硬件和编程
- 预备知识：
  - 对计算机的组成有一定的概念（有修过计算机组成原理更好）
  - C、C++ 编程
  - 操作系统（Linux）
  - 线性代数
- 课程考核形式：笔试、编程作业（通过课程网站发布）

# 课程大纲

- 高性能计算的基本概念
- 高性能计算的发展历史
- 高性能计算的基本架构原理
- 高性能计算的编程方法
- 并行计算案例
- 并行程序优化方法
- 高性能计算应用案例

什么是超级计算机

# 什么是超级 (Super)

- 联想到什么？
- 什么是超级？
  - 汉语解释：特别;超出一般等级的。
  - 英文解释：especially; particularly.



# 计算机 vs 超级计算机



- 思域（Civic）民用版 vs 高性能版（Type R）

# 超级计算机 和 高性能计算

- 超级计算机（Supercomputer）
  - 超级计算机是当前具有最前沿处理能力的计算机。
  - A supercomputer is a computer at the frontline of current processing capacity.
  - 传统上，超级计算机主要用于解决大型的计算密集型科学问题，例如天气预报、气候变化研究、油气田勘探、分子动力学建模、物理模拟仿真等。
  - 现在和未来，超级计算机和大数据、人工智能紧密结合解决更多的计算密集型问题。
- 高性能计算（High Performance Computing）
  - 应用超级计算机解决普通计算机难以处理的大型计算问题。
  - "High-Performance Computing," or HPC, is the application of "supercomputers" to computational problems that are either too large for standard computers or would take too long.



# 衡量性能的指标是什么？

- Floating-point Operations Per Second (Flops)
  - 每秒钟进行的浮点计算操作次数
- 世界超算TOP500 排行榜
  - 采用统一的测试集HPL(High Performance Linpack)
  - Mega → Giga → Tera → Peta → Exa
  - $10^6 \rightarrow 10^9 \rightarrow 10^{12} \rightarrow 10^{15} \rightarrow 10^{18}$

# Top500 排行榜

- Top500--世界超算500强排行榜
  - 世界最快的计算机，公开数据，基于Linpack结果
  - 每年发布2次， 6月德国ISC， 11月美国SC
  - 1993年开始
  - <http://www.top500.org>
- 其他榜单
  - Green500, Graph500, HPCG.....



# Top500 排行榜

- Rmax :  
Maximal  
LINPACK  
performance  
achieved
- Rpeak:  
Theoretical  
peak  
performance

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	<b>Supercomputer Fugaku</b> - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,299,072	415,530.0	513,854.7	28,335
2	<b>Summit</b> - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148,600.0	200,794.9	10,096
3	<b>Sierra</b> - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
4	<b>Sunway TaihuLight</b> - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
5	<b>Tianhe-2A</b> - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, NUDT National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482

# 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 100-1000 m<sup>2</sup>
  - 能耗超级高 ~10 MW
  - 造价超级贵 ~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

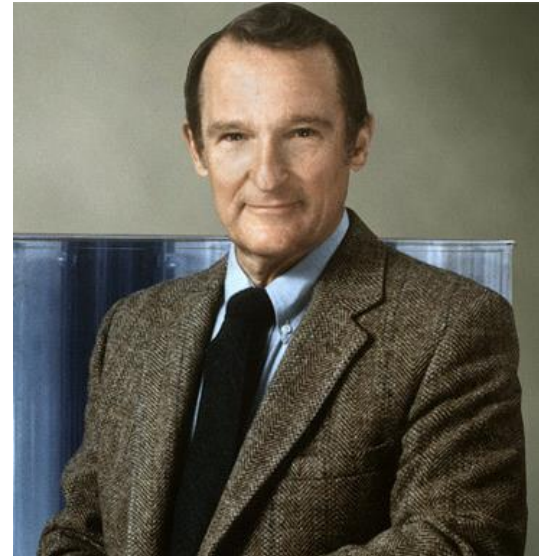
# 世界上第一台超级计算机

- CDC 6600, 1964年
  - 约3 MFlops
  - 性能是当时最快的计算机IBM 7030的10倍
  - 价格为800万美元



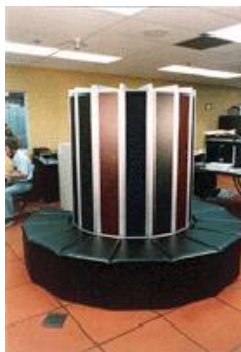
# 超算之父 – Seymour Cray

- 超算之父 1925~1996
- 创立了克雷公司（Cray），多次研制了世界上最快的超级计算机
- Cray公司研制了橡树岭实验室Titan超算，2013年被天河2号超越。



*"Anyone can build a fast CPU. The trick is to build a fast system."*

*"Computers should obey a square law — when the price doubles, you should get at least four times as much speed."*



Cray-1  
1976 160Mflops



Intel ASCI Red  
1997, 1Tflops



Cray Jaguar  
2009, 1.76Pflops



IBM Sequoia  
2012, 16.3Pflops



中国 太湖之光  
2016. 93Pflops



Cray-YMR  
1988, 2.3Gflops



日本地球模拟器  
2002, 40Tflops



中国NUDT TH-1A  
2010, 4.7Pflops



Cray Titan  
2012, 17.6Pflops



Summit  
2018, 148Pflops

40年系统性能增长5.8亿倍



CrayT3D  
1993, 19Gflops



IBM BlueGene/L  
2004, 70Tflops



日本 K computer  
2011, 8Pflops



中国NUDT TH-2A  
2013. 61.4Pflops



日本 Fugaku  
2020, 416Pflops



# 超算发展之路

- 1976, Cray-1
  - 向量处理技术, 80M
  - 射板耦合逻辑电路 (ECL)
  - 当时世界上最快的计算机
  - 133Mflops, 115KW
- 1982, CrayX-MP
  - 并行技术
  - 四个向量处理器
  - ECL器件, 200MHZ, 941MFlops
  - 液冷技术





