

AMiner

超级计算机研究报告

Supercomputer Research Report

2018



目录

概述篇

技术篇

人才篇

市场篇

趋势篇

01

概述篇

超级计算机相关概念



- (1) 什么是超级计算机?
- (2) 超级计算机评价体系有哪些?
- (3) 超级计算机有何研究价值?

超级计算机发展脉络



- (1) 国防驱动阶段
- (2) 公司主导阶段
- (3) 蓬勃发展阶段
- (4) 多向发展阶段

典型超级计算机简介



- (1) Summit
- (2) 神威·太湖之光
- (3) Sierra

1.1 超级计算机相关概念

什么是超级计算机？

- 超级计算机又称高性能计算机、巨型计算机等，在计算速度、存储容量等方面有着普通计算机所不具备的超高性能。

超级计算机有何研究价值？

- 超级计算机在天气预测、污染检测与防控、石油气勘探与地震预测、工程仿真、动画效果渲染等领域已经创造了不可估量的价值。

定义

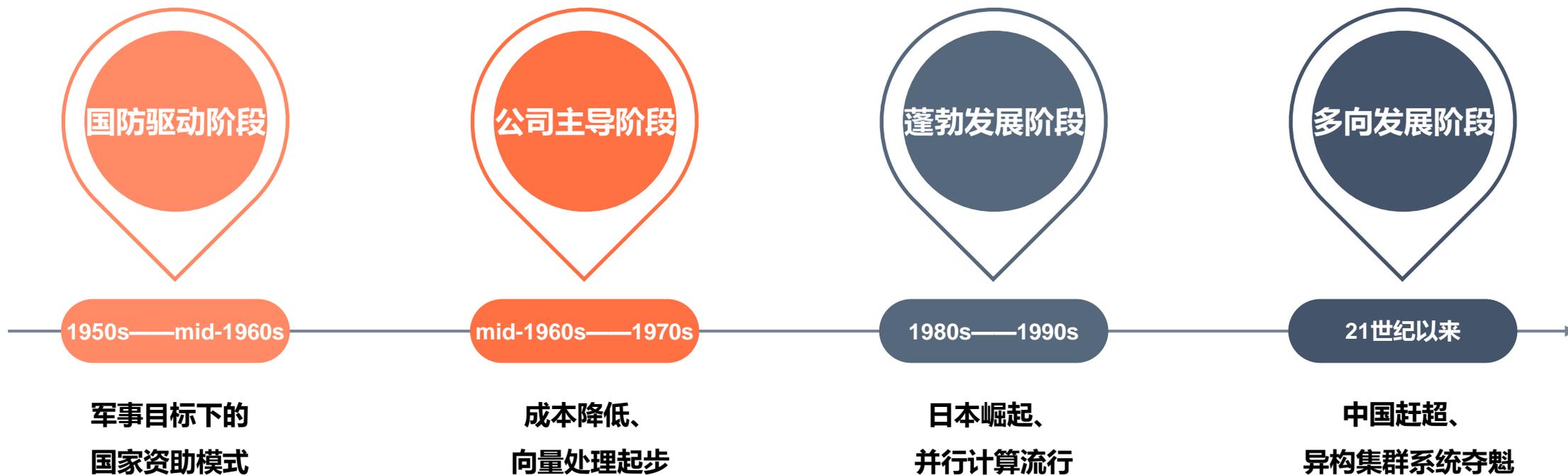
超级计算机评价体系有哪些？

- TOP500是业界公认的超级计算机性能排行榜。
- Green500是针对超级计算机能效的排行榜。
- 戈登·贝尔奖被认为是超级计算应用领域的诺贝尔奖。

现状

- 超级计算机当前以每秒钟浮点运算速度 (flops) 为主要衡量单位。
- 目前，超级计算机领域顶尖研究机构正在针对E级 (Eflops = 10^{18} flops) 系统的研发进行激烈竞争。

