

上海第二工业大学

并行体系结构和并行计算

Parallel Architecture and Parallel Computing

主讲人 陈林

Spring, 2021

Overview

- ~~Course Administration~~
- Course Style and Structure
- Intro to Parallel Computing

Course Administration

Reference Book:

- . Kai Hwang, Zhiwei Xu, "Scalable Parallel Computing", McGraw-Hill, 1998
- . J. JaJa, "Introduction to Parallel Algorithms", Addison Wesley, 1992
- . V. Kumar et al, "Intro to Parallel Computing", Benjamin Cummings, 1994
- . 陈国良, 并行算法的设计与分析 (第3版), 高等教育出版社, 2009
- . 陈国良等, 并行计算机体系结构, 高等教育出版社, 2002
- . 陈国良等, 并行算法实践, 高等教育出版社, 2003
- . J. Dongarra et al, "Sourcebook of Parallel Computing" (莫则尧等译), 电子工业出版社, 2005
- . Shameem Akhter, et. al. 著, 李宝峰等译. 多核程序设计技术, 电子工业出版社, 2007
- . Richard Gerber, et. al. 著, 王涛等译. 软件优化技术, 电子工业出版社, 2007)

Grading

.Grade breakdown

. Final Exam:	50%
. Homework Assignments	20%
. Experiments	30%

.

Overview

- Course Administration
- ~~Course Style and Structure~~
- Intro to Parallel Computing

- . Parallel Architectures**
- . Parallel Algorithms**
- . Parallel Programming**
- . Parallel Applications**

Abstract

- . **Generally speaking**, parallel computing deals with the parallel computer architectures, parallel algorithms and parallel programming. In this lecture we will discuss briefly them separately.
- . **In part I**, we will discuss the contemporary parallel parallel system interconnections and parallel system performance evaluation.
- . **In part II**, we will discuss the parallel computational models, the design methods, techniques and methodology of parallel algorithms, as well as some parallel numerical algorithms.
- . **In part III**, we will discuss the parallel programming models, shared-memory, message-passing and data parallel programming, as well as parallel programming environment and tools.
- . **In part IV**, we will discuss ...

Part I : Parallel Hardware System

Hardware Platform for Parallel Computing:

- . System Architectures and Models
- . System Interconnections
- . Performance Evaluation

System Architectures and Models

- Parallel Computer System Architectures
 - **PVP**: Parallel Vector Processors
 - **SMP**: Symmetric Multiprocessors
 - **MPP**: Massively Parallel Processors
 - **DSM**: Distributed Shared Memory
 - **COW**: Cluster Of Workstations
- Parallel Computer Memory Access Models
 - **UMA**: Uniform Memory Access
 - **NUMA**: Non-Uniform Memory Access
 - **COMA**: Cache-Only Memory Access
 - **NORMA**: NO-Remote Memory Access

System Interconnections

▪ Network Environments

- Inter-node Interconnections(Buses , Switches)
- Inter-node Interconnections(SAN)
- Inter-system Interconnections(LAN , MAN , WAN)

▪ Interconnection Topologies

- Static-Connection Networks(LA,RC,MC,TC,HC,CCC)
- Dynamic-Connection Networks (Buses, Crossbar, MIN)

▪ Wide-Band Networks

- FDDI(Fiber Distributed Data Interface)
- FE/GE(Fast Ethernet / Gigabit Ethernet)
- ATM(Asynchronous Transfer Mode)
- SCI(Scalable Coherence Interface)

Performance Evaluation

- Speed up of Systems
 - Amdahl's Law
 - Gustafson's Law
 - Sun and Ni's Law
- Scalability of Systems
 - Iso-efficiency
 - Iso-speed
 - Average Latency
- Performance of Systems: Benchmarks
 - LINPACK
 - SPEC
 - PARMBENCH
 - NAS etc

Part II : Parallel Algorithms

Theoretical Base for Parallel Computing:

- . Computational Models
- . Design Policy
- . Design Methodology
- . Design Techniques
- . Parallel Numerical Algorithms

Computational Models

- . **PRAM** : Parallel Random Access Machines
- . **APRAM** : Asynchronous PRAM
- . **BSP** : Bulk Synchronous Parallel
- . **LogP** : Latency , Overhead , Gap ,
Processors

Design Policy

- . **Parallelizing a Sequential Algorithm**
- . **Designing a new Parallel Algorithm**
- . **Borrowing Other Well-known Algorithm**

Design Methodology

- . **PCAM** : **P**artitioning
- . **PCAM** : **C**ommunication
- . **PCAM** : **A**gglomeration
- . **PCAM** : **M**apping

Design Techniques

Balanced Trees

- Balanced Trees
- Doubling Technique
- Partitioning Strategy
- Divide and Conquer
- Pipelining

Parallel Numerical Algorithms

- Dense Matrix Algorithms
- Solving Systems of Linear Equations
- Fast Fourier Transform

Part III : Parallel Programming

Software Support for Parallel Computing:

- . Programming Models
- . Shared-Memory Programming
- . Message-Passing Programming
- . Data-Parallel Programming
- . GPU CUDA Programming and OpenCL

Part IV : Parallel Applications

Applications for Parallel Computing:

- . Application Backgrounds
- . Parallel Computing for Atmospheric Model
- . Software Packages of Numerical Computing
- . Others: 3D Fourier Transform, Image Feature extraction, Seepage Computing, ...

Overview

- Course Administration
- Course Style and Structure
- ~~Intro to Parallel Computing~~

什么是并行计算？

- A parallel computer is a “collection of processing elements that communicate and cooperate to solve large problem fast”. -David E. Culler
- Or all processors cooperate to solve a single problem
- Daily life examples:
 - House construction //综合：并发、分布、流水
 - Car manufacturing //流水线
 - Grocery store operation //分布