# 超算发展之路

- 80年代末-90年代，百花齐放
  - Stanford DASH
  - SGI, Origin
  - IBM
  - Intel
- 1997, Intel ASCI Red

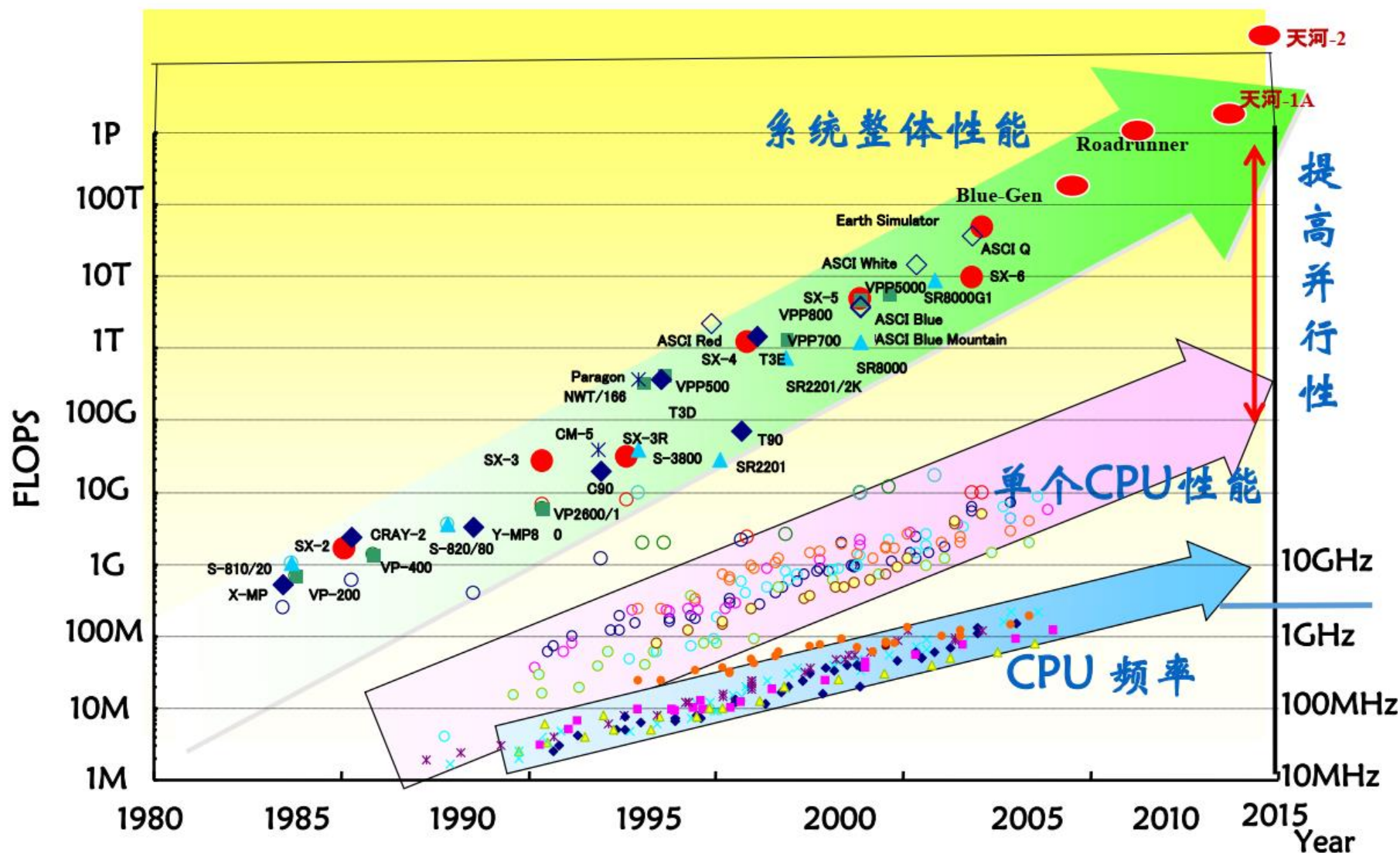


# 超算发展之路

- IBM Blue Gene
  - 2004
  - 嵌入式处理器
  - 100TFLOPS
  - 功耗仅为2.9MW
- 天河一号、二号、神威
  - 2010开始

