

并行计算

第一讲 并行计算概述

何克晶

kejinghe@gmail.com

华南理工大学
计算机科学与工程学院

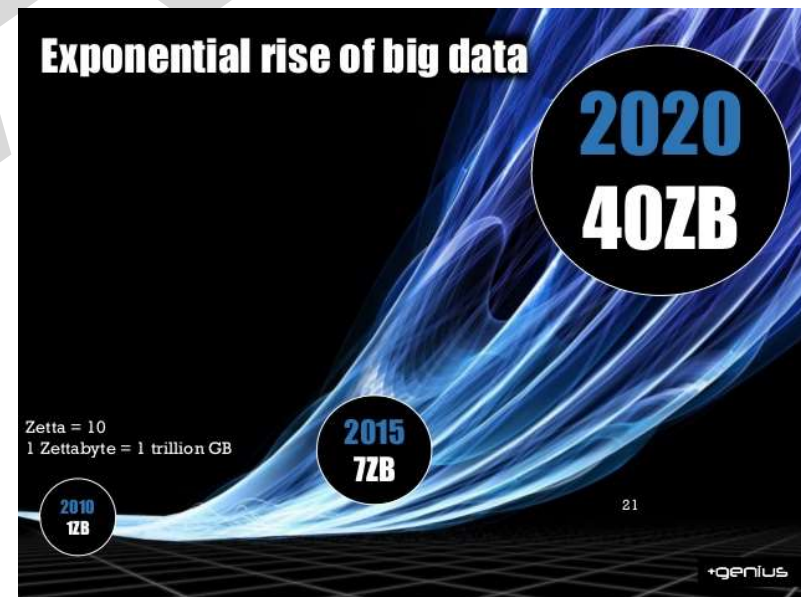
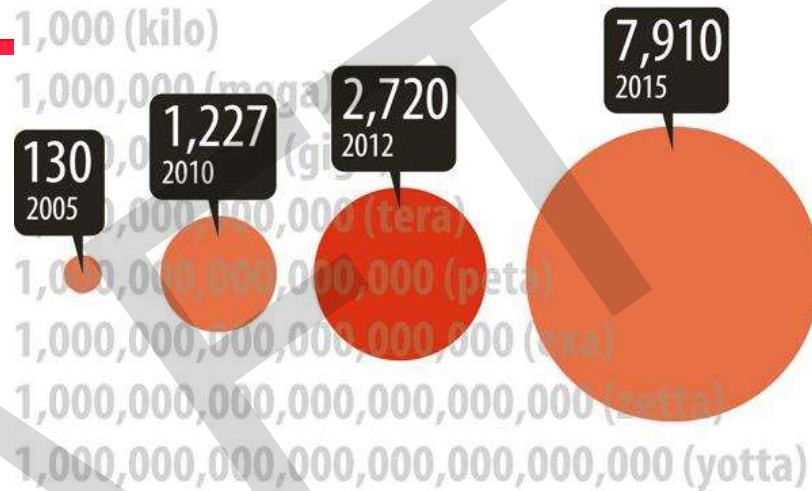


华南理工大学
South China University of Technology

IDC预测，2025 年预计将
创建 163ZB 的数据 (1ZB
等于 1 万亿 GB)，相当于
2016 年所产生 16.1ZB 数
据的十倍

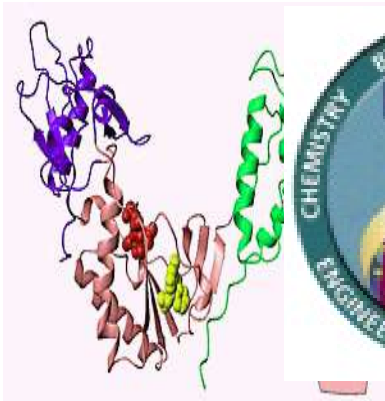
Exponential

Quantity of global digital data, exabytes

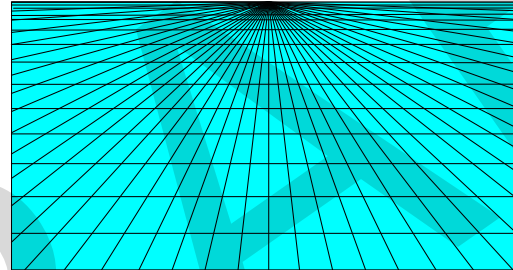


科学研究与工程应用

 Solving grand challenge applications using *modeling*, *simulation* and *analysis*



