并行计算
Parallel Computing

主讲 孙经纬

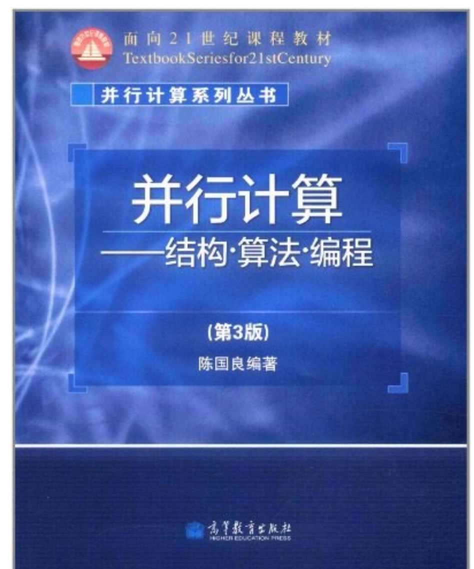
2024年 春季学期

概要

- **课程基本信息**
- 课程内容架构
- 并行计算简介

课程基本信息

- 授课教师：孙经纬 sunjw@ustc.edu.cn
- 助教：林俊卿 沈周毅
- 教材：《并行计算——结构·算法·编程》，陈国良 编著，高等教育出版社，2011