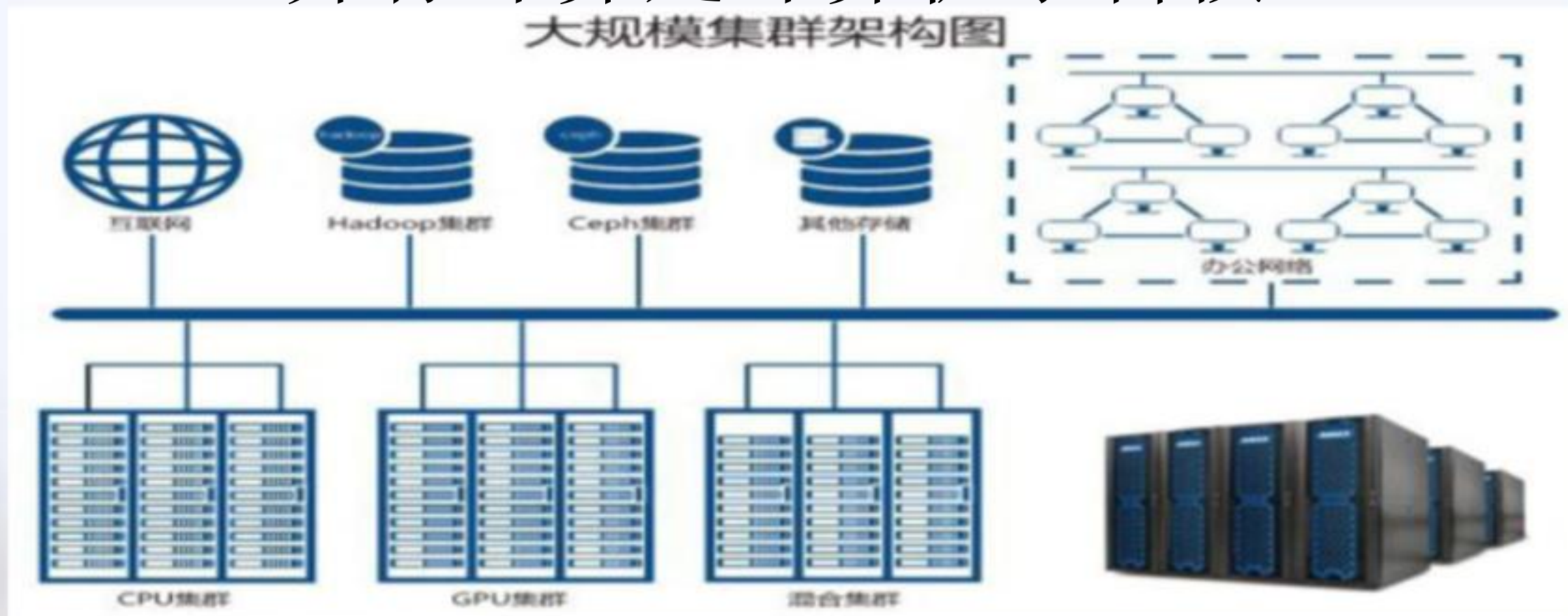
为什么需要并行计算？（1）

- Interest in parallelism since the very ancient era of computers(e.g. **ILLIAC IV of 1967 had 64 processors**)
- Parallel Processing is **an effective answer** for the tremendous future computing requirements.
- **applications impulses:**
 - Data-intensive applications: videoconferencing, virtual reality, large database and data mining, speech recognition, biology, image and signal processing, etc
 - Computing-intensive applications: numerical simulation(e.g. forecasting, manufacturing, chemistry, aerodynamics)
 - Network-intensive applications:
 - **Multicore and manycore and cloud computing:**

为什么需要并行计算？（2）

- **Grand challenges:**
 - Science today: experimentation, theory, simulation(or computation)
 - Simulation relies heavily on parallel processing
- **America HPCC project, ASCI project**
- **In one words: Parallel processing promises increase of**
 - Performance(e.g. large, fast, cost)
 - Reliability
 - Large set of computational problems are inherently parallel in nature. But their existing applications are designed for uniprocessor systems. Their parallelization is required.

并行计算是计算机学科核心



串行软件 = 数据结构 + 串行算法 + 串行编程 + 文档

跨越、非渐变  并行思维需要学习

并行软件 = 数据结构 + 并行算法 + 并行编程 + 文档 +
并行体系结构

并行计算是大数据和人工智能支撑技术



核心支撑： 算力 + 算法 + 数据

并行计算 并行算法 并行存储

- 算力 = 并行体系结构（硬件）+ 并行程序
- 算法 = 并行分布式算法（思维）+ 并行编程
- 数据 = 分布式存储技术（数据组织）+ 存储系统



高性能计算用于天气预报

气象预报: 天气预报是高性能应用的一个传统领域, 复杂的气候三维模型对超级计算机有着巨大需求。日本的地球模拟器就是用来做气象预报的超级计算机, 我国新一代的数值天气预报系统**Grapes**就是这类典型应用。这类应用程序对系统的通信性能要求较高。



高性能计算在中央气象台

中央气象台国内一流的计算机设备，峰值运算速度达3840亿次/秒的国产“神威”超级计算机就安装在这里，另外，还安装了国产银河III和曙光并行机，以及IBM cluster1600、IBM SP、SP2、CRAY、CRAY、HP、SUN、SGI、ALPHA等高性能巨型机。

通过气象卫星、气象雷达等大气观测探测系统收集到的数据，以国际通信网、国内通信网、互联网传输到气象信息网络系统中，将全球资料分析同化处理之后，形成数值天气预报。

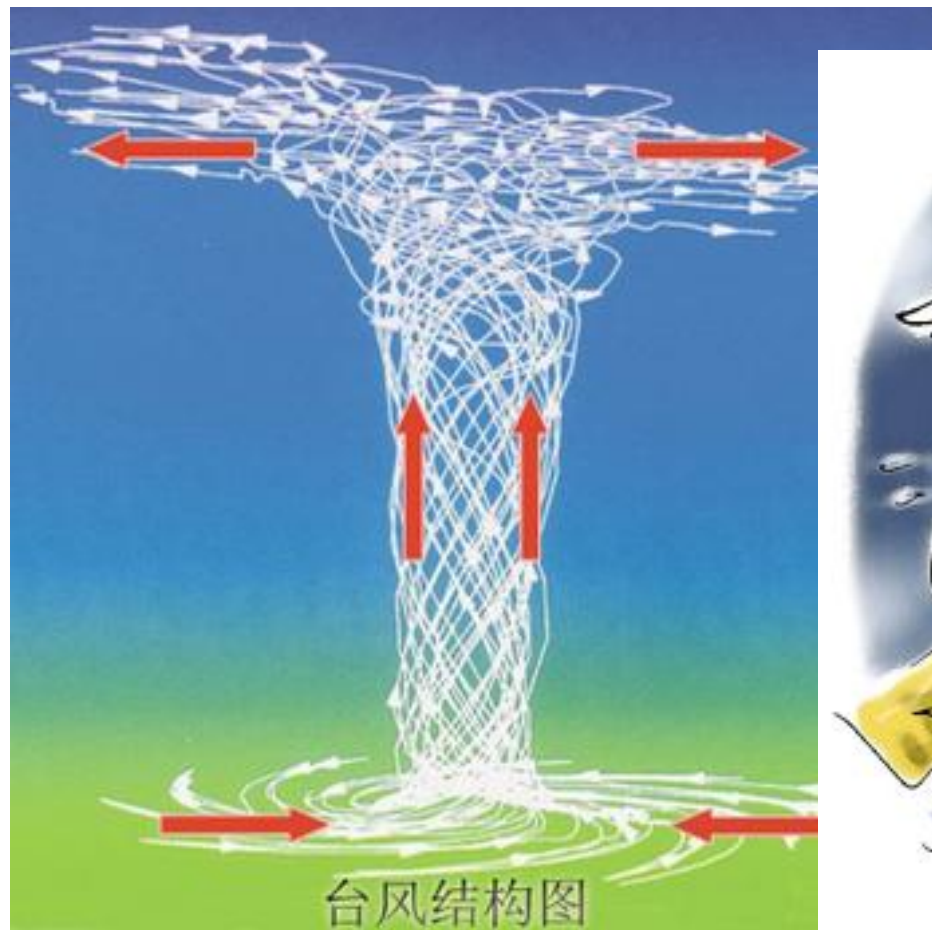
高性能计算是分析天气预报的主要手段。

每天要接收世界各地的气象报告，其中天气报告二万余份，地面航空报告近三万份，船舶报告三千余份，探空报告约一千六百份，高空报告约二千五百份，飞机探测报告约三千二百份，此外还有大量的卫星和雷达探测资料。