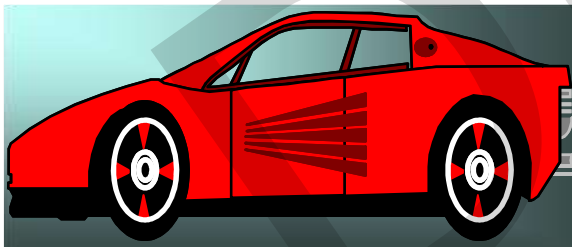
Life Sciences



Aerospace



Internet &
Ecommerce



CAD/CAM



Digital Biology



Military Applications



咋办?

提高计算机性能:

- 研究**CPU**运算速度更快的量子计算机、生物计算机与光计算机。
- 用多个处理器来协同求解同一问题，即将被求解的问题分解成若干个部分，各部分均由一个独立的处理机来计算

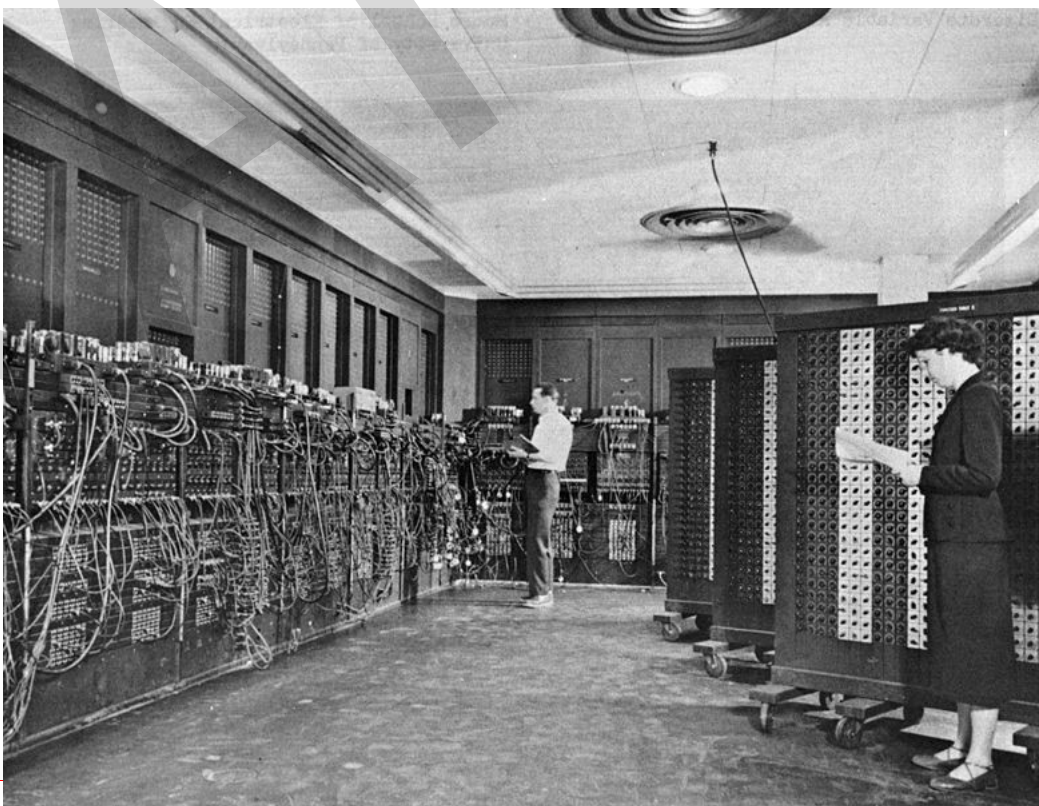


并行计算机的发展

□ 始于70 年代

- 1946年第一台计算机 ENIAC
(Electronic Numerical Integrator And Computer)

