



北京交通大学  
BEIJING JIAOTONG UNIVERSITY



# 《大数据概论》

## 大数据架构与处理

鲍鹏  
软件学院





# 目录

---

- 集中式分布架构
- 分布式计算架构
- 大数据计算处理加速技术





# 大型主机的特点和优势

- 传统的集中式数据计算架构包括大型主机和超级计算机系统。
- 大型主机相比于其他计算机系统, 主要特点在于其 **RAS** (Reliability, Availability, Serviceability; 高可靠性、高可用性、高服务性) 的特性。





# 大型主机的特点和优势

## 1) RAS

- 大型主机一般在系统内集成了**高强度的冗余和错误检查技术**，防止系统发生灾难性问题。每个处理器核心都有两个执行通道来同时执行每一条指令。如果两个通道的计算结果不一致，CPU的状态就会复原。重新执行该条指令，结果还是不一致的话，一个空闲状态的CPU 将会被激活替代当前CPU。除了CPU，其他元件例如记忆芯片、内存总线、I/O通道电源等，都有相应的冗余设计，确保系统的高可靠性、高可用性。即使出错，许多组件的**热拔插**特性也能确保系统的高服务性，在系统运行的同时被更新。



# 大型主机的特点和优势

## 2) I/O吞吐量

- 除了RAS外，大型主机还被设计用来处理大容量I/O的应用。大型主机的设计中包括一些辅助电脑来管理I/O吞吐量的通道，让CPU解放出来只处理高速内存中的数据，每一个I/O通道都能同时处理许多I/O操作和控制上千个设备。使用大型主机处理数据中心GB级或是TB级的文件是常见的，相比于普通的个人电脑，大型主机经常是同时处理上千个数据流，并且能保证每个数据流的高速运转。



# 大型主机的特点和优势

## 3) ISA系统指令架构

- 作为大型主机市场的绝对霸主，IBM大型机的整体指令集保持了对应用程序的**向后兼容**。这样客户使用新的硬件就更为容易，只需换上新系统而**无须做额外的软件测试工作**。大型主机的投资回报率像其他的计算机平台一样，取决于其规模、所支持的工作负载。



# 大型主机所面临的问题

- 20世纪80年代以来。计算机系统向网络化和微型化的发展日趋明显，传统的大型主机处理模型越来越不能适应人们的需求：
  - 1) 大型主机的人才培养成本非常高，通常一台大型主机汇集了大量精密的计算机组件，操作非常复杂，这对一个运维人员掌握其技术细节提出了非常高的要求；
  - 2) 大型主机的价格通常非常昂贵，通常一台配置较好的IBM大型主机其售价达到上百万美元甚至更高，因此也只有像政府、金融和电信等企业才有能力采购大型主机。



# 大型主机所面临的问题

- 3) 大型主机虽然在性能和稳定性方面表现卓越，但并不代表永远不会出故障。一旦一台大型主机出现了故障，那么整个系统将处于不可用的状态，后果相当严重。随着业务的不断发展，用户访问量迅速提高，计算机系统的规模也在不断扩大，在单一大型主机上进行扩容往往比较难；
- 4) 随着PC机性能的不断提升和网络技术的快速普及，大型主机的市场份额变得越来越小，很多企业开始放弃原来的大型主机，改用小型机和普通的PC服务器来搭建分布式计算机。





# 超级计算机的发展历史

- 超级计算机与通用计算机相比具有较高的计算性能。
- 超级计算机的性能以每秒执行的浮点运算(FLOPS)次数来衡量。超级计算机在计算科学领域发挥着重要的作用，被用于各个领域计算密集型任务中，包括量子力学、天气预报、气候研究、石油和天然气勘探、分子建模(分析化合物、生物大分子和聚合物的结构和性质)以及物理模拟(如宇宙早期模拟、飞机和宇宙飞船的空气动力学、核武器爆炸和核聚变模拟)。