▪ 教材作者：陈国良教授，中科院院士，国家高性能中心（合肥）主任，第一届全国高等学校“国家级教学名师奖”、2009年国家教学成果二等奖获得者

▪ 并行算法教学体系：并行算法的设计与分析、并行计算机体系结构、并行算法实践、并行计算

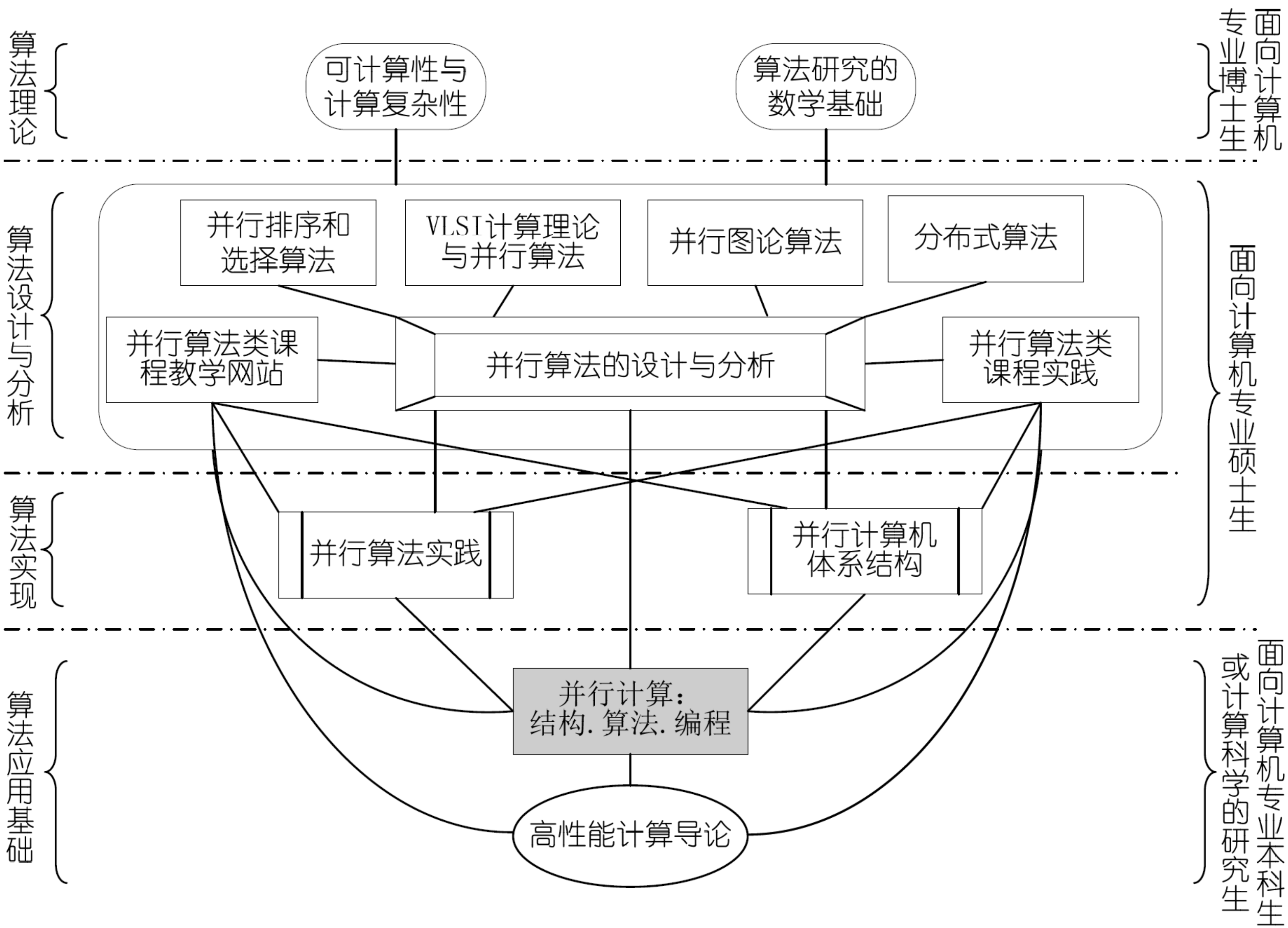
课程基本信息

- 评分结构
  - 期末考试（闭卷） 60%
  - 书面作业（随堂2次，课后4次） 15%
  - 课程实验（一共4题） 25%

概要

- 课程基本信息
- **课程内容架构**
- 并行计算简介

课程内容架构



算法理论
可计算性与计算复杂性
算法研究的数学基础
面向计算机专业博士生
算法设计与分析
并行排序和选择算法
VLSI计算理论与并行算法
并行图论算法
分布式算法
并行算法类课程教学网站
并行算法的设计与分析
并行算法类课程实践
面向计算机专业硕士生
算法实现
并行算法实践
并行计算机体系结构
算法应用基础
并行计算：结构. 算法. 编程
高性能计算导论
面向计算机专业本科生或计算科学的研究生

课程内容架构

- 并行计算
  - 第一部分：并行体系结构 Parallel Architectures
  - 第二部分：并行算法 Parallel Algorithms
  - 第三部分：并行编程 Parallel Programming

第一部分：并行体系结构

- 并行计算系统的硬件结构
  - 系统结构与模型 System Architectures and Models
  - 系统互联 System Interconnections
  - 性能评测 Performance Evaluation

第二部分： 并行算法

- 并行计算的理论基础
  - 计算模型 Computational Models
  - 设计策略 Design Policy
  - 设计方法学 Design Methodology
  - 设计技术 Design Techniques
  - 并行数值算法Parallel Numerical Algorithms

第三部分：并行编程

- 并行计算的支撑软件:
  - 编程模型 Programming Models
  - 共享内存编程 Shared-Memory Programming
  - 消息传递编程 Message-Passing Programming
  - 数据并行编程 Data-Parallel Programming
  - 编程环境与工具 Programming Environment and Tools
  - GPU编程

概要

- 课程基本信息
- 课程内容架构
- **并行计算简介**

什么是并行计算?

"A parallel computer **is a collection of processing elements** that cooperate and communicate to solve large problems **fast**" — Almasi and Gottlieb, Highly Parallel Computing, 1989

什么是并行计算？

UC Berkeley 2006年关于并行计算的技术报告

https://people.eecs.berkeley.edu/~krste/papers/BerkeleyView.pdf

The Landscape of Parallel Computing Research: A View from Berkeley

Krste Asanovíc, Rastislav Bodik, Bryan Catanzaro, Joseph Gebis, Parry Husbands, Kurt Keutzer, David Patterson, William Plishker, John Shalf, Samuel Williams, and Katherine Yelick

EECS Technical Report UCB/EECS-2006-183

December 18, 2006

Abstract

The recent switch to parallel microprocessors is a milestone in the history of computing. Industry has laid out a roadmap for multicore designs that preserves the programming paradigm of the past via binary compatibility and cache coherence. Conventional wisdom is now to double the number of cores on a chip with each silicon generation.