超级计算机用于预报台风飓风

计算云层的演变以预报台风、飓风

<http://cimss.ssec.wisc.edu>



“地球模拟器”用于研究全球天气

右图就是日本曾经排名第一的超级计算机“地球模拟器”。

在地球模拟器中，地球立方体格子边长为10公里，实现了比原来精细了1000倍。

“地球模拟器”用来进行全球范围的大气和海洋系统模拟运算，而不是仅仅限于局部地区。



高性能计算用于天体物理模拟

因为天体物理学基本上没有可以进行的实体实验，所以除了天文观测之外，天体物理学家只能通过计算机模拟的方法来证实自己的理论。

星系和星球的动力学演化计算以及星球的相互作用的轨道计算都是计算密集的任务。在对计算速度的需求之外还有高精度和高可靠性的需求。未来，这一领域需要1~10PFlops的计算能力，主要用于银河系形成模拟、行星形成模拟等。

由Max-Planck协会的天体物理学家领导的一个国际研究小组就利用超级计算机完成了当时世界上最大的一个宇宙景象模拟，他们使用了当前最精确的星系形成和超大质量黑洞的理论模型。这项成果发表在一期《Nature》（2 June 2005）上，并被选作了封面。

高性能计算用于核武器模拟

核武器模拟：采用以计算为基础的方法代替以地下核试验为基础的方法，利用数字方程和三维建模仿真核武器爆炸的效果，确保现有库存核武器的性能、安全和可靠性。国际上的超级计算机很多是用来做核武器研究的，比如美国能源部的ASC(Advanced Simulation and Computing)计划支持的大部分系统、法国的CEA(Atomic Energy Commission)工程。在核武器研究方面有很多大规模并行模拟程序需要运行在超级计算机上。因此，一些公开的核武器模拟并行程序可以作为高性能计算机评测的代码。比如，公开的程序SAGE、Sweep3D等。

高性能计算用于艺术设计

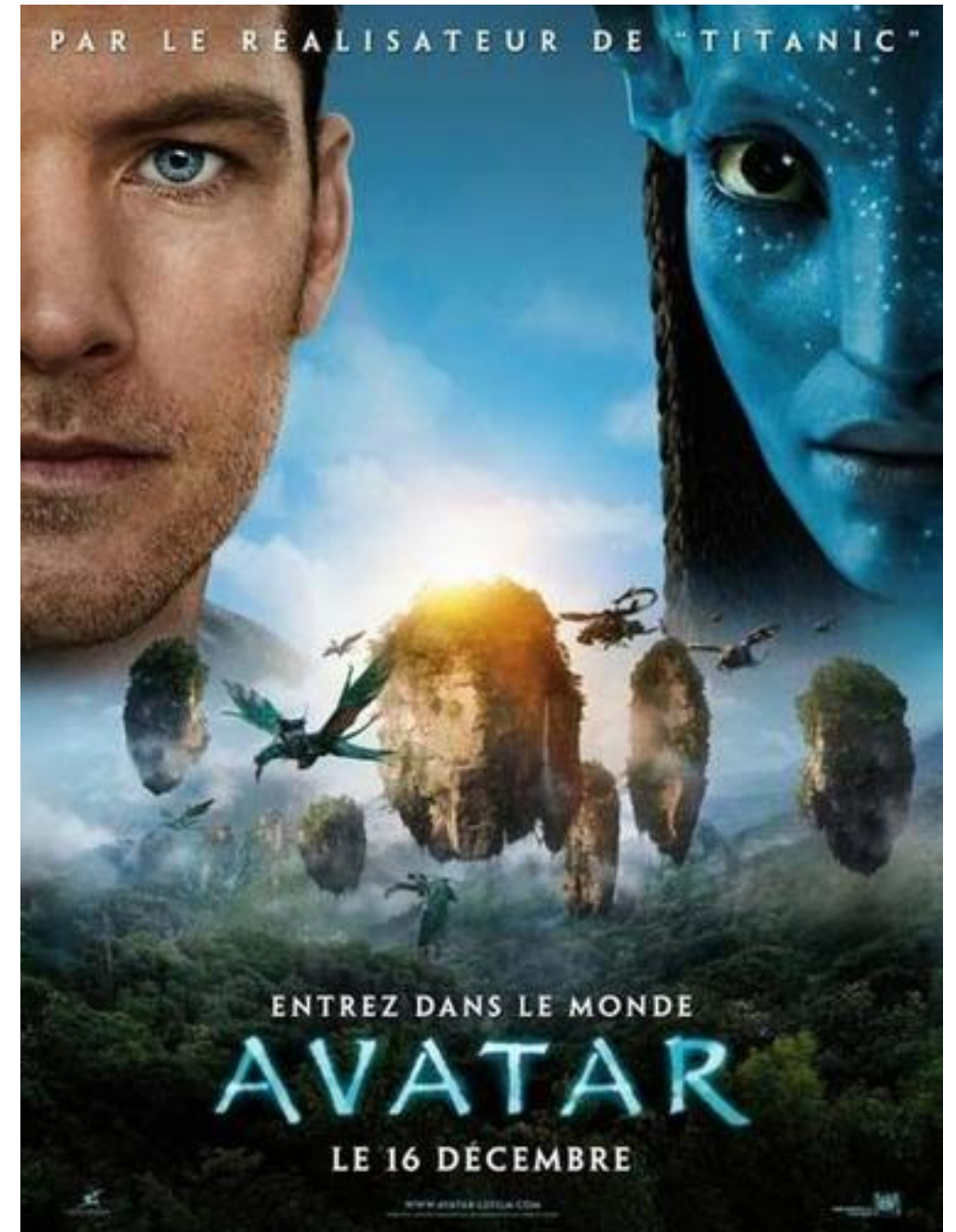
渲染：对三维动画、数字影片进行剪接、加效果、加字幕、音乐等操作，当生成影片时需要将后加入的素材融合到影片中并压缩成为影片最终格式，这个过程就是渲染。渲染属于计算高度密集型业务，对计算机硬件、软件设备的投入，以及后期维护和管理都有较高要求。

随着中国创意产业的发展，影视动画、工业设计、建筑环境设计、平面设计、工艺美术设计、服饰设计等领域对渲染的需求凸显，特别是随着全高清节目的逐渐流行，三维动画精细建模手段的不断丰富，要满足越来越高的影视后期制作需要，需要更高的计算能力做支撑。

在使用**CPU**进行多媒体相关领域的计算的同时，越来越多的超级计算机使用专门为渲染和多媒体设计的图形处理器**GPU**加速计算。**GPU**计算是高性能计算的又一个热点。

电影与高性能计算

炫目的特效成就了《阿凡达》。除了实景镜头外，电影场景的制作要通过数字建模、虚拟摄像、渲染。渲染环节需要极大的运算能力和数据吞吐量。为《阿凡达》制作特效的数据中心，每秒就要处理7-8GB的数据。特效使用的超级计算机是6套惠普的HP BladeSystem BL2x220c超高密度刀片服务器系统，每套处理器数都在5000以上，峰值计算能力都在30000GFlops以上。