- 第一代：电子管
- 占地170平方
- 重约 30 吨
- 5000 次加法/秒
或500次乘法/秒
- 15分钟换一个零件
- 主要用于弹道计算
和氢弹研制



并行计算机的发展

□ 1964年

● 第一台超级计算机 CDC 6600

- ✓ 西蒙·克雷（Seymour Cray, 1925-1996）——“巨型机之父”
- ✓ 35万个晶体管
- ✓ 第二代计算机
- ✓ 标量计算机，使用了一个中央处理器（Central Processor, CP）和多个外围处理器（Peripheral Processor, PP）的结构

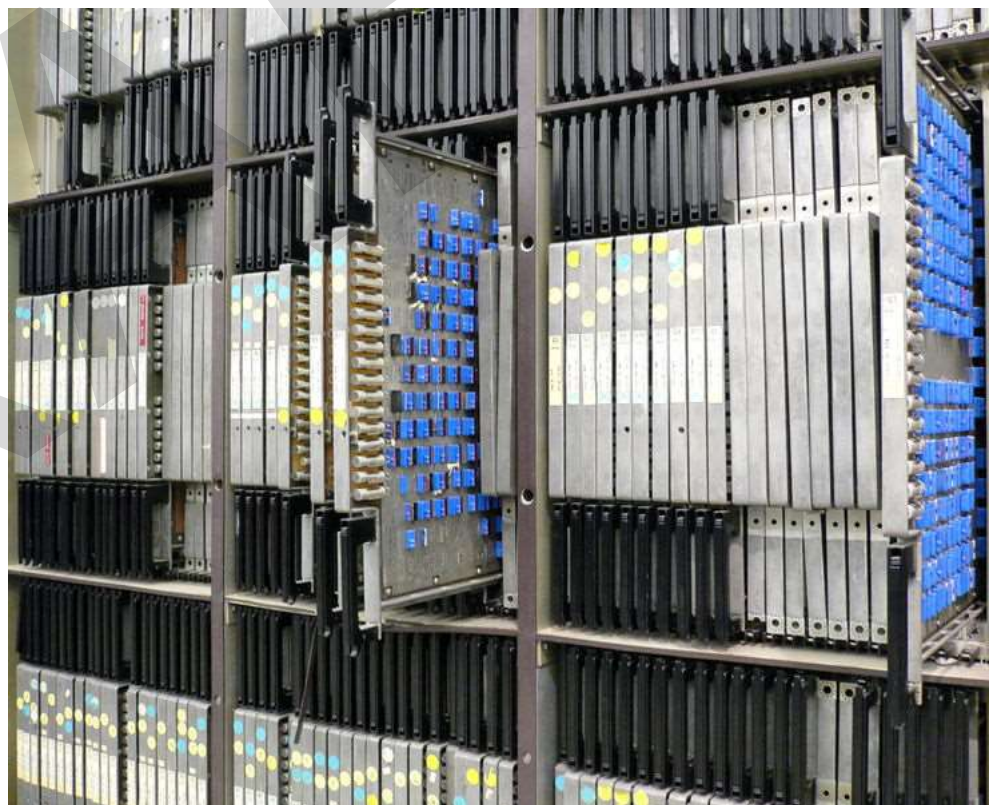


并行计算机的发展

□ 始于70 年代

● 1972年第一台并行计算机 ILLIAC IV（伊利诺依大学）

- ✓ 60 年代末开始建造
- ✓ 72 年建成, 74 年运行第一个完整程序, 76年运行第一个应用程序
- ✓ 64 个处理器, 是当时性能最高的CDC7600机器的 2-6倍
- ✓ 被认为是MPP计算机的鼻祖
- ✓ 公认的1981年前最快
- ✓ 1982年退役
- ✓ 可扩展性好, 但可编程性差