A multidisciplinary group of Berkeley researchers met nearly two years to discuss this change. Our view is that this evolutionary approach to parallel hardware and software may work from 2 or 8 processor systems, but is likely to face diminishing returns as 16 and 32 processor systems are realized, just as returns fell with greater instruction-level parallelism.

什么是并行计算？

- 伯克利报告总结了并行计算领域关注的7个关键问题
- 其中，应用（Applications）和硬件（Hardware）是两大重点

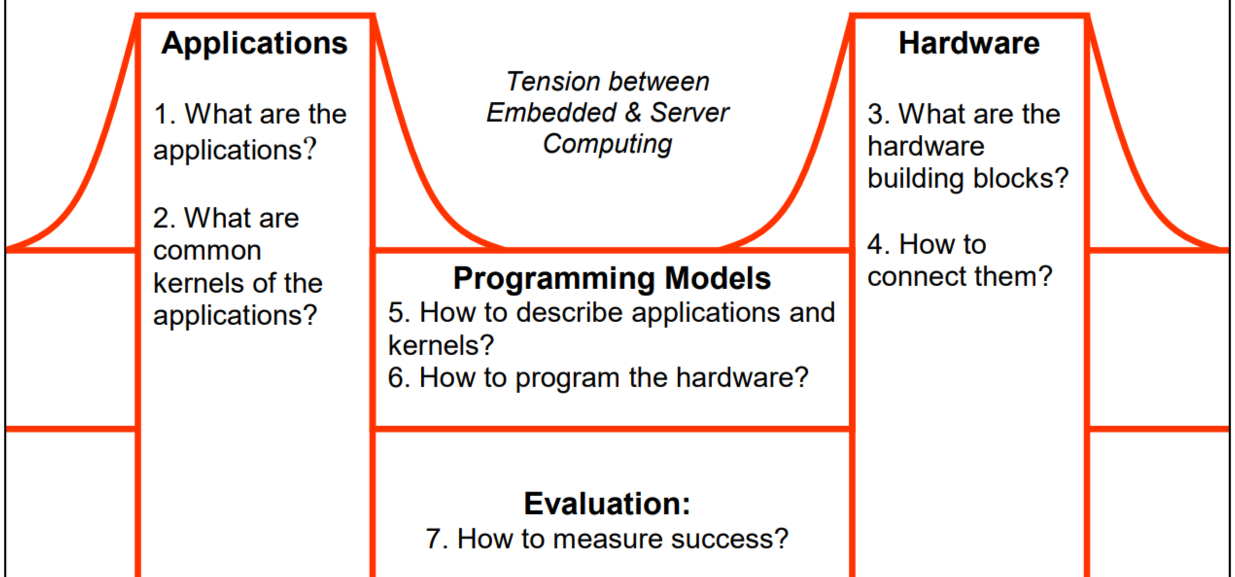
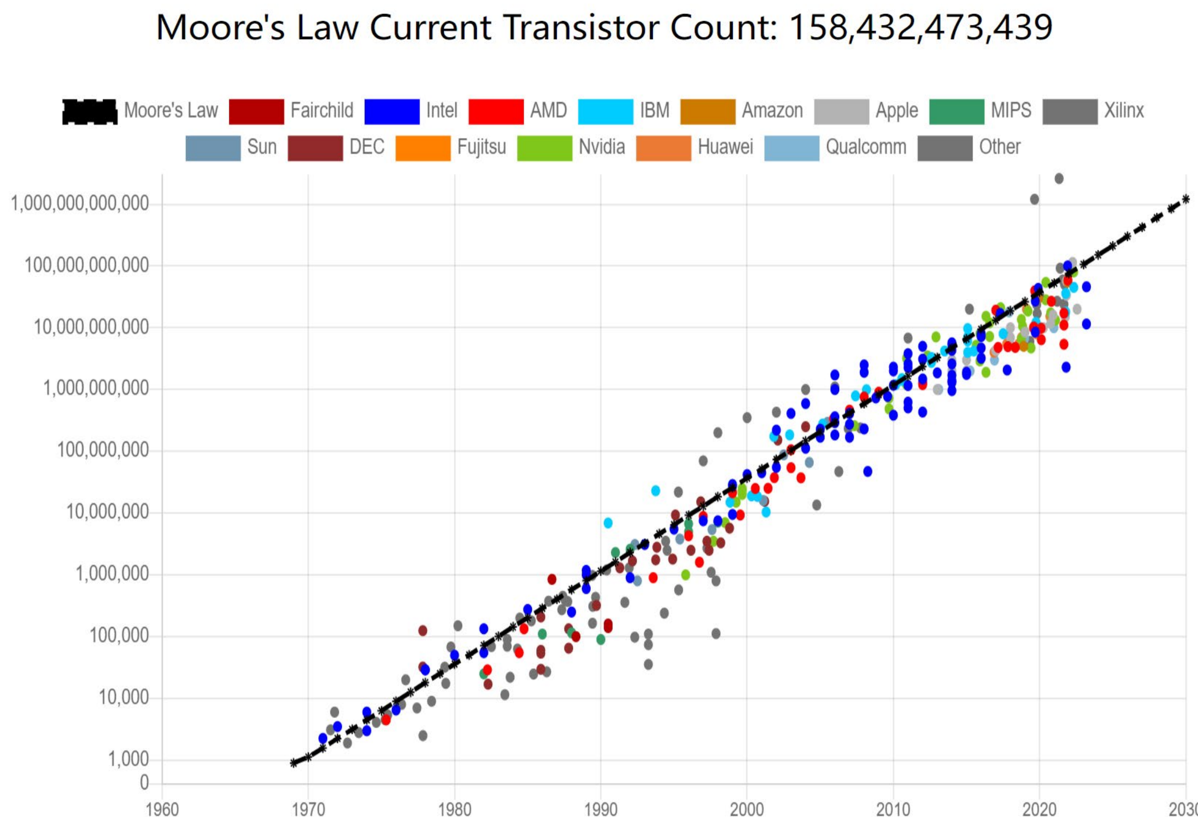