1.2 超级计算机发展脉络



1.3 典型超级计算机简介



Summit

Summit, 代号“OLCF-4”, 擅长人工智能、机器学习和深度学习方面的运算。



神威·太湖之光

神威·太湖之光是中国大陆首个不使用英特尔等美国公司的核心产品而登上TOP500第一名宝座的超级计算机。



Sierra

Sierra, 代号ATS-2, 是IBM为美国能源部下属的劳伦斯利福摩尔国家实验室建造的超级计算机。

02

技术篇

基础层：以异构并行为基础的超级计算机组成



- (1) 基本原理
- (2) 架构分类
- (3) 最新发展

中间层：六类设备+三大网络



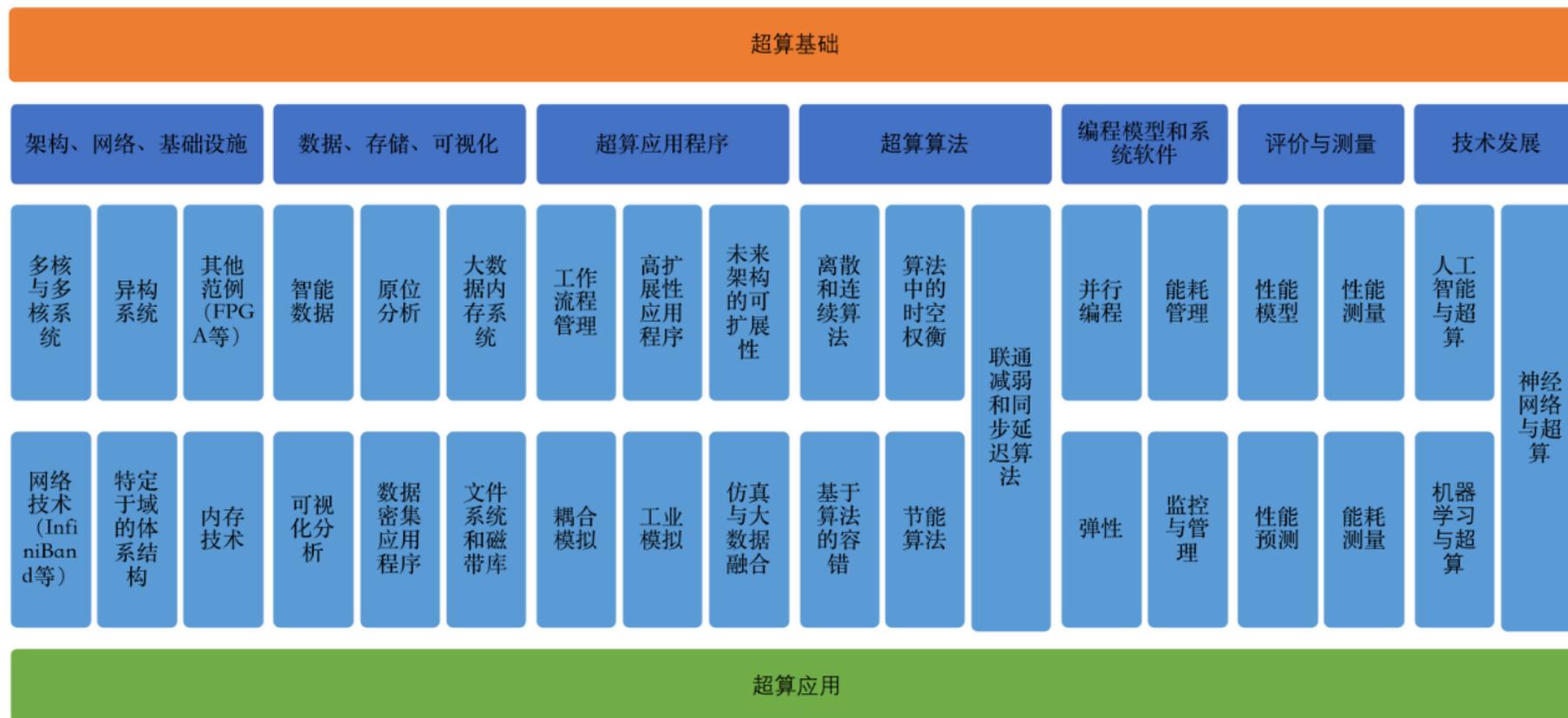
- (1) 六类设备
- (2) 三类网络
- (3) 基本原理

应用层：解决方案



- (1) 石油气勘探
- (2) 生物医药与智能医疗
- (3) 工程仿真与航天器研发
- (4) 天气预报与雾霾预警
- (5) 海洋环境工程
- (6) 建筑信息模型
- (7) 基础科学研究

超级计算机技术分层



超级计算机技术有三个层次：基础层、中间层和应用层。其中，基础层主要应用层涉及超级计算机使用场景，而中间层则包含了数据存储、计算、管理等多重技术支持，正是有了中间层的链接，超级计算机原理才能落实到应用问题解决之中。

2.1 基础层：以异构并行为基础的超级计算机组成

基本原理

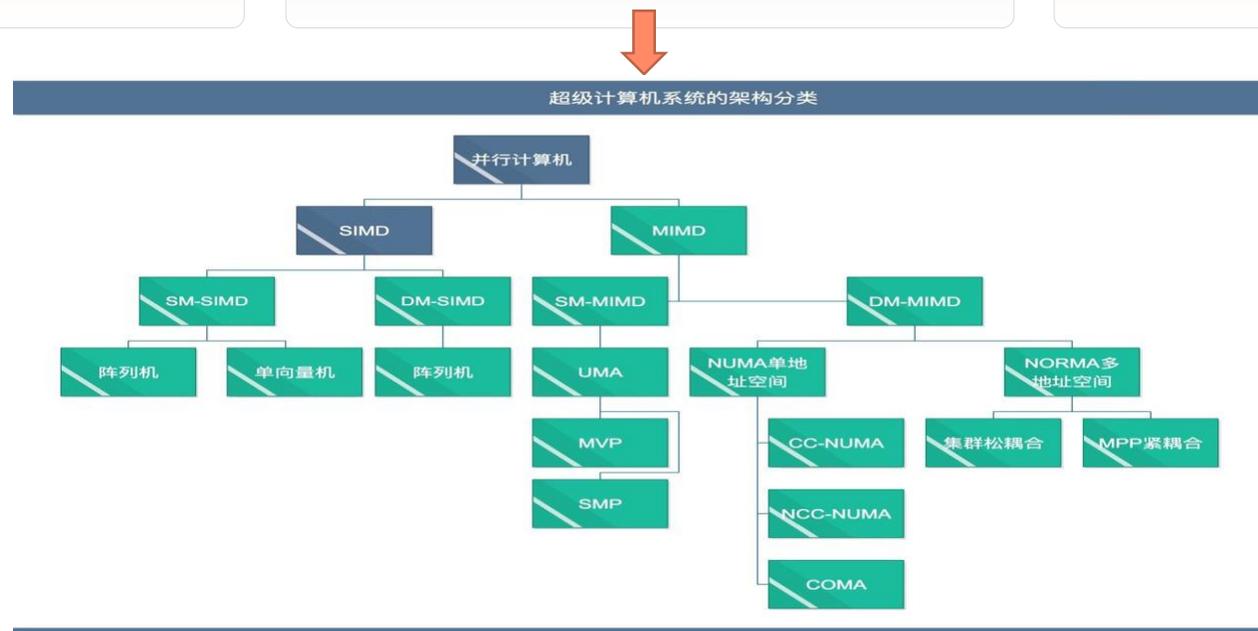
- 超级计算的基本原理是并行计算。
- 超级计算的性能优化之一是提高并行可扩展性。目前来看，硬件层面并行化的实现手段为：多重执行单元等。