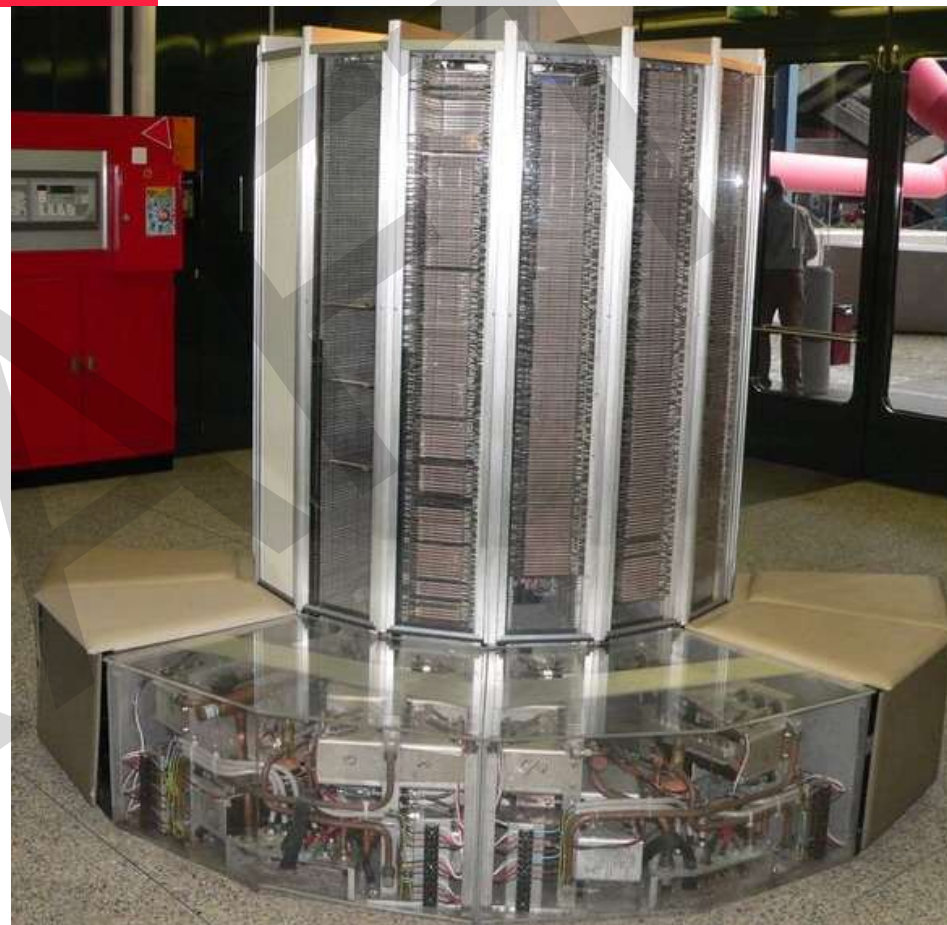


并行计算机的发展

□ 始于70 年代

● 向量机 Cray-1

- ✓ 一般将 Cray-1 投入运行的 1976 年称为“**超级计算元年**”
- ✓ 编程方便，
但可扩展性差
- ✓ 以 Cray 为代表的
向量机称雄超级
计算机界十几载



收藏于 Deutsches Museum 德意志博物馆的 Cray-1原型



并行计算机的发展

□ 80 年代百家争鸣

● 早期：以 MIMD 并行计算机的研制为主

- ✓ Denelcor HEP （1982年）第一台商用 MIMD 并行计算机
- ✓ IBM 3090 80 年代普遍为银行所采用
- ✓ Cray X-MP Cray 研究公司第一台 MIMD 并行计算机



西摩·克雷 Seymour Cray (1925-1996)，
电子工程学学士，应用数学硕士，
超级计算机之父，Cray研究公司的创始人，
亲手设计了Cray机型的全部硬件与操作系统，
作业系统由他用机器码编写完成。1984年时，
公司占据了超级计算机市场 70%的份额。
1996年Cray研究公司被SGI收购，2000年被
出售给Tera计算机公司，成立Cray公司。