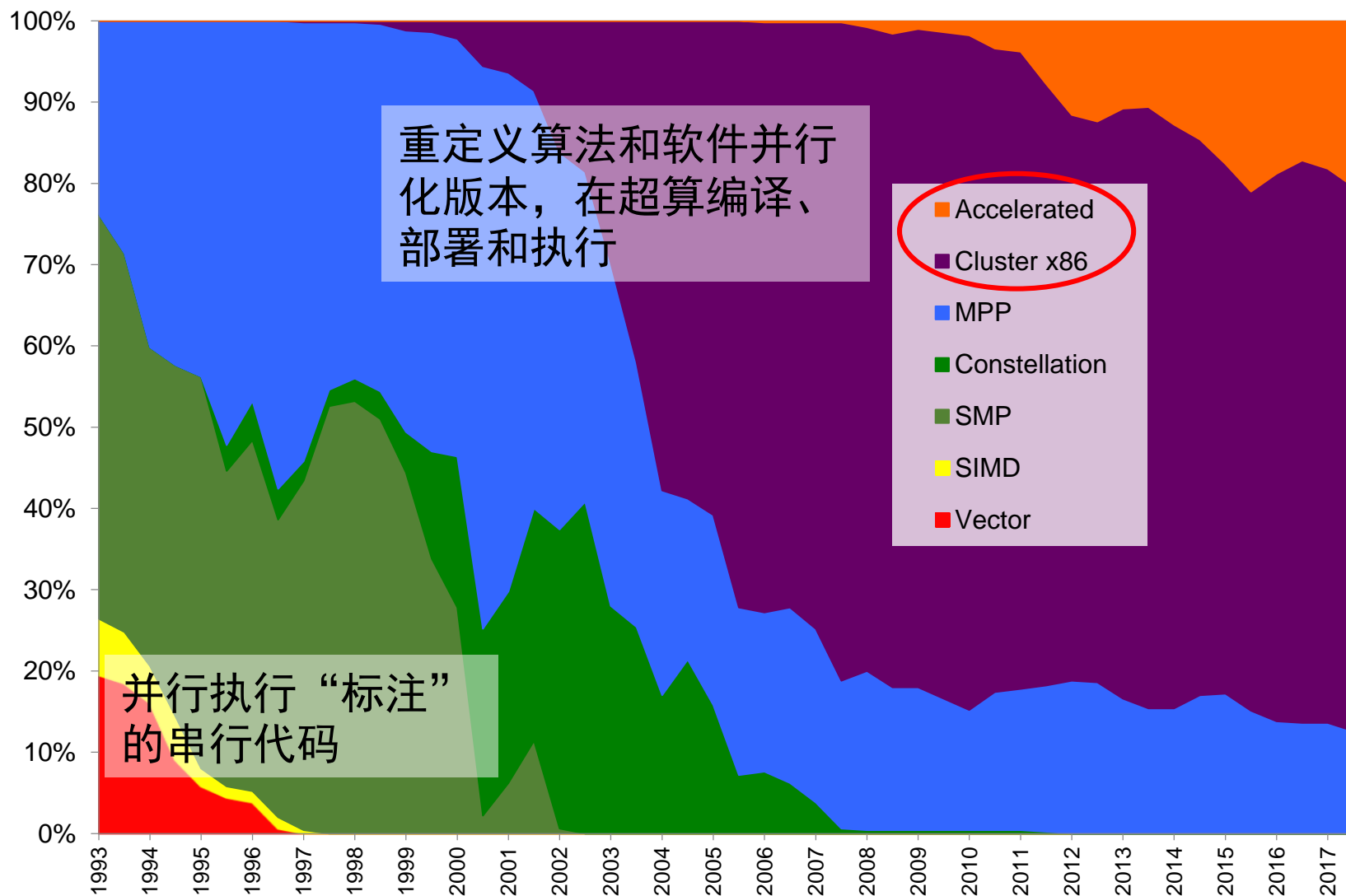


# 并行 通往超级计算的唯一途径



# 超级计算机演化

## 向量机→大规模并行加速器





# Summit (#2 超算)



## System Performance

- Peak performance of 200 petaflops for modeling & simulation
- Peak of 3.3 ExaOps for data analytics and artificial intelligence

## Each node has

- 2 IBM POWER9 processors, 44 cores
- **6 NVIDIA Tesla V100 GPUs**
- 608 GB of fast memory (512 GB DDR4 + 96 GB HBM2)
- 1.6 TB of NVMe memory

## The system includes

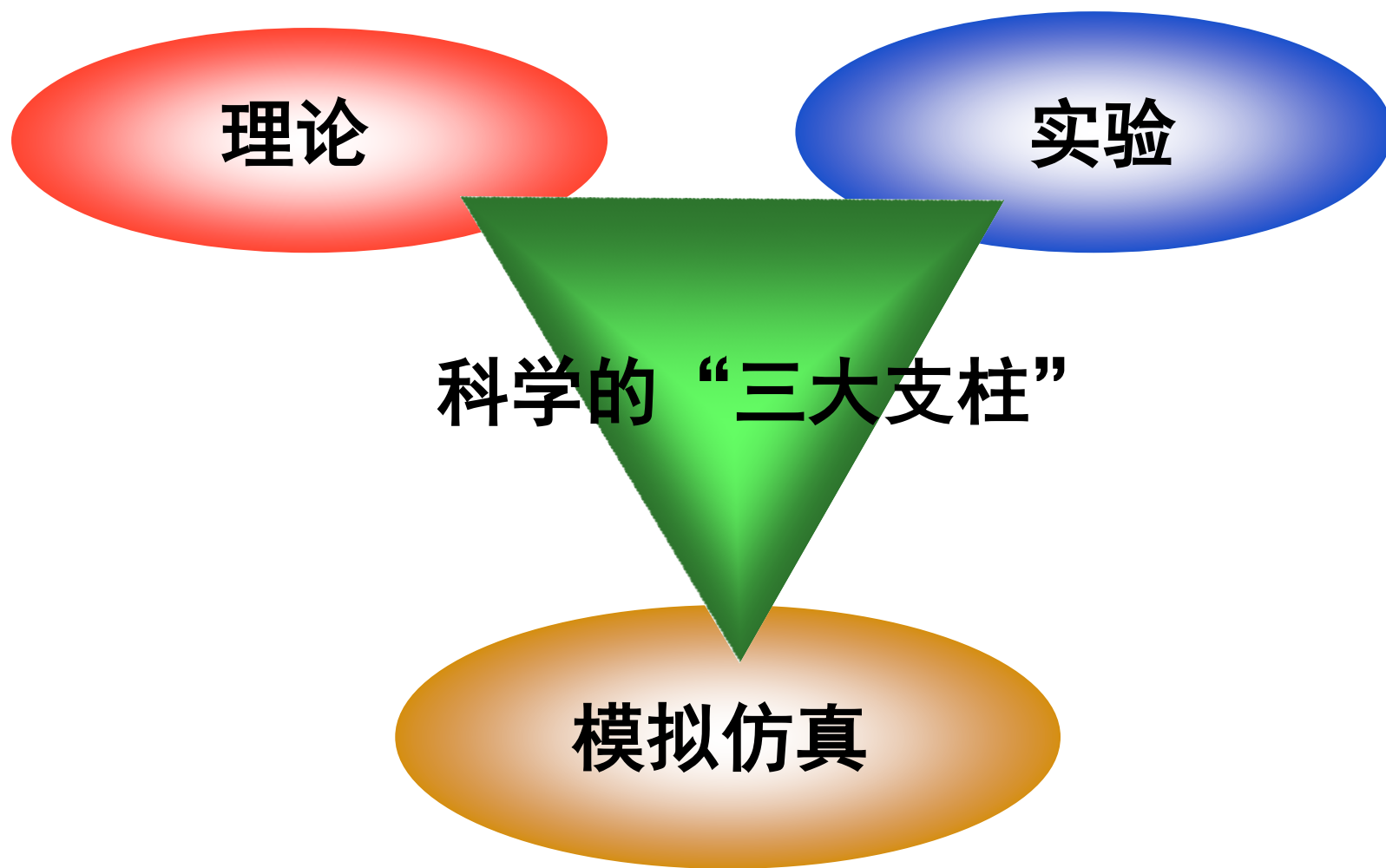
- 4608 nodes
- Dual-rail Mellanox EDR InfiniBand network
- 250 PB IBM Spectrum Scale file system transferring data at 2.5 TB/s