Figure 1. A view from Berkeley: seven critical questions for 21st Century parallel computing. (This figure is inspired by a view of the Golden Gate Bridge from Berkeley.)

为什么需要并行计算：硬件视角

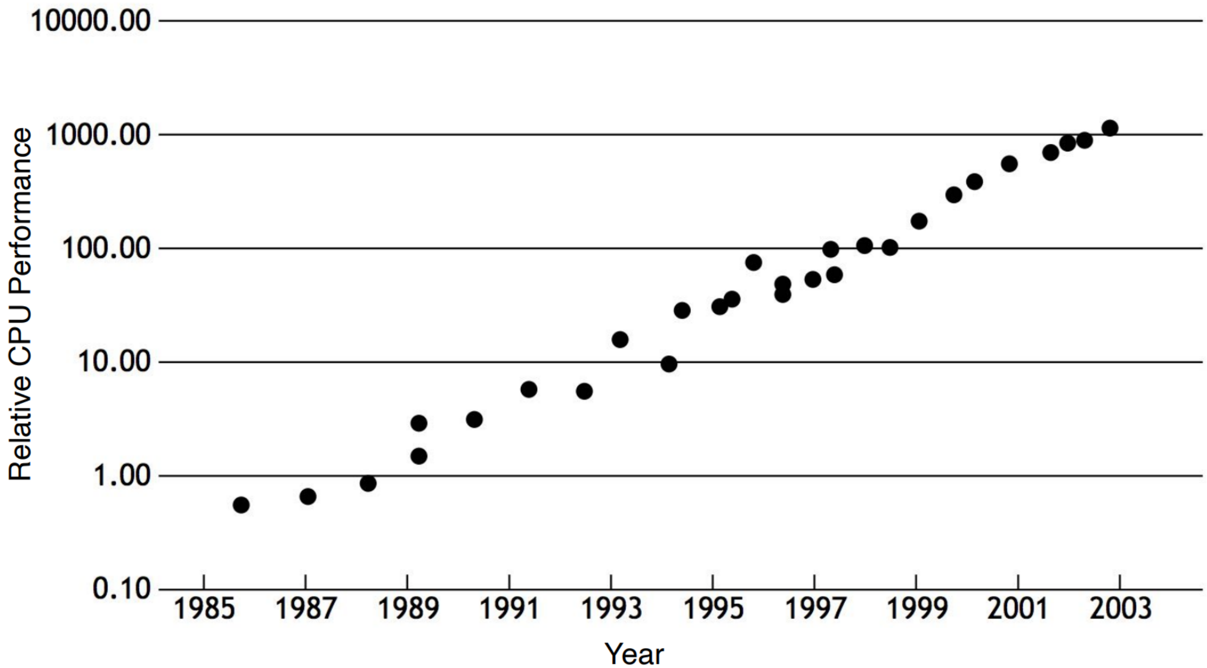
- 摩尔定律
- 在不同的发展阶段，决定CPU性能提升的关键因素是不同的



http://www.transistorcount.com/

为什么需要并行计算：硬件视角

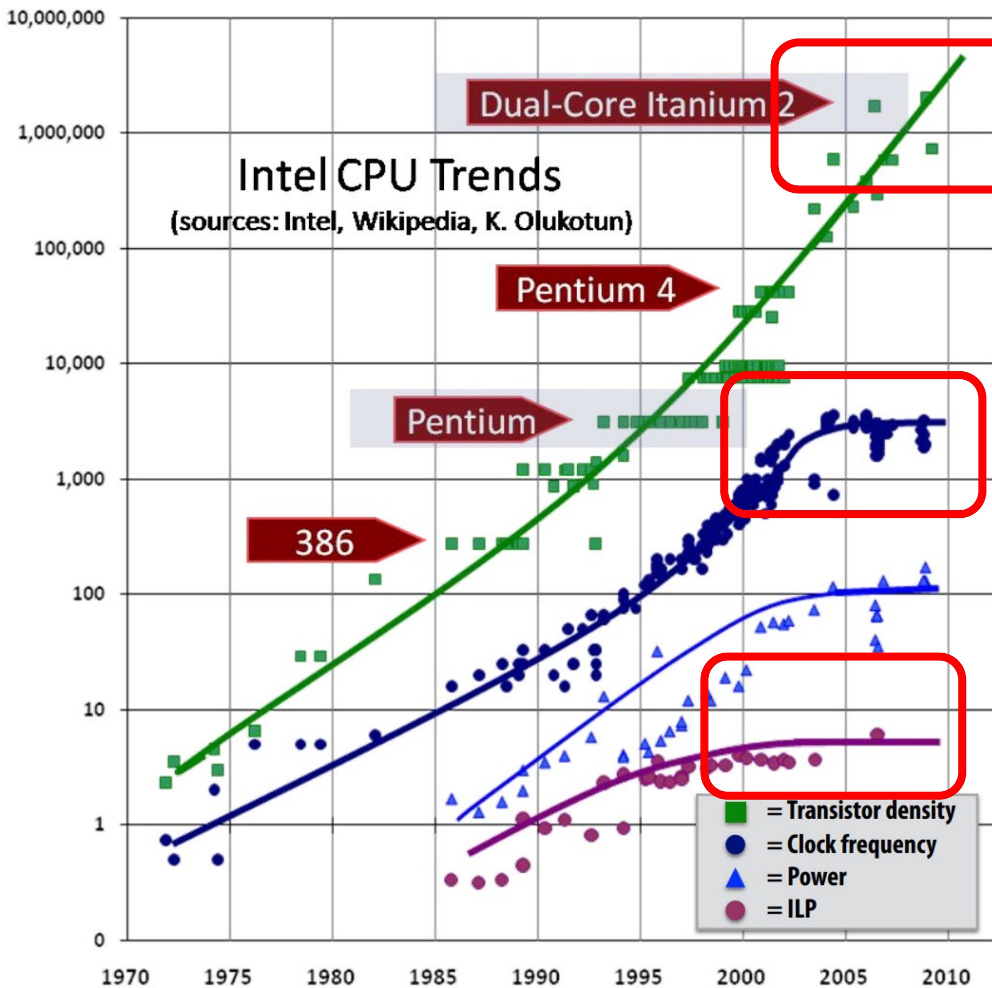
1985-2003 单核CPU性能



在大约2005年之前，CPU性能提升主要来源于：

- 主频提升
- 指令级并行（Instruction-Level Parallelism，ILP）

为什么需要并行计算：硬件视角



免费午餐结束了