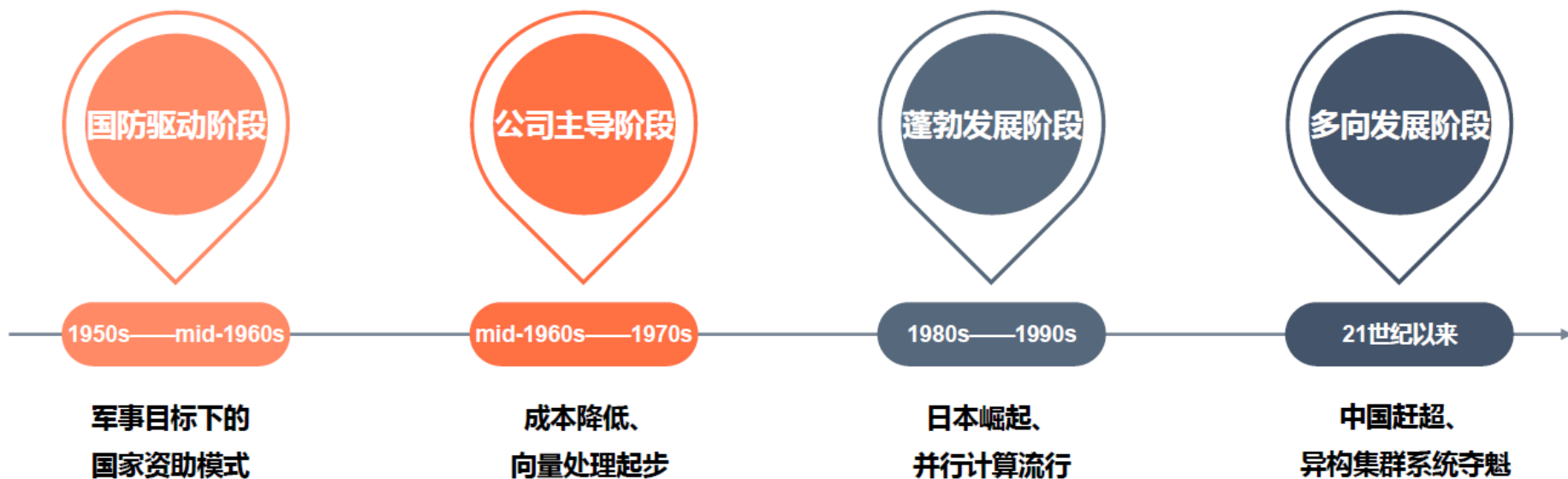


# 超级计算机的发展历史

	结构	典型代表
第一代	单芯片系统	CDC7600, IBM360
第二代	向量处理系统	Cray XMP, Cray YNP, NEC SX2和我国的银河一号, 银河二号
第三代	大规模并行处理(MPP)体系	IBM SP2, Intel Paragon和我国的曙光2000, 3000等
第四代	共享内存结构	SUN E10000/15000, SGI Origin 2000/3000和我国的银河三号, 神威一号等
第五代	机群系统(CLUSTER)	国外的“雷鸣”, “闪电”, “红色风暴”和国内的曙光4000系列, 联想1800/6800, 浪潮天梭TS10000, 巨星10000等



# 超级计算机的发展历史





# 超级计算机的发展历史

- 长期以来，美国一直是超级计算机领域的领军者，20世纪80年代和90年代，日本在这一领域取得了重大进展，中国在超算领域的影响力也变得越来越强。国防科技大学研制的天河二号超级计算机连续六次蝉联TOP500桂冠，2017年11月，同样来自中国的神威·太湖之光超过了天河二号，摘得TOP500桂冠。神威·太湖之光也是中国首度自行设计不使用英特尔等美国公司的核心产品而登TOP500第一名宝座的超级计算机。



# 超级计算机的发展历史

- 超级计算机能够为大数据应用提供强大的计算能力、海量的存储空间。但是，超级计算机是面向高性能计算应用设计的计算系统，高性能计算与大数据处理存在多方面差异。表5-1从应用领域、密集型、数据存储、软件平台等不同角度对高性能计算(high-performance computing,HPC)应用和大数据应用特点进行比较，可见超级计算机在处理大数据和高性能计算时所表现出的性能具有明显的差异。