User <https://docs.olcf.ornl.gov/>



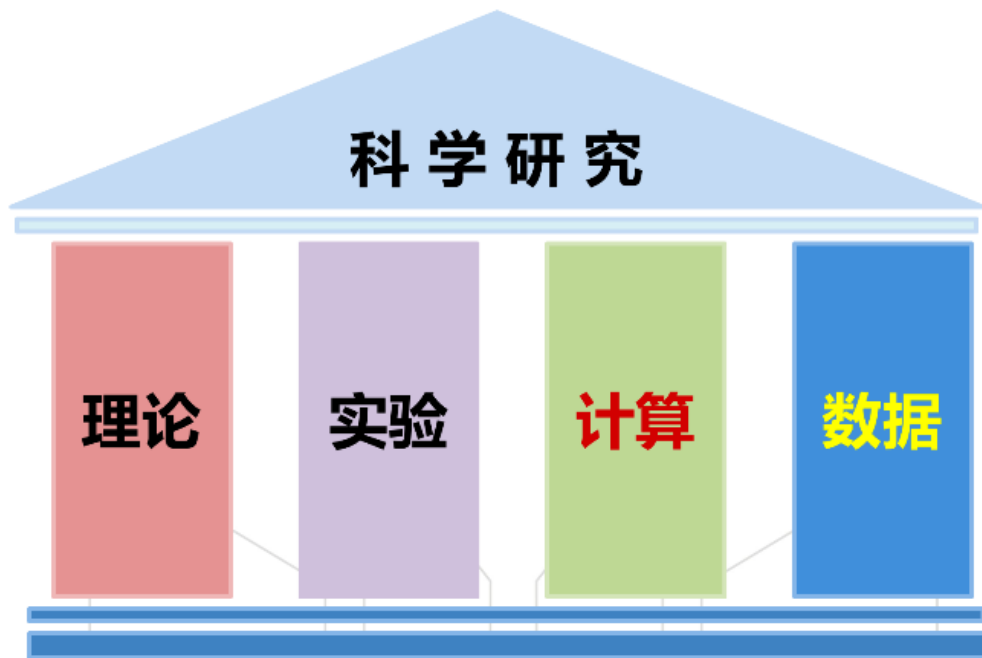
超级计算机能干什么？

# 超级计算机能干什么（传统）



# 超级计算机能干什么

- 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

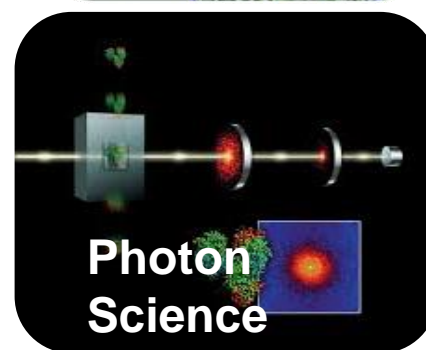
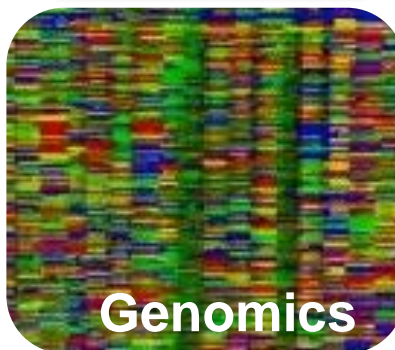
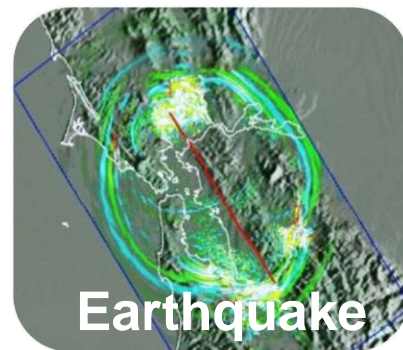
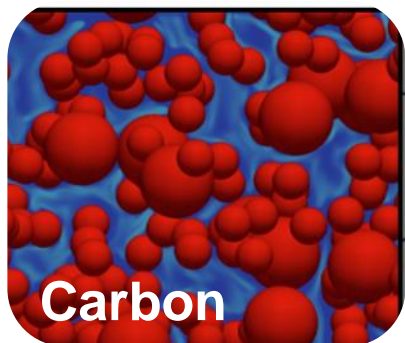
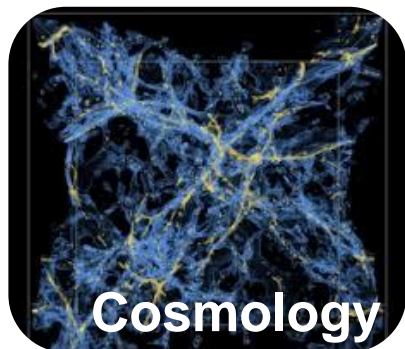
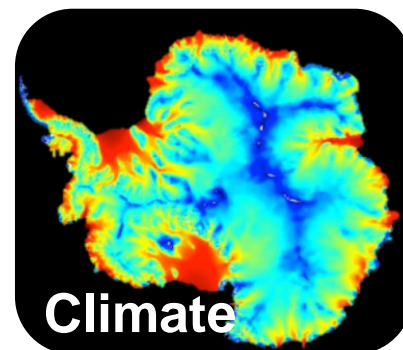
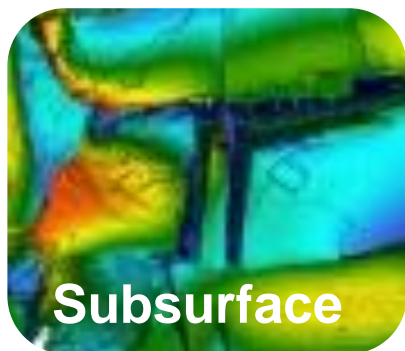