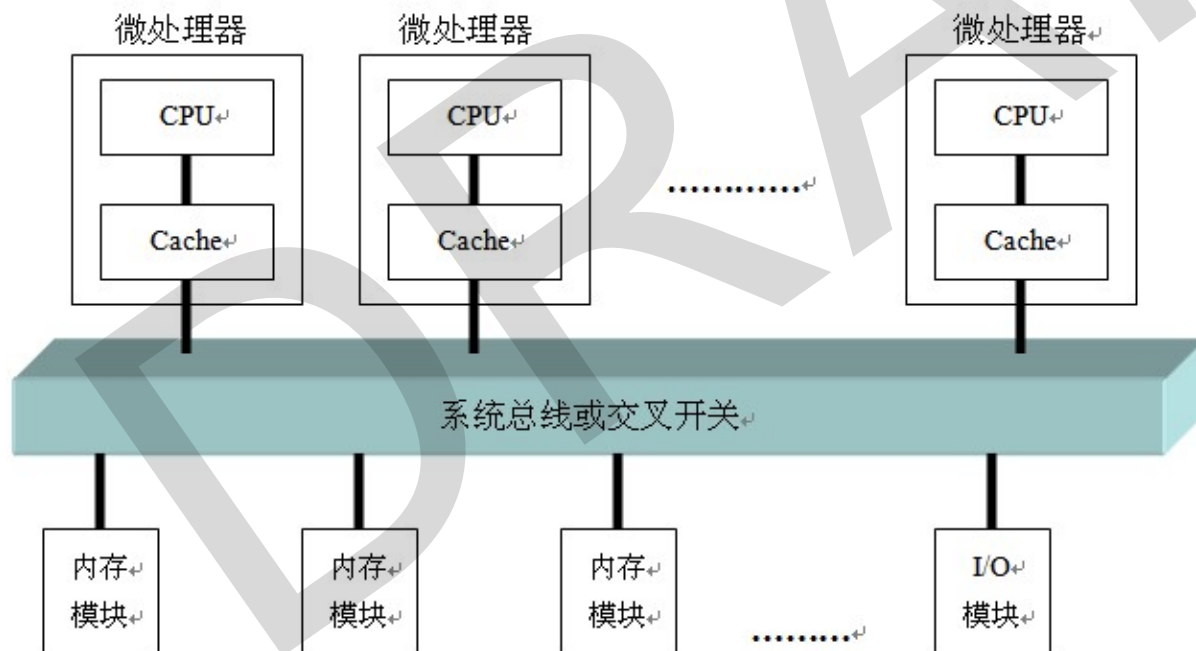


并行计算机的发展

□ 80 年代百家争鸣

- 中期：共享存储多处理机 Shared-Memory MultiProcessor

✓ **SMP** (Symmetrical Multi-Processing): 在一个计算机上汇集一组处理器，各处理器对称共享内存及计算机的其他资源，由单一操作系统管理，极大提高整个系统的数据处理能力。