HP BladeSystem BL2x220c中每个机箱可以容纳32台服务器节点，最多有128台服务器、1024个CPU核，以及4 TB内存。这个计算密度是传统1U机架服务器密度的三倍。

每个服务器节点最多支持2颗Intel至强4核CPU和 32GB RAM，每个服务器节点标配 2 个千兆以太网端口，可升级到万兆以太网或 Infiniband。

高清影片的数据量非常大，一帧的数据是 12 MB，一秒钟 24 帧，每分钟的数据就有 17.28 GB，而整部《阿凡达》长达160多分钟，而这还仅仅是剪辑完成的成片的数据，加上前期拍摄的以及后期制作所需的图片、素材等等，整部《阿凡达》电影产生的数据据说在3PB左右。

（1PB=1024TB, 1TB=1024GB, 1GB=1024MB）

超级计算机用于国防安全

国防安全: 这类应用主要是国防上用于密码破译、间谍侦探等方面。这类应用程序主要是信息处理和信息分析，对浮点计算需求比较小，一般对单次计算的响应时间要求比较高。

- 曾经在top500中排名第一的IBM Roadrunner就是的一个军用计算机。主要用于解决机密军事问题，以确保美国核武器储备的持续发展。

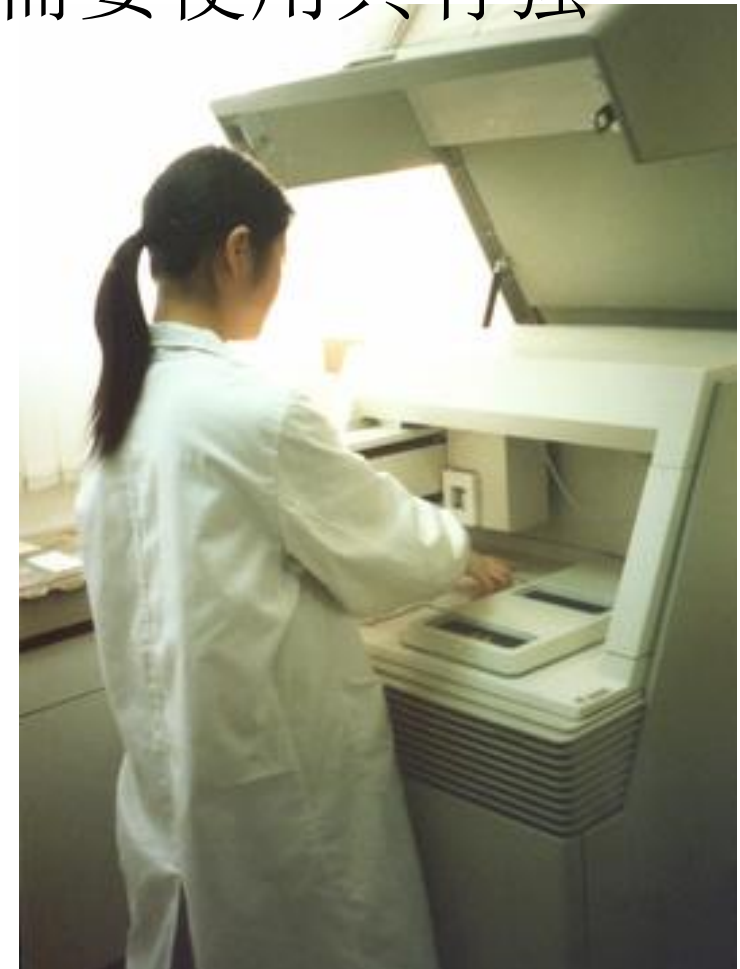


军事超级计算机 Roadrunner

超级计算机用于生物信息学

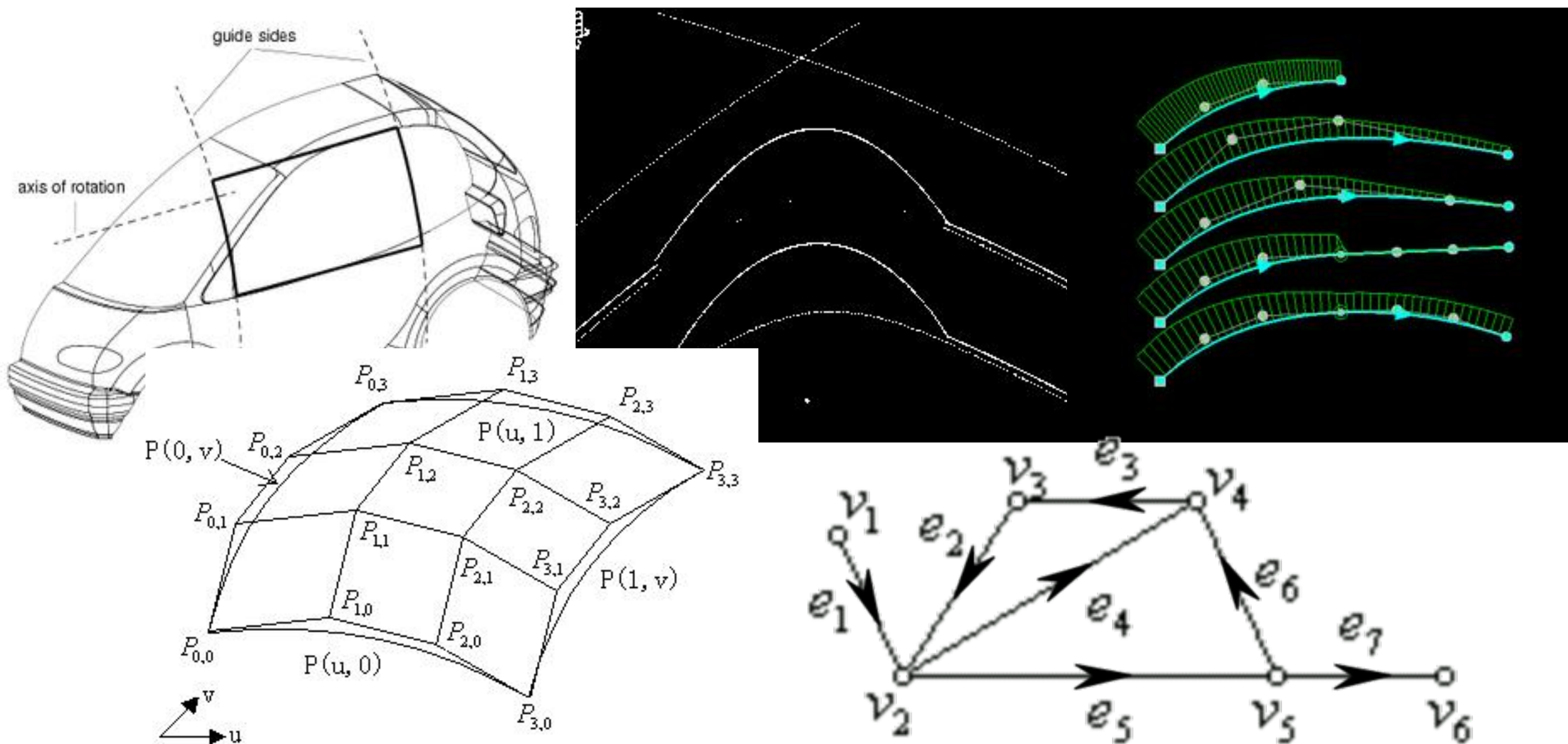
生物信息学: 通过超级计算机强大的计算能力来解决生物领域的难题是近些年研究的热点。比如基因组学、蛋白质组学、基因诊断等都是生物信息学的重要方向。