# 超级计算机能干什么

- 超级计算可以帮助人们解决一系列重要问题
  - 尺度超大 Too big
  - 尺度超小 Too small
  - 时变超快 Too fast
  - 时变超慢 Too slow
  - 过程超昂贵 Too expensive
  - 过程超危险 Too dangerous

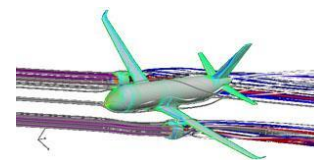
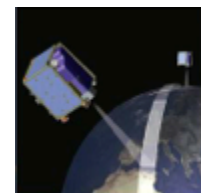
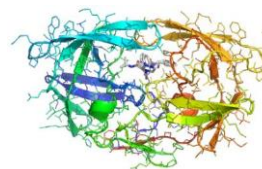
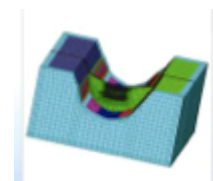
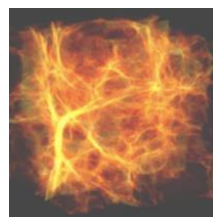
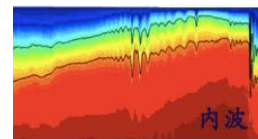
# 超级计算机能干什么



# 超级计算机的应用领域

- 气候环境
- 天文物理
- 新型能源
- 材料化工
- 航空航天
- 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



# 中国超算的发展之路

# 中国超算的发展之路

- 高性能计算是战略性、前沿性的高技术
  - 国家创新体系的重要组成部分
  - 产生原始创新和高端技术，影响下游产业
- 国际态势
- 美国长期保持领先优势
  - 顶级水平超级计算机，计划三大E级高性能计算基础设施
  - 应用面广，水平高，国家计划持续支持，如最近的ECP
- 日本多年紧随美国
  - Fugaku, K-Computer, HPCI日本国家计算基础设施，国家持续资
- 欧盟在基础理论、低功耗超级计算机、应用方面有优势
  - 低功耗E级计算机研究，欧洲PRACE高性能计算基础设施
  - 大规模仿真应用，欧盟协调主要成员国的力量

# 中国高性能计算发展战略计划

- 中国国家863计划的持续努力
  - 2002-2005年：高性能计算机及核心软件(863重大专项)
- 强调资源共享与协同工作，以网格支持多领域应用
- 成功研发10万亿次量级计算机和中国国家网格实验床
  - 2006-2010年：高效能计算机与应用服务环境(863重大项目)
- 性能结合开发效率、程序可移植性、系统的鲁棒性等
- 强调**机器、环境、应用**三位一体的发展，强调环境的服务特征
- 成功研发千万亿次量级计算机天河一号，建立国家高性能计算服务环境
  - 2010-2016年：高效能计算机及应用服务环境(863重大项目)
- 强调环境新的运行模式和机制，探索建立计算服务业的途径
- 发展应用社区，更好地支持应用
- 研制世界领先的计算系统：天河二号、神威·太湖之光