架构分类

- 当今的超级计算机系统大多以MIMD方式工作。
- 按照并行计算方式存储器是共享还是分布，可以将超级计算机系统的架构作如下划分：

最新发展

- 自1996年以来，由于挑战性应用问题的急切需求，加快了计算机系统结构的演变和发展进程。
- 进入21世纪之后，采用专用处理器或者附属加速处理器的方式加以实现。



2.2 中间层：六类设备+三大网络

当前HPC的主要架构包括Cluster集群和MPP（大规模并行处理）两种，典型的HPC集群系统主要由五类计算（或网络）设备和三类网络组成。



2.2 中间层：六类设备+三大网络

管理网络



管理网络，用于管理节点和各个计算节点、I/O节点的互联。

计算网络



计算网络，用于各计算节点之间的互联，是并行任务执行时进程间通信的专用网络。

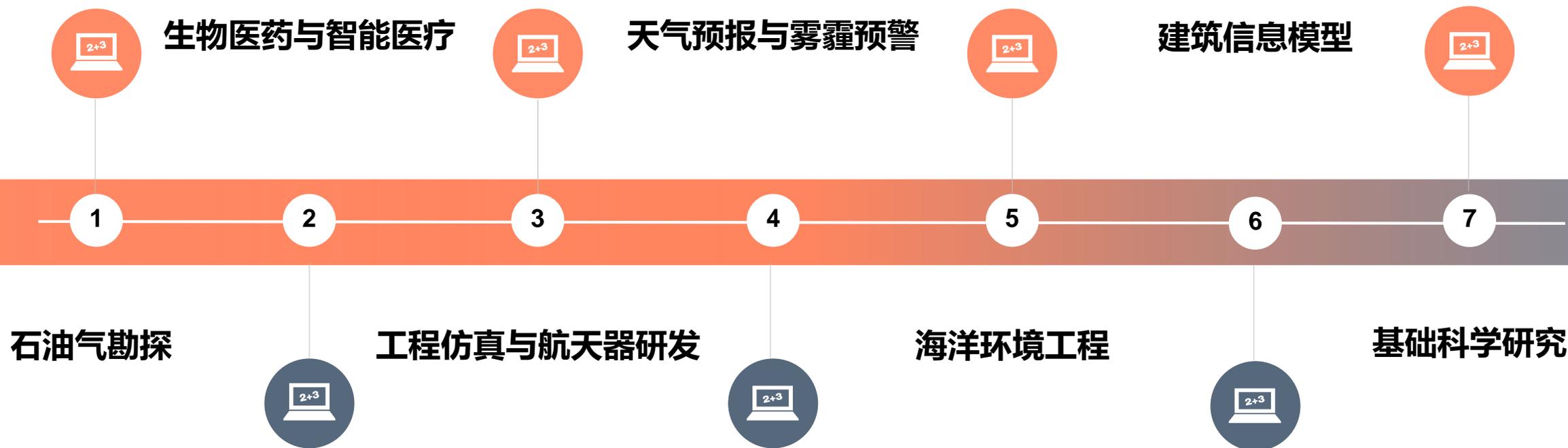
存储网络



存储网络，需要向高性能计算机的节点提供数据访问服务。

三大网络

2.3 应用层：解决方案



2.3 应用层：解决方案

石油气勘探

长期以来，油气地球物理勘探技术的发展与应用高度依赖于包括高性能计算技术在内的信息技术的发展。

- 基于“天河一号”开展的石油勘探数据处理程序，实现了复杂地质条件下上千平方公里数据的逆时偏移处理。
- 解决方案与服务内容包括：①测试服务；②偏移处理服务，③应用并行优化开发服务。

生物医药与智能医疗

依托高性能计算、云计算、大数据及人工智能等技术的高度发展，使国民健康行业走向真正意义上的智能化。

生物医药 (Biological Medicine)

依托高性能计算相关科学研究和项目合作，可以整体提高行业和企业竞争力。

智能医疗 (Intelligent Medical)