生物领域有很多海量数据处理应用程序, 比如基因拼接、多序列比对、蛋白质三维结构预测等, 它们都需要使用具有强大的计算能力的服务器来解决。



超级计算机更多应用

此外，机械设计、化工、交通调度安全、计算流体力学等领域都有着对超级计算机的巨大需求。



计算机性能的计量单位

Flops 即每秒浮点数运算(实数计算) 指令数

Byte 字节 $1\text{B}=8\text{bit}$, bit为计算机存储中的最小单位

1K (Kilo 千) $=1024$, 1M (Mega 兆) $=1024\text{K}$,

1G (Giga 吉) $=1024\text{M}$, 1T (Trillion 万亿) $=1024\text{G}$

1P (Peta 千万亿) $=1024\text{T}$, $1\text{E} = 1024\text{P} \dots$

1.026PetaFlops即浮点计算1.026千万亿次每秒

计算问题与时间要求

- 问题: 科学和工程问题的数值模拟与仿真
 - 计算密集
 - 数据密集
 - 网络密集
 - 三种混合
- 要求: 在合理的时限内完成计算任务
 - 秒级 制造业
 - 分钟级 短时天气预报(当天)
 - 小时级 中期天气预报(3~10日)
 - 尽可能快 长期天气预报(气候)
 - 可计算 湍流模拟

几种并行实现方案

- Multi core PC and GPU (lowest cost)
- Cluster of workstations (lower cost)
- Multiprocessor workstations (\$60,000)
 - DEC Firefly, Apollo DN 10000, SUN SPARCstation 20
- Shared memory multiprocessors (\$200,000-400,000)
 - Sequent Symmetry, Encore Multimax, SGI Challenge, SUN SPARCserver 2000
- Distributed memory multicomputers (\$200,000-400,000)
 - Intel iPSC/860, NCUBE/2, Meiko
- Massively parallel processors (\$5,000,000)
 - Intel Paragon, TMC CM-5, CRAY T3D, IBM SP-2

典型的多核处理器系统

- 通用处理器:
 - Intel——Xeon 5300 (四核)
 - IBM——POWER5(双核)
 - SUN——NIAGARA(8核)、UltraSpac(双核)
- 网络处理器
 - Intel——IXP2400
 - Motorola——C-5
- 嵌入式系统
 - TI——OMAP, Davinci
 - ARM——ARM11MP
- 图像处理
 - Nvidia——GF6800
- 多媒体处理
 - IBM, Sony和Toshiba——Cell处理器
 - Stanford大学——Imagine
- 实验系统
 - Ambric——AM2045(360核)
 - Intel——80个核的实验处理器

多核时代软件的挑战：在于并行算法和程序的并行化

并行计算的粒度

- Coarse-grained(粗粒度): Level of jobs
- Middle-grained(中等粒度): Level of processes
- Fine-grained(细粒度): Level of machine instructions(or lower)

并行计算的研究领域

- **Design of parallel computers:** How to the number of processors, communication throughput, data sharing, etc.
- **Design of parallel algorithms:** Parallel algorithms may be quite different from their sequential counterparts.
- **Design of parallel software:**
 - Operating systems
 - Compiles
 - Libraries
 - Tools: debuggers, performance analyzers
- **Applications of parallel computing**

<http://www.top500.org>

上海第二工业大学

[HOME](#)[LISTS ▾](#)[STATISTICS ▾](#)[RESOURCES ▾](#)[ABOUT ▾](#)[MEDIA KIT](#)

TOP500 Expands Exaflops Capacity Amidst Low Turnover

Nov. 16, 2020

FRANKFURT, Germany; BERKELEY, Calif.; and KNOXVILLE, Tenn.—The 56th edition of the TOP500 saw the Japanese Fugaku supercomputer solidify its number one status in a list that reflects a flattening performance growth curve. Although two new systems managed to make it into the top 10, the full list recorded the smallest number of new entries since the project began in 1993.

The entry level to the list moved up to 1.32 petaflops on the High Performance Linpack [HPL] benchmark, a small increase from 1.23 petaflops recorded in the June 2020 rankings. In a similar vein, the aggregate performance of all ...

[read more »](#)



TOP500 LIST



ANNIVERSARY



NEWSLETTER
SIGN UP



EuroHPC
The European
High Performance Computing
Infrastructure

EuroHPC
Summit Week 2021



Tweets by [top500supercomp](#)



TOP500 NEWS

TOP500 Event at ISC

High Performance

2020 Digital

June 23, 2020

Find the highlights of the 55th TOP500 list



THE LIST

11/2020 Highlights

After a make-over of the Top10 in June we again see some interesting changes driven by two system upgrades (#1 Fugaku and #5 Selene) and two new systems (#7 JUWELS Booster Module and #10 Dammam-7). The full list however recorded the smallest number of new entries ever since the project started in 1993.