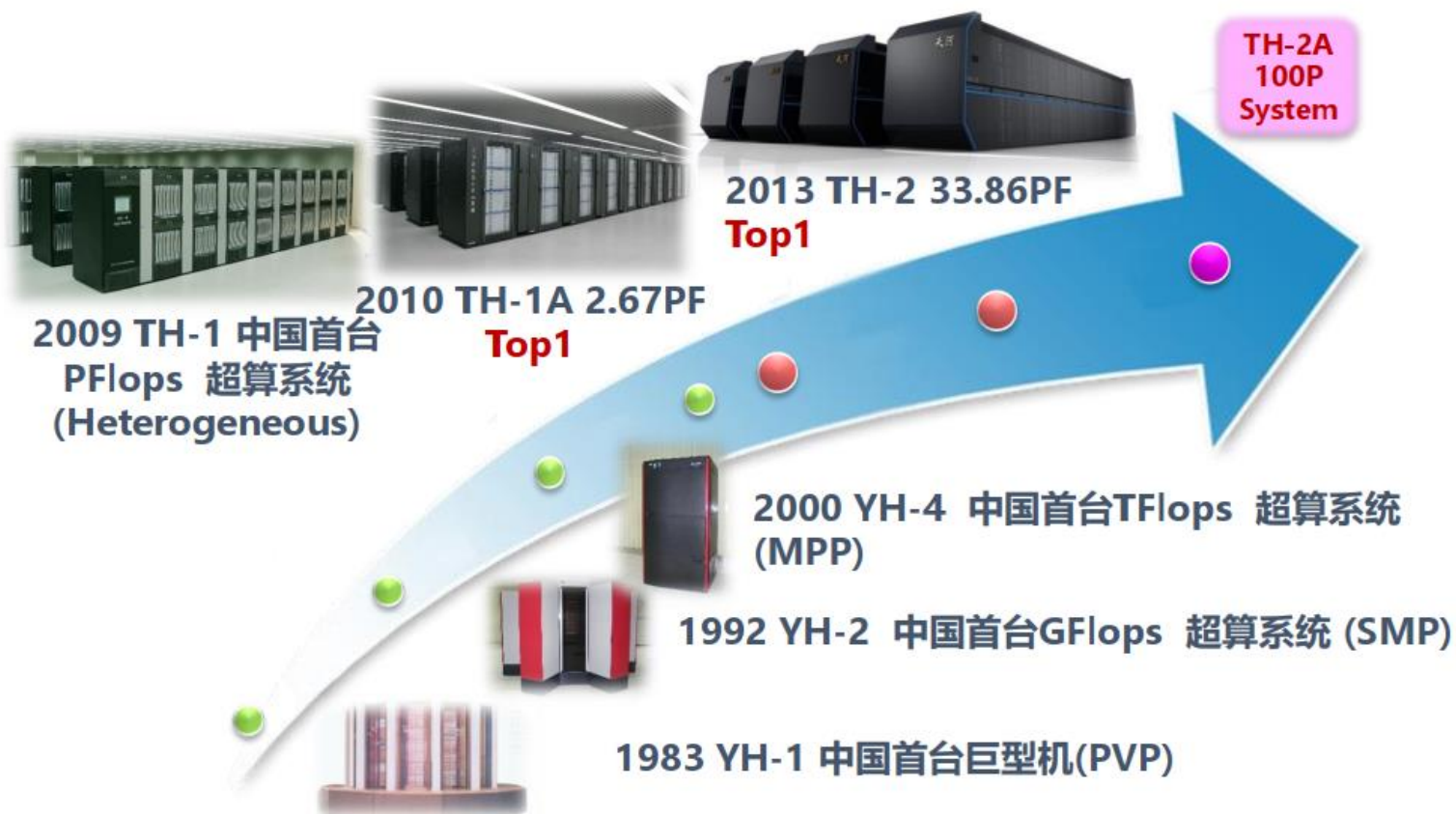
# 中国巨型计算机之父 - 慈云桂

- 慈云桂(1917-1990)，  
中国科学院院士，中国  
计算机界老前辈之一
- 主持研制成功我国首台  
亿次级巨型计算机“银河-I”





# 银河天河系列超级计算机





# 中国登顶TOP500的超算系统

- 首次登顶，天河一号A
  - 计算性能： 2.57 PFLOPS
  - 2010年11月
- 世界上唯一一台六连冠的超算系统，天河二号
  - 计算性能： 33.9 PFLOPS
  - 约 16,000个节点
  - 约 3,000,000个核
  - 约是个人电脑外星人 i7 4770K的36万倍



# 国家级的超算中心分布



NSCC-广州, 2013 天河2



NSCC-无锡, 2016 太湖之光



NSCC-长沙, 2012 天河1A



NSCC-济南, 2012 神威蓝光



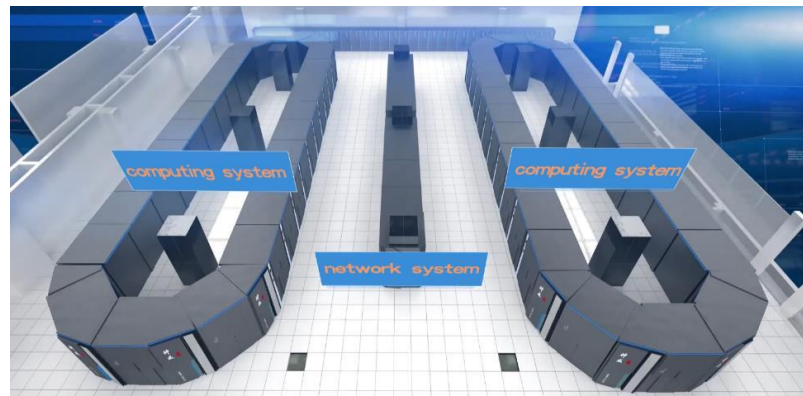
NSCC-天津, 2010 天河1A



NSCC-深圳, 2011 曙光 6000

# “神威太湖之光” 超级计算机

- 神威太湖之光
  - 采用国产申威众核处理器
  - 全系统4万个结点， 1000万核
  - 峰值性能12.5亿亿次 / 秒， Linpack 9.3亿亿次 / 秒
  - 能效比 6 GFlops/W
  - 装备国家超算无锡中心
- 世界超级计算机Top500排行榜 “三连冠”
  - 2016.6~2018.6