通过打造医疗信息平台等，结合大数据等关键技术，提高临床疾病的诊断效率和精度。

案例：“天河一号”支持华大基因开展大规模生物基因处理及数据存储：

①开发了基于GPU的高效基因测序处理软件，并利用该软件进行了3000株水稻的基因组重测序分析

2.3 应用层：解决方案

工程仿真与航天器研发

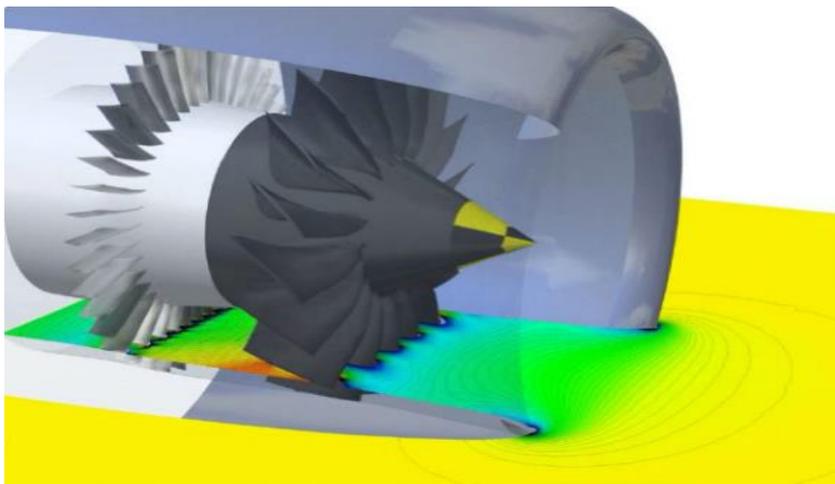
工程仿真CAE利用计算机辅助求解力学性能的分析计算结果。其核心思想是结构离散化，得出满足工程精度的近似结果。

基础科学研究

人们可以模拟越来越大规模的微观系统，从而极大地增强了对自然的认知能力。用来模拟微观世界中原子和分子的相互作用和行为。

案例

案例：在航空发动机研制中，气动稳定性是最重要的技术标准之一。



某型航空发动机内部三维流动工程仿真效果图

案例：

- ①使用并行程序进行密度泛函理论（DFT）计算已经成为材料科学、固体物理、计算化学、计算生物学等领域内必不可少的研究手段之一；
- ②并行实现的高精度耦合簇理论（CC）和组态相互作用（CI）方法被许多量子化学计算程序采用，成为计算化学的主要工具；

2.3 应用层：解决方案

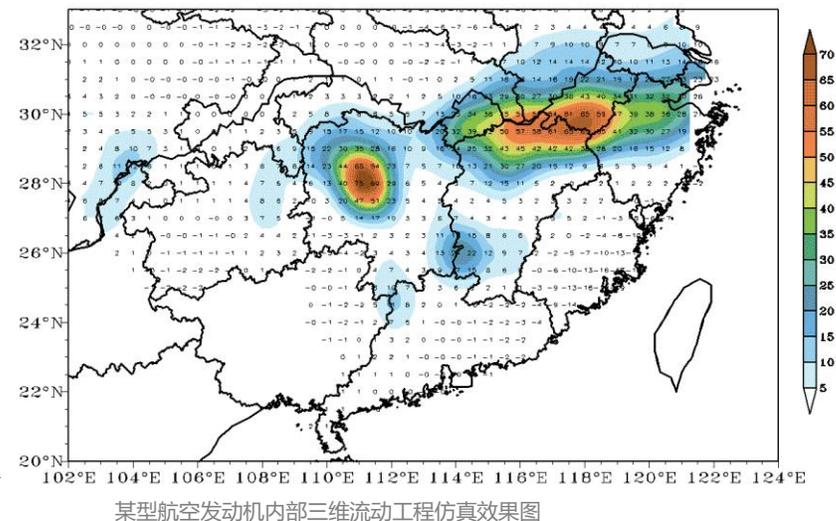
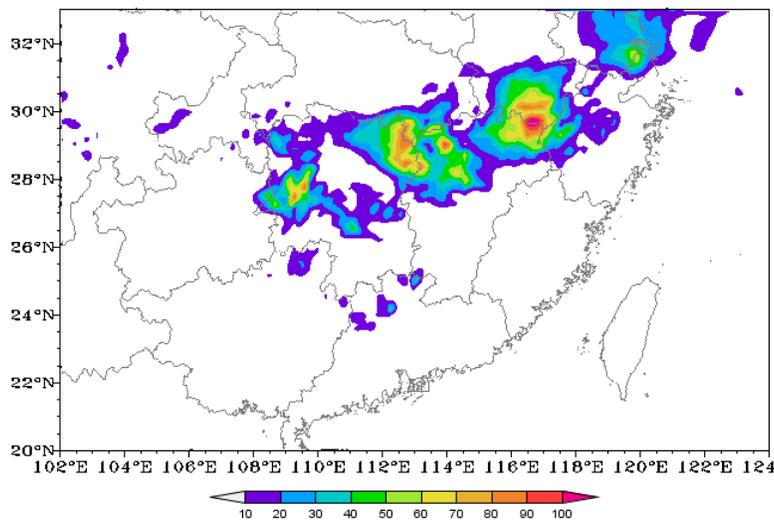
天气预报与雾霾预警

数值天气预报将气象数据和边界参数导入方程求解，从而预测大气变化和状态的科学。

- 业务流程大致为：气象数据收集和预处理、数值天气预报流程、综合数值天气预报、天气学与统计学等输出预报结果。它是典型的计算密集型应用（Computing-Intensive），