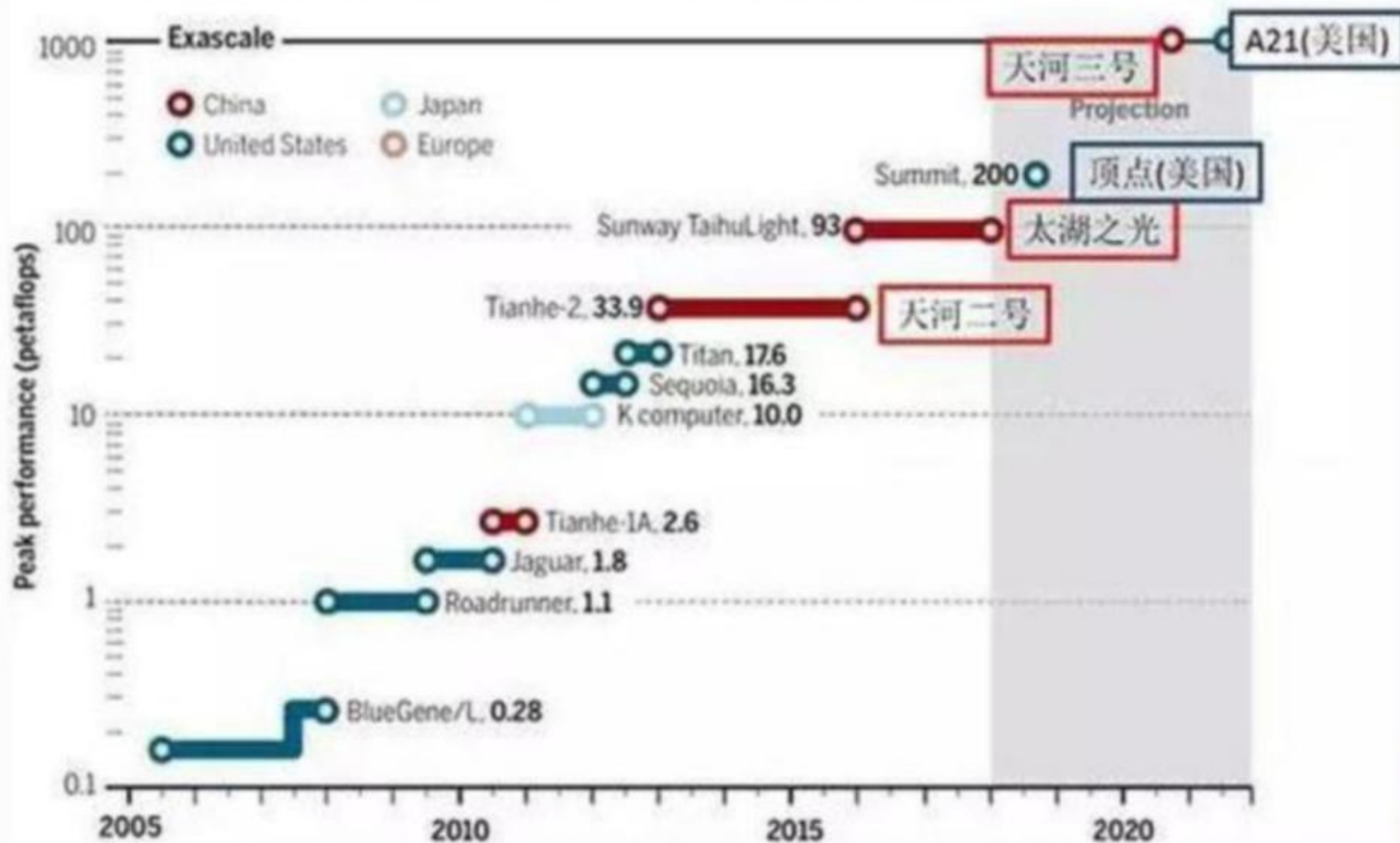
上海第二工业大学

TOP500 <http://www.top500.org>

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442,010.0	537,212.0	29,899
2	Summit - 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	Sierra - 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	Sunway TaihuLight - 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	Selene - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	555,520	63,460.0	79,215.0	2,646

上海第二工业大学

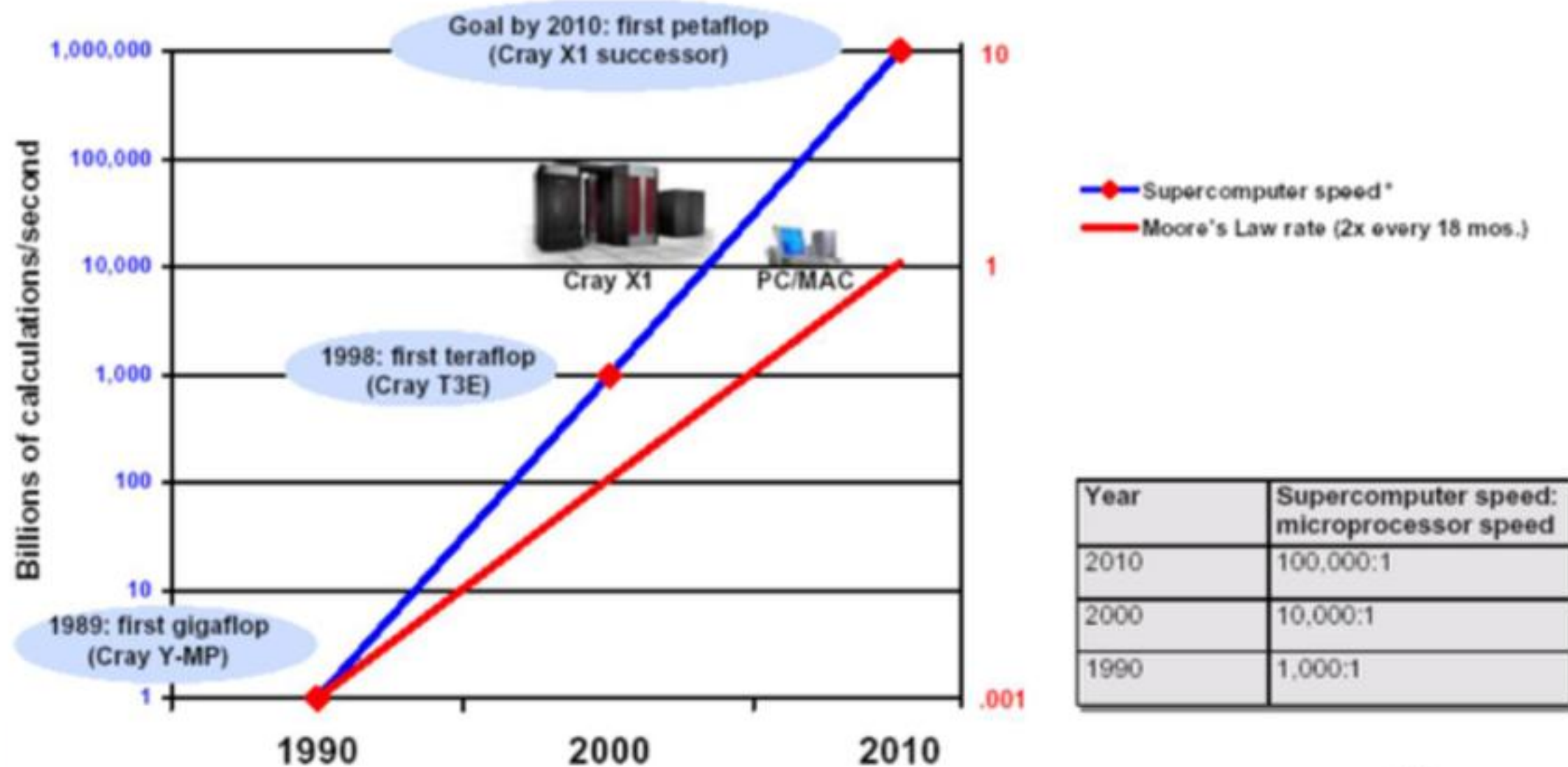
Summit is likely to reclaim the title for the United States this year. China is on track to unveil the first exascale computer in 2020.



上海第二工业大学

超级计算机的增长速度

——超过摩尔定律100倍



Cite from CRAY Inc.



Trends?

动物食物链



Trends?