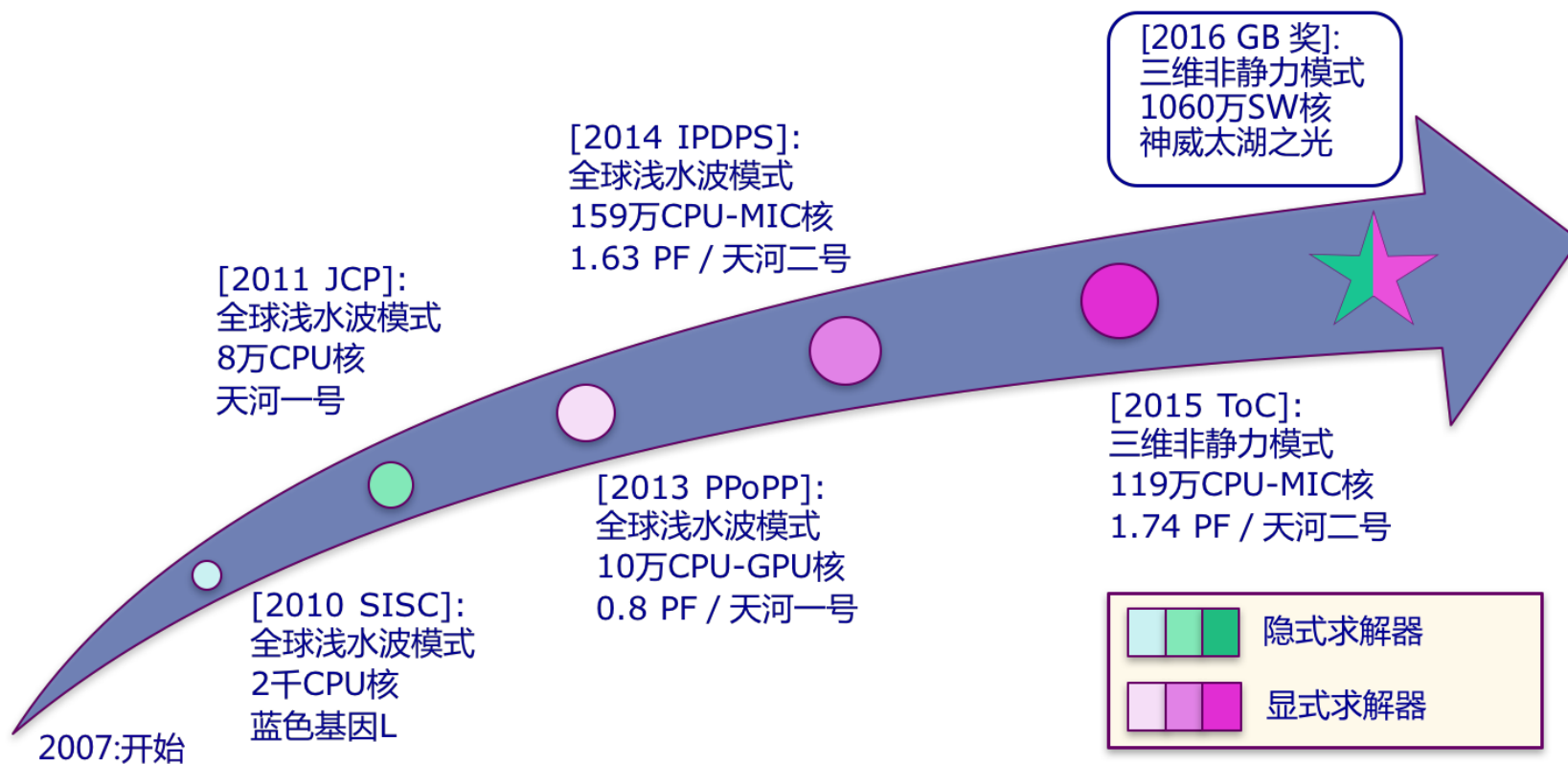


# 神威超级计算机的特点

- 神威众核处理器
  - 核心工作频率： 1.5 GHz
  - 双精度浮点峰值运算速度： 3.0 Tflops
  - 工艺： 28nm
  - 采用异构众核架构，芯片集成260个运算核心，集成存储控制器和网络接口
- 主核 (MPE)64-bit RISC core
  - 支持用户态和系统态
  - 256-bit 向量指令
  - 32 KB L1 instruction cache, and 32 KB L1 data cache
  - 256 KB L2 cache, 8x8 CPE mesh
- 从核(CPE)64-bit RISC core
  - 只支持用户态
  - 256-bit 向量指令
  - 16 KB L1 instruction cache, and a Scratch Pad Memory (SPM)

# 戈登·贝尔奖

- 颁发给高性能应用领域最杰出成就



# 中国超算发展的重要经验

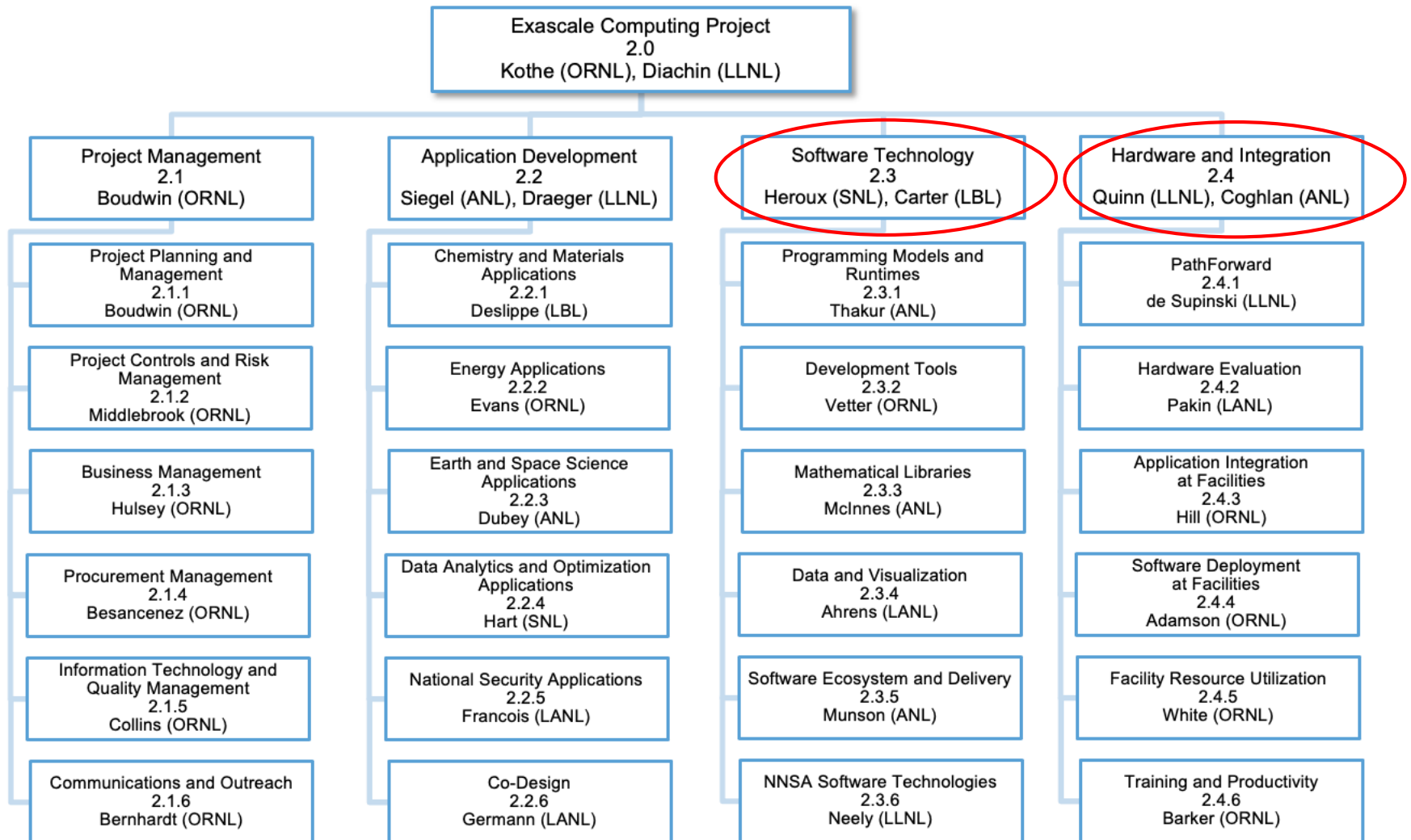
- 国家科技计划与地方、应用部门的发展计划相结合
  - 多渠道筹资研制高效能计算机
  - 科技部和地方政府共同支持建立国家超算中心
- 产学研用的结合
  - 高性能计算中心在高效能计算机研制中发挥作用
    - 遴选研制团队 / 提出系统指标
  - 企业参与国家科技计划，提高自身水平，促进研制工作
    - 浪潮、曙光、联想参与千万亿次和亿亿次计算机的研制
  - 应用部门牵头发展应用软件
- 机器、环境、应用三者均衡发展，相互促进