- ✓ 扩展性较差
- ✓ 可靠性较差
- ✓ 内存访问瓶颈

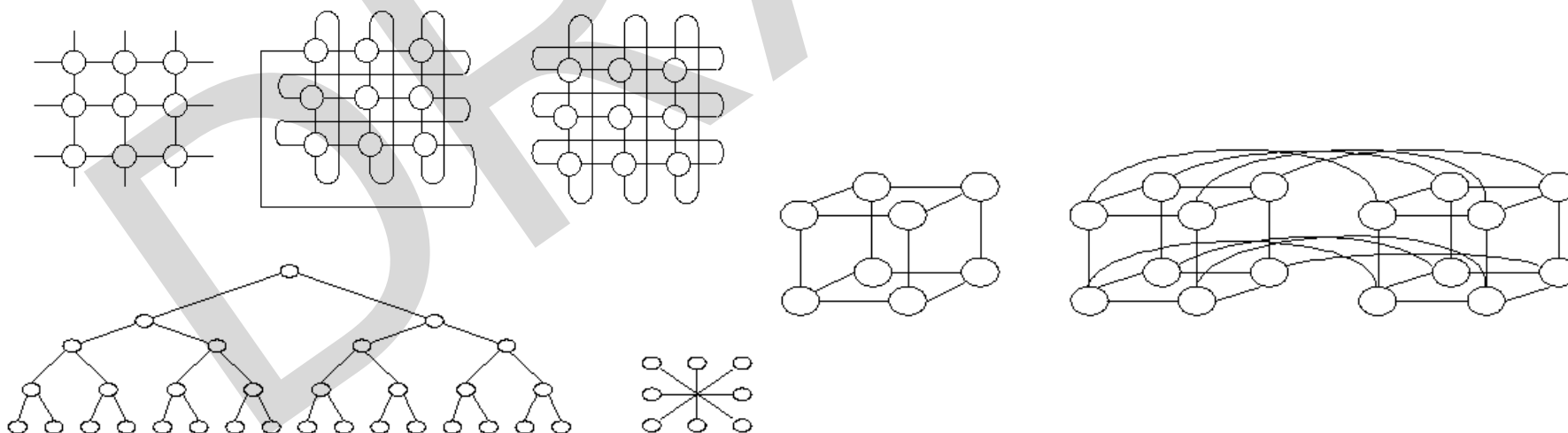


并行计算机的发展

□ 80 年代百家争鸣

● 后期：具有强大计算能力的并行机

- ✓ 通过二维Mesh连接的Meiko（Sun）系统
- ✓ 超立方体连接的 MIMD 并行机：nCUBE-2、iPSC/80
- ✓ 共享存储向量多处理机 Cray Y-MP
- ✓



并行计算机的发展

□ 90 年代：体系结构框架趋于统一 (DSM、MPP、NOW)

- DSM (Distributed Shared Memory) 分布式共享存储
- MPP (Massively Parallel Processing) 大规模并行处理结构
 - ✓ DM-MPP 分布式存储 MPP：每个结点仅包含一个微处理器
 - ✓ SMP-MPP：每个结点是一台 SMP 并行机
 - ✓ DSM-MPP：每个结点是一台 DSM 并行机
- NOW (Network of Workstations) 工作站机群
 - ✓ 也称为 COW (Cluster of Workstations)
 - ✓ NOW (COW) 与 MPP 之间的界线越来越模糊



并行计算机的发展

□ 2000 年至今：前所未有大踏步发展

● Cluster 机群

- ✓ 每个结点含多个商用处理器，结点内部共享存储
- ✓ 采用商用机群交换机通过前端总线连接结点，结点分布存储
- ✓ 各个结点采用 Linux 操作系统、GNU编译系统和作业管理系统

● Constellation 星群

- ✓ 每个结点是一台子并行机
- ✓ 采用商用机群交换机通过前端总线连接结点，结点分布存储
- ✓ 各个结点运行专用的结点操作系统、编译系统和作业管理系统

● MPP

- ✓ 专用高性能网络，大多为政府直接支持



HPC TOP500 List – June. 2019

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	Summit , United States	2,414,592	148,600.0	200,794.9	10,096
2	Sierra , United States	1,572,480	94,640.0	125,712.0	7,438
3	Sunway TaihuLight , China	10,649,600	93,014.6	125,435.9	15,371
4	Tianhe-2A , China	4,981,760	61,444.5	100,678.7	18,482
5	Frontera , United States	448,448	23,516.4	38,745.9	
6	Piz Daint , Switzerland	387,872	21,230.0	27,154.3	2,384
7	Trinity , United States	979,072	20,158.7	41,461.2	7,578
8	AI Bridging Cloud Infrastructure (ABCI) , Japan	391,680	19,880.0	32,576.6	1,649
9	SuperMUC-NG , Germany	305,856	19,476.6	26,873.9	
10	Lassen , United States	288,288	18,200.0	23,047.2	



硬件数据	
核心数	2397824个
处理器	IBM POWER9 22C 3.07GHz
主存储容量	1310720GB
互连技术	Dual-rail Mellanox EDR Infiniband
性能	
Linpack峰值	143500.0TFlop/s
理论运行速率峰值	200795.0 TFlop/s
Nmax	16693248
HPCG测试	2925.75TFlop/s
能源效率	
功率	9783.00千瓦
功率测量等级	3
测量的核心数	2397824个