1984计算机食物链



Mainframe



Vector Supercomputer



Mini Computer



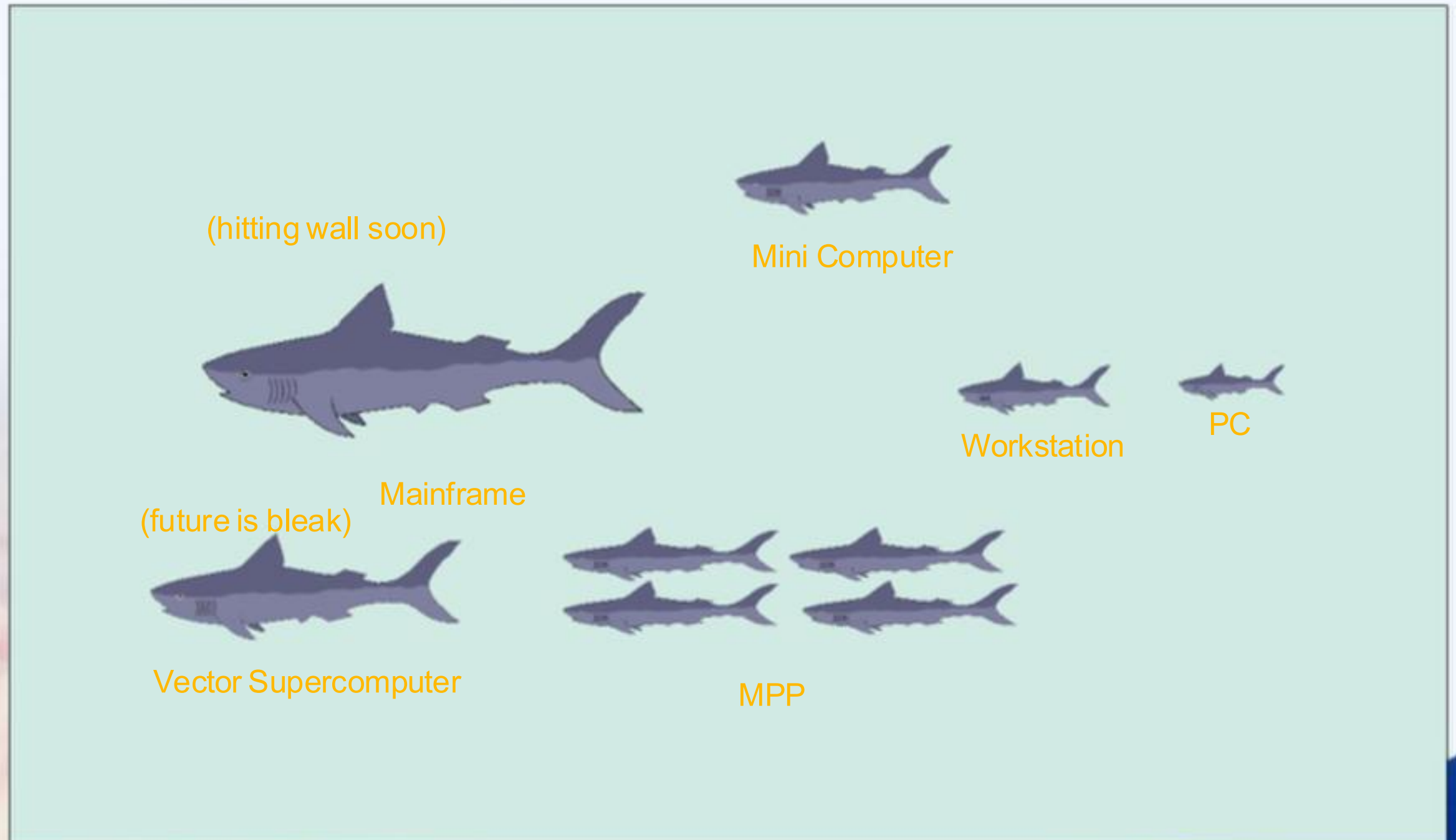
Workstation



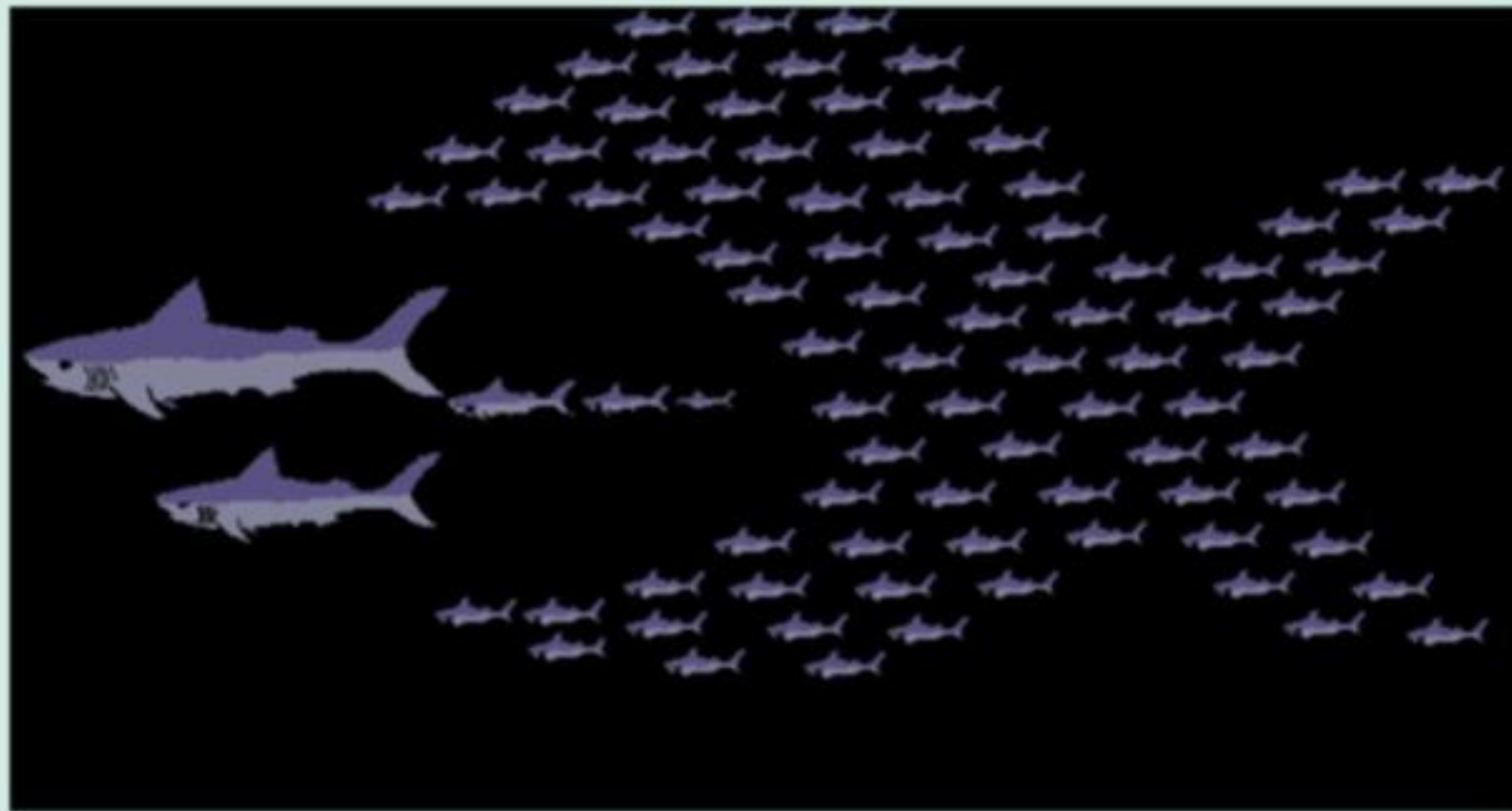
PC

Trends?

1994 计算机食物链



计算机食物链(现在和将来)



上海第二工业大学

Trends?