案例

案例：国家超级计算长沙中心为中部某省气象局提供了数值天气预报计算的平台支持，以提高天气预报、气候预测的及时性、准确性、可靠性和精细化。



某型航空发动机内部三维流动工程仿真效果图

2.3 应用层：解决方案

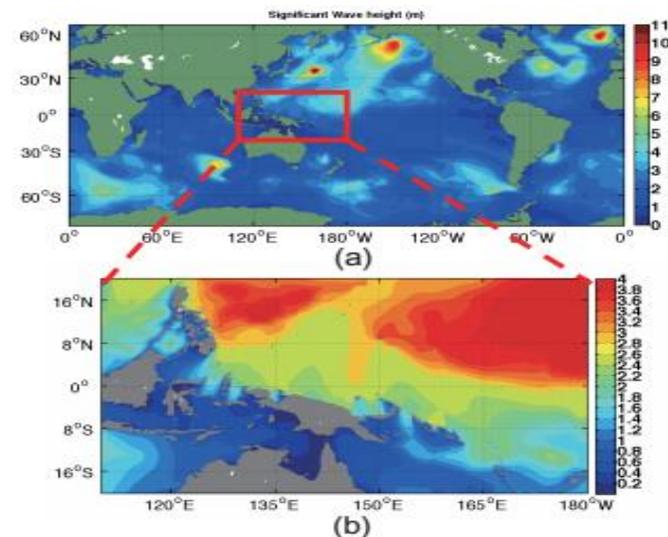
海洋环境工程

从海洋环流数值模拟到空气质量实时监测，再到海洋灾害预报等，实现海洋环境数值预报的精确性。

- 广州市香港科大霍英东研究院建立的CMOMS数值模拟系统，为应对中国海域及其附近海域的气候变化提供参考与支持。

案例

案例：“神威·太湖之光”超级计算机实现了MASNUM海浪数值模拟的 $(1/60)^\circ$ 高分辨率的全球海洋模式，模式成功扩展到8,519,680核数，达到最高30.07Pflops的峰值性能。



利用“神威·太湖之光”模拟的全球 (a) 与区域 (b) 重要海浪高度分布

2.3 应用层：解决方案

建筑信息模型

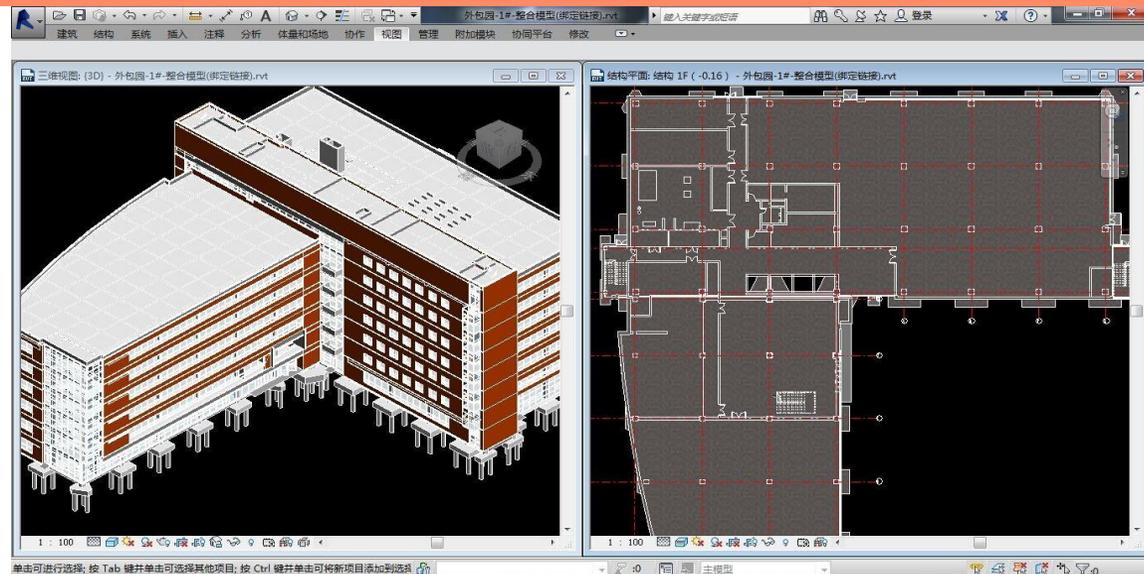
建筑信息模型（Building Information Modeling，简称BIM）是一项将建筑与信息相结合的综合技术。