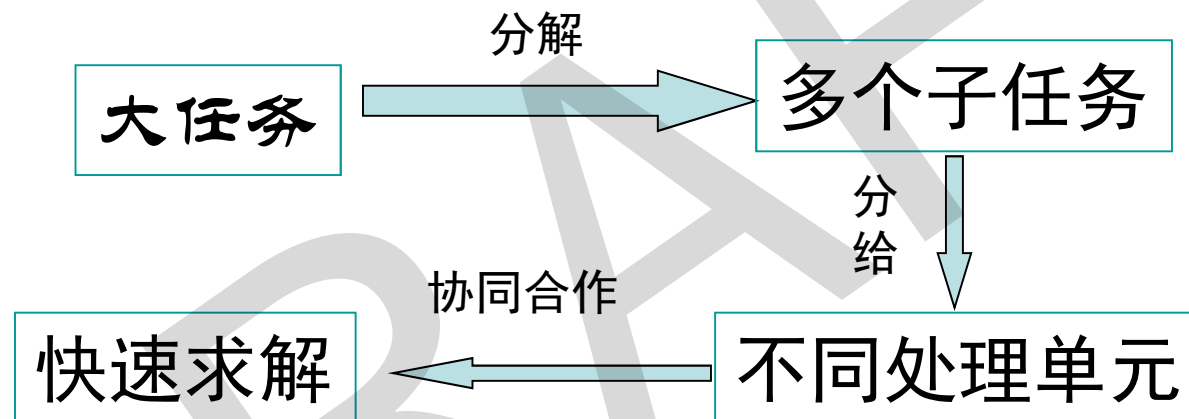
并行计算 (Parallel Computing)

什么是并行计算？



并行计算 (Parallel Computing)

□ 并行计算，同义词：高性能计算（High Performance Computing）、超级计算（Super Computing）



在并行机上，将一个应用分解成多个子任务，分配给不同的处理器，各个处理器之间相互协同，并行地执行子任务，从而达到加速求解速度，或者求解应用问题。



并行计算2

□ 基本条件：

- ✓ 硬件（并行机）：并行机至少包含两台或两台以上处理机，这些处理机通过互连网络相互连接，相互通信。
- ✓ 并行算法设计：也就是说，应用可以分解为多个子任务，这些子任务可以并行地执行。将一个应用分解为多个子任务的过程，称为并行算法的设计。
- ✓ 并行编程环境：在并行机提供的并行编程环境上，具体实现并行算法，编制并行程序，并运行该程序，从而达到并行求解应用问题的目的。

□ 主要目标：

- ✓ 提高求解速度：例如，在单处理器上，串行执行需要2 个星期（14 天），借助并行计算，使用100 台处理器，加速50 倍，将执行时间缩短为6.72 个小时。
- ✓ 扩大问题规模：例如，在单处理器上，受内存资源2GB的限制，只能计算10 万个网格，也可以借助并行计算，使用100 个处理器，将问题求解规模线性地扩大100 倍。



并行计算 vs 分布式计算

分布式计算（distributed computing）：

- 通过网络相互连接的两个以上的处理机相互协调，各自执行相互依赖的不同应用，从而达到协调资源访问，提高资源使用效率的目的。
- 无法达到并行计算所倡导的提高求解同一个应用的速度，或者提高求解同一个应用的问题规模的目的。
- 分布式计算和并行计算通常相互配合，既要通过分布式计算协调不同应用之间的关系，又要通过并行计算提高求解单个应用的能力。



并行计算 vs 分布式计算2

并行计算:

- 给定二维规则区域上的Dirichlet问题 $\Delta u = f$, 采用标准5点有限差分格式离散。平均分配 $N = N_x \times N_y$ 个网格单元给 $P = P_x \times P_y$ 台处理机。所有处理机并行计算, 执行Jacobi迭代, 求解Dirichlet问题, 从而达到缩短求解问题的时间, 或者扩大网格规模 N 的目的。

分布式计算:

- 观众点播, 远程驾驭式可视化, 电视会议等。