中国科学院计算技术研究所
Institute of Computing Technology, Chinese Academy of Sciences

[ENGLISH](#) | [新闻动态](#) | [关于我们](#) | [研究与开发](#) | [研究生培养](#) | [学术园地](#) | [技术转移与合作](#) | [创新文化](#) | [继续教育](#) | [招贤纳士](#) | [搜索引擎](#)

[返回首页](#) 

[学术园地](#) Academic Field

首页 > 学术园地 > 虚拟情报中心 > 信息技术快报

学术刊物
学术论文
虚拟情报中心
学会联接

[搜索](#) [高级](#)

 [留言板](#)



[信息技术快报](#) | [虚拟图书馆](#)

内部刊物

信息技术快报

Information Technology Letter

2003年第5期(总第5期)
出版日期:2003年9月29日

信报查询:

本期主题:

高性能计算机研究的现状与展望

樊建平 陈明宇

摘要 高性能计算机的研制受永无止境的探求复杂的物理世界与人类社会本身的应用计算需求的驱动及研制者所处环境(人员、经验、经费等)及当时的可选择实现使能技术的影响。回顾历史,任何时刻研制的最高性能的计算机总是服务于当时的科学计算的需求(材料模型、药物设计、气候模拟、核武器模拟、电磁学等)或者称是以科学计算为最初应用的靶子进行设计的(当前最快的日本Earth Simulation 与IBM BLUE/Gene 两个项目是很

Parallelism in Microprocessor

微处理器中的并行计算

- 抢占式或时间片轮转的多任务OS
- 指令级并行(Instruction-level Parallelism ILP): 乱序(out of order)执行、分支预测、指令多发射、硬件预取等;
- 同时多线程技术(simultaneous Multi-Threading, SMT): Intel公司实现的SMT技术就是超线程(Hyper-Threading, HT)技术, 超线程技术实际上只有一个物理处理器, 但从软件的角度来看, 存在多个逻辑处理器。
- 多核处理器技术: 采用单芯片多处理器(Chip Multiprocessor, CMP)的设计, 多核多线程(有别于单核上的超线程)

ILP、TLP与超标量技术（1）

- 超标量和多发射技术可以使得处理器在一个时钟周期内运行多条指令，可以分为：
 - 支持单线程的**ILP(Instruction-Level Parallelism)**处理器：超标量技术利用率不高。
 - 支持线程级并行的**TLP(Thread-Level Parallelism)**处理器：有较好的超标量技术利用率。
- **TLP**处理器分有：
 - 多线程处理器(**Multithreaded Processor**)：可以有效地减少垂直浪费情形。
 - 单片多处理器**CMP(Chip MultiProcessor)**：是指在单个芯片上的多个处理器核构成的多处理器系统，允许线程在多个处理器核上并行执行。
 - 同时多线程**SMT(Simultaneous Multithreaded)**处理器：允许在一个时钟周期内发射多个线程的多条指令执行。可以同时减少水平和垂直浪费。

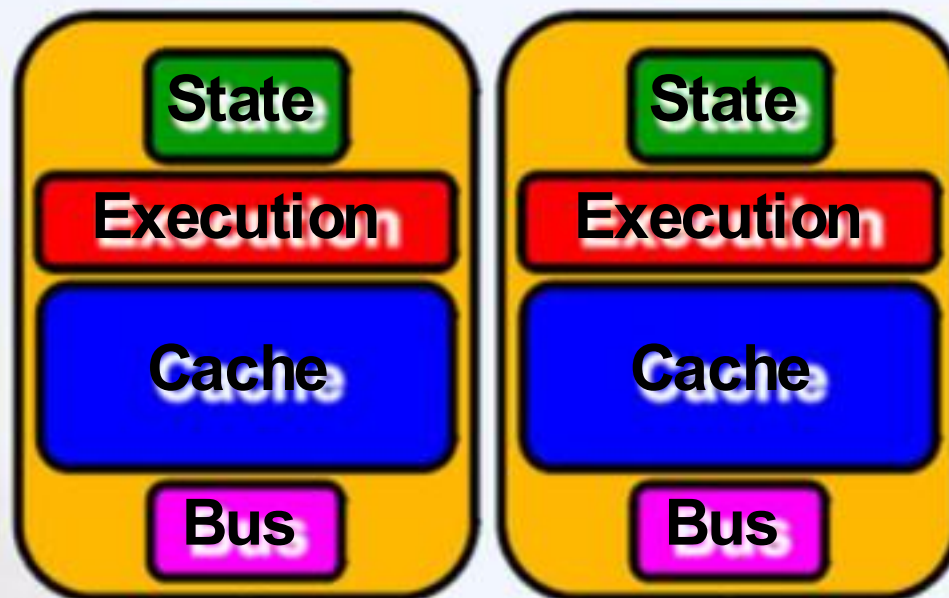
Intel多核技术的发展

**Pentium 4
with HT**



2 Threads
1 Package

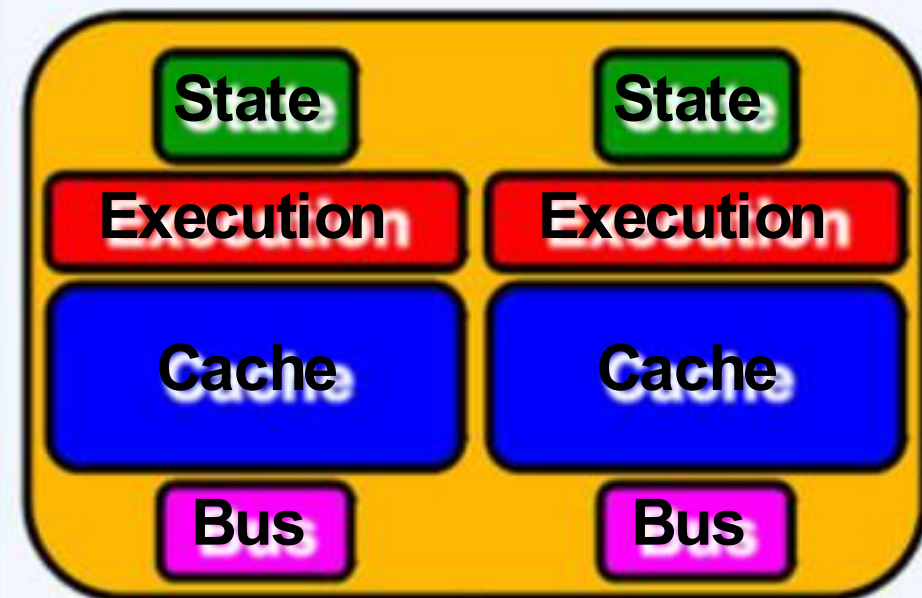
Dual Xeon Processors



2 Threads
2 Packages

Dual Core

*Two independent execution
cores in the same processor*



2 Threads
1 Package

几种不同的技术

- **HT – Hyper Threading:** 2 threads running on the same processor core
 - 处理器上的某些资源会被共享
 - 使用相同的缓存和运算器
- **DP/MP – Dual/Multi-Processing:** 2 or more processors in the same system
 - 只共享系统总线，独立缓存
 - 高性能，资源冲突少
- **DC – Dual Core:** 2 execution cores in the same processor package
 - 共享系统总线和缓存
 - 与双处理器的性能相同