- 通过高性能计算机系统对建筑过程中产生的主要数据进行存储处理，提高工程效率与质量，为推动智慧城市的建设与发展起到重要作用。

案例

案例：天河BIM云平台以工程仿真系统、大数据平台和云平台为依托，为产业园建设提供四类服务：

- ①BIM远程可视化建模平台，即可进行BIM远程可视化建模。
- ②BIM协同设计服务，实现不同参与方、不同专业的多地点、实时协同工作。



远程可视化建模

学者概况

学者概况

国外学者

国内学者

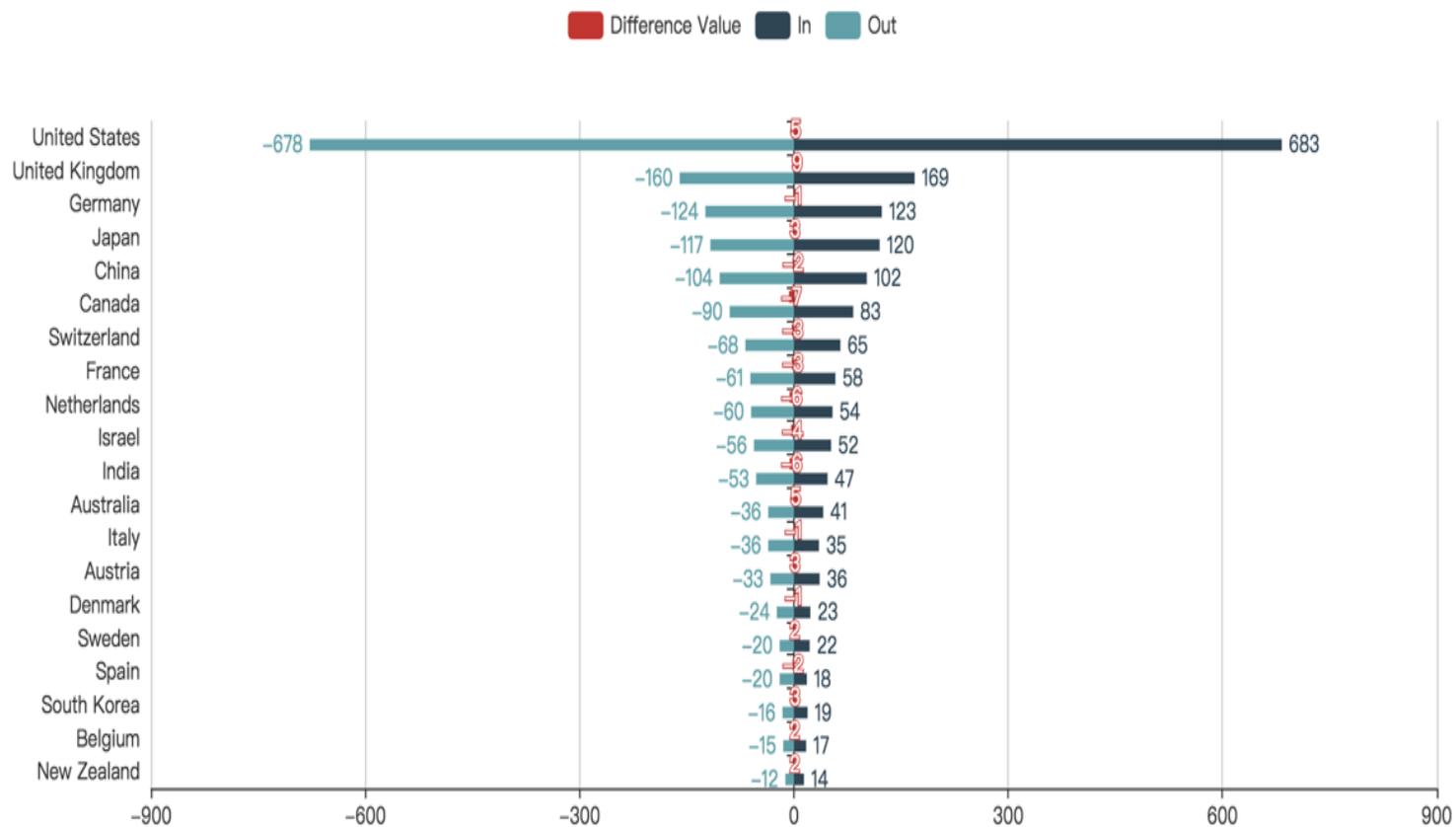




AMiner以“supercomputer”为关键词，对超级计算领域全球TOP1000的学者进行分析。

由上图可得：美国人才最多、最为集中，中国紧随其后，英国等也有不少超级计算机人才。这与国家超级计算机发展水平呈正相关。

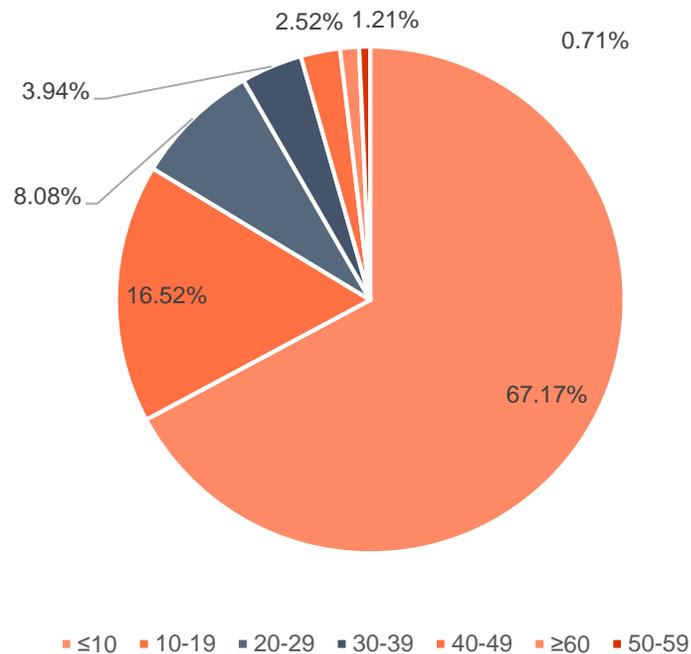
全球学者迁徙图



AMiner选取超级计算机领域排名TOP1000的专家学者的迁徙做了分析。

由上图可得，美国是超级计算机领域人才流动大国。

英国、中国等国人才流动量落后于美国，英国和中国的人才流入量大于流出量，德国和加拿大有轻微的人才流失现象。

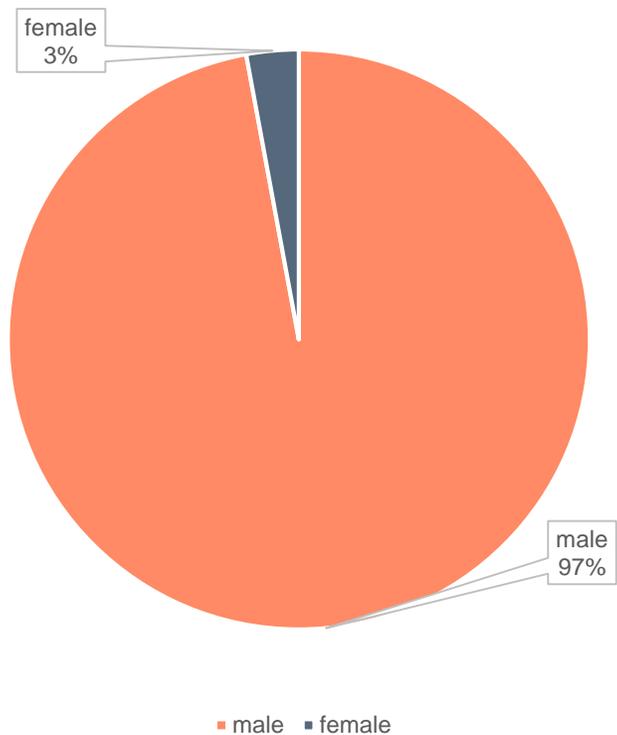


全球超级计算机领域TOP学者h-index指数 ≤ 10 的人数最多，占比67.17%，

h-index指数在10-19之间的学者人数次之，占比16.52%，

h-index指数 ≥ 60 和在50-59之间的学者人数最少，前者占比1.21%，后者占比0.71%。

全球学者男女比例图



全球超级计算机TOP学者男性占97%，女性占3%，男女人数相差悬殊。

中国学者分布图



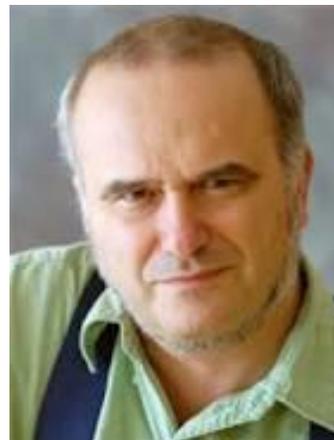
由全国学者分布图放大后，绘制了该领域学者的中国分布地图。由图可知：超级计算机学者在中国集中于北京及广州等地方。

**Jack Dongarra**

美国田纳西大学电子工程和计算机科学系的计算机科学杰出教授。在曼彻斯特大学的计算机科学和数学学院拥有图灵奖学金。

**Ian T. Foster**

新西兰裔美国人计算机科学家。数据科学与学习部门的主任，也是芝加哥大学计算机科学系的教授。

**Marc Snir**

以色列美国计算机科学家。擅长计算复杂性、并行算法、并行体系结构、互连网络等细分领域。

**Bronis R. De Supinski**

劳伦斯利弗莫尔国家实验室项目的首席技术官，主要研究包括数据挖掘技术及其在性能分析和建模上的应用