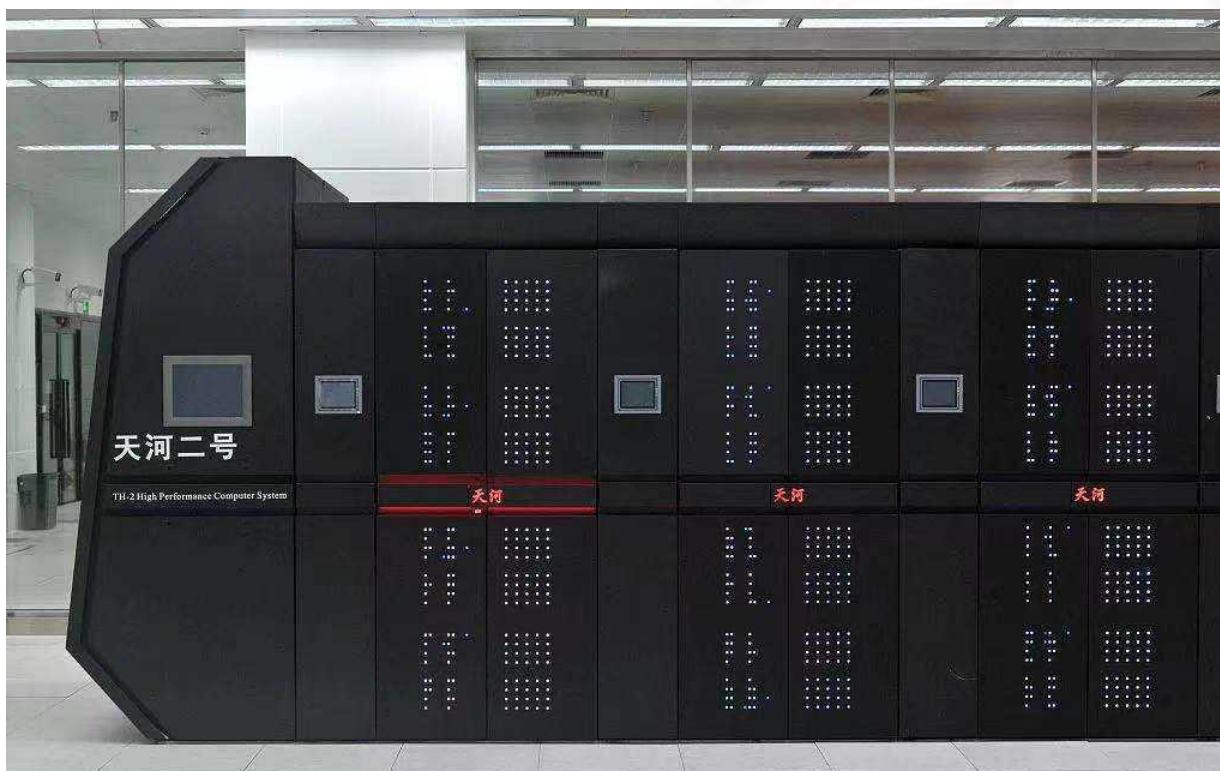
表5-1 大数据应用与高性能计算应用的差异

	大数据应用	高性能计算应用
应用领域	政府、商业、金融	科学与工程
密集型	数据	计算
数据存储	计算节点附近	与计算分离
并行性	隐式并行(数据并行)	显式并行
粒度	粗粒度	细粒度
耦合度	松耦合	紧耦合
软件平台	Hadoop, Spark, Storm等	MPI, OpenMP, Lustre等



# 大数据对超级计算机提出的挑战

- 以天河二号为例，针对大规模数据处理，天河二号采用的优化措施包括**互联通信优化**、**I/O增强**及**软件栈增强**等技术。







# 大数据对超级计算机提出的挑战

- 目前,天河二号计算机系统的互联通信已升级成自制的**TH Express-2**高速网络技术,该通信网络能提供更**高的通信性能和TCP/IP协议数据传输性能**。在I/O系统中,通过部署**混合的存储结构和层次式的文件系统**来提高I/O吞吐能力。在天河二号上提供**异构Mapreduce**并行处理框架,支持CPU和MIC的混合Mapreduce处理能力。天河二号在高性能系统设计过程中专门针对数据领域的应用特征进行了多项体系结构调优,包括I/O增强、软件栈增强、通信协议增强及并行处理模式优化等。