# 美国现阶段超算进展 (ECP)

- ECP: Exa-scale Computing Program
  - 2016—mid-2020s
  - 美国E级计算实施计划，美国能源部
  - ECP Enables US Revolutions in Technology Development:
    - 硬件设施与集成
      - 规划三大E级计算机：ANL (Aurora), ORNL (Frontier), LLNL (El Capitan)
      - 基于新型硬件的应用集成、硬件评测、基于新型硬件的软件部署
    - 软件系统
      - 规划建设面向异构系统的编程模型，编译环境，算法和数学库，快速开发和部署工具，大数据分析及可视化，计算和I/O框架等。
    - E级应用建设
      - 协同设计中心，物理和材料，能源，数据分析和优化，地球和空间科学，国家安全



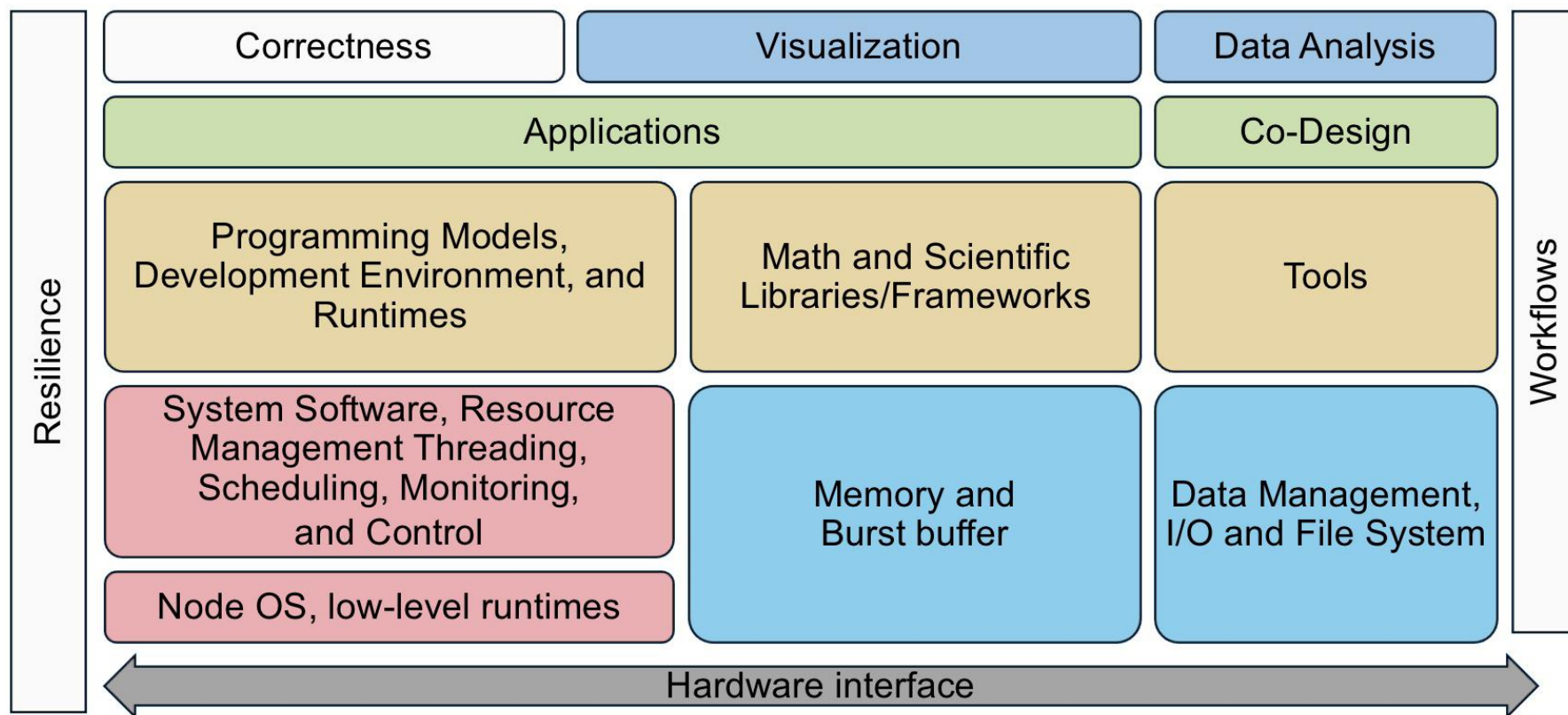
# 美国现阶段超算进展 (ECP)





# 美国现阶段超算进展 (ECP)

- ECP, Software Technology



# 小结

- 超级计算机是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
- 超级计算可以帮助人们解决一系列重要问题
  - 尺度超大 尺度超小 时变超快 时变超慢 过程超昂贵 过程超危险