**Satoshi Matsuoka**

日本东京工业大学教授，研究大型超级计算机和类似基础设施的系统软件。曾获得“戈登·贝尔”奖

**杨广文**

清华大学博导，高性能计算技术研究所所长、计算机系学术委员会副主任。他擅长分布式信息服务与集成技术等，曾获“戈登·贝尔”奖。

**廖湘科**

现任国防科技大学计算机学院院长、天河二号总设计师。长期从事高性能计算机系统软件与通用操作系统的科研工作。

**钱德沛**

北京航空航天大学教授，中山大学数据科学与计算机学院院长、教授。研究高性能计算机体系结构、分布式系统等。

**王恩东**

中国工程院院士，高效能服务器和存储技术国家重点实验室主任。在开放计算领域都达到了国际领先水平。

**杨超**

中国科学院软件研究所研究员，荣获“戈登·贝尔”奖。在大气动力学模拟的全隐式求解器算法和软件研究方面辛勤耕耘。

04

市场篇

各国排名进入TOP500的计算机数量情况

超级计算机TOP500前十的情况

国际高性能计算大会榜单

各国排名进入TOP500的计算机数量情况（2018年6月）

排名	国家和地区	排名进入TOP500的计算机数量	所占百分比	处理器总数
1	中国	206	41.2%	25,107,680
2	美国	124	24.8%	15,829,720
3	日本	36	7.2%	7,170,408
4	英国	22	4.4%	1,678,220
5	德国	21	4.2%	1,422,814
6	法国	18	3.6%	1,816,720
7	荷兰	9	1.8%	424,800
8	韩国	7	1.4%	843,300
9	爱尔兰	7	1.4%	334,824

由榜单可得，中国已成为全球拥有最多超级计算机的国家，全球TOP500的超级计算机中，有206台是中国研制，美国则拥有124台。

TOP500前十排名

2018年6月超级计算机TOP500前十

排名	名称	国家	场所	供应商
1	Summit	美国	橡树岭国家实验室	国际商业机器公司
2	神威·太湖之光	中国	国家超级计算无锡中心	国家并行计算机工程技术研究中心
3	Sierra	美国	劳伦斯利福摩尔国家实验室	国际商业机器公司
4	天河-2A	中国	国家超级计算广州中心	中国人民解放军国防科学技术大学
5	ABCI	日本	国家先进工业科学技术研究所	富士通株式会社
6	代恩特峰	瑞士	瑞士国家超级计算中心	克雷公司
7	泰坦	美国	橡树岭国家实验室	克雷公司
8	红杉	美国	劳伦斯利福摩尔国家实验室	国际商业机器公司
9	Trinity	美国	洛斯阿拉莫斯国家实验室	克雷公司
10	Cori	美国	国家能源研究科学计算中心	国际商业机器公司

由图可知，除了排名第二的神威·太湖之光，前五名中的其余四个都是新建或是大幅升级过的系统。

05

趋势篇

速度大幅提升

当前，世界TOP500的超级计算机，基本都可以达到P级

(1Pflot= 10^{15} flops) 运算速度。

与AI、VP融合

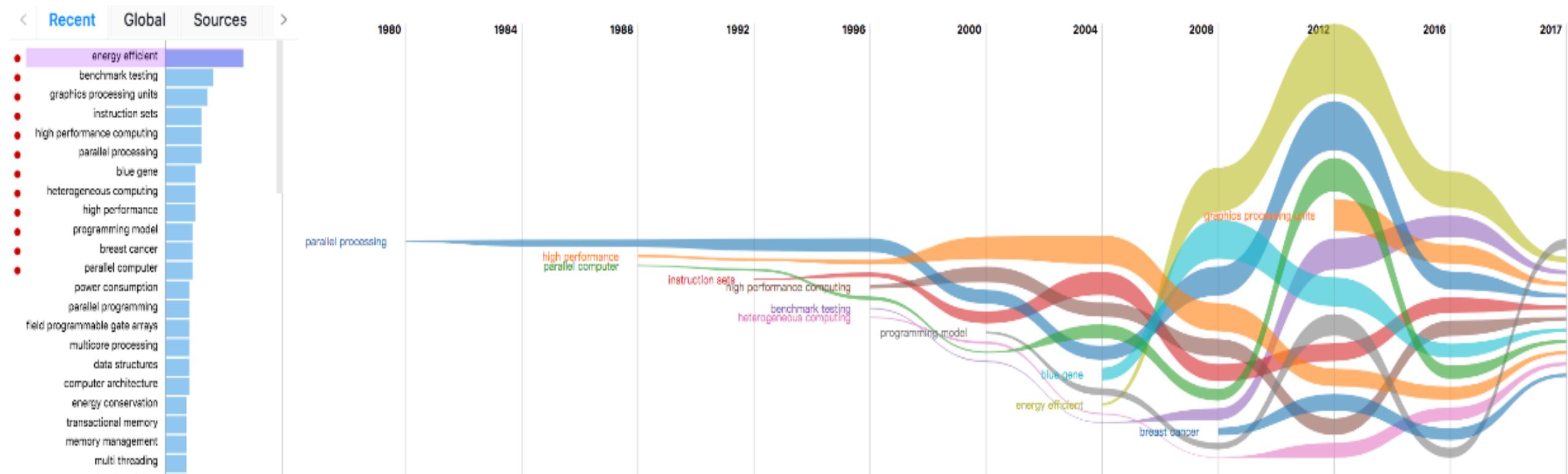
预计2020年全球数据量将超过4万亿GB，这些海量的大数据通过最新的深度学习技术将为人类社会创造难以估量的价值。

量子计算机

2017年年底，IBM宣布全球首台50个量子比特的量子计算原型机，被媒体认为最快超级计算机“神威·太湖之光”。

生物计算机

生物计算机也称仿生计算机，主要原材料是生物工程技术产生的蛋白质分子，能模仿人脑的机制等。



AMiner通过挖掘1980-2017年间超级计算机领域论文信息，计算出超级计算机近期热点。图中每个彩色分支表示一个关键词领域，其宽度表示该关键词的研究热度，各关键词在每一年份（纵轴）的位置是按照这一时间点上所有关键词的热度高低进行排序。

版权声明

AMiner研究报告版权为AMiner团队独家所有，拥有唯一著作权。AMiner咨询产品是AMiner团队的研究与统计成果，其性质是供用户内部参考的资料。

AMiner研究报告提供给订阅用户使用，仅限于用户内部使用。未获得AMiner团队授权，任何人和单位不得以任何方式在任何媒体上（包括互联网）公开发布、复制，且不得以任何方式将研究报告的内容提供给其他单位或个人使用。如引用、刊发，需注明出处为“AMiner.org”，且不得对本报告进行有悖原意的删节与修改。

AMiner研究报告是基于AMiner团队及其研究员认可的研究资料，所有资料源自AMiner后台程序对大数据的自动分析得到，本研究报告仅作为参考，AMiner团队不保证所分析得到的准确性和完整性，也不承担任何投资者因使用本产品与服务而产生的任何责任。

公司	名称	国家	时间
IBM	Summit	美国	2018.06 –今
国家并行计算机工程技术研究中心	神威·太湖之光	中国	2016.6-2017.11
国防科技大学	天河-2	中国	2013.06-2016.06
克雷公司 (Cray)	Titan	美国	2012.11-2013.06
IBM	蓝色基因/Q	美国	2012.06-2012.11
理化学研究所	京(超级计算机)	日本	2011.06-2012.06
国防科技大学	天河-1	中国	2010.11-2011.06
Cray	美洲虎(超级计算机)	美国	2009.11-2010.11
IBM	走鹃(超级计算机)	美国	2008.06-2009.11
IBM	蓝色基因/L	美国	2004.11-2008.06
日本电气 (NEC)	地球模拟器	日本	2002.06-2004.11
IBM	ASCI White	美国	2000.11-2002.06
英特尔	ASCI Red	美国	1997.06-2000.11
日立	CP-PACS	日本	1996.11-1997.06
日立	SR2201	日本	1996.06-1996.11
富士通	数值风洞	日本	1994.11-1996.06
英特尔	Paragon XP/S140	美国	1994.06-1994.11
富士通	数值风洞	日本	1993.11-1994.06
TMC	CM-5	美国	1993.06-1993.11



扫码关注 下载报告

<https://www.aminer.cn/>

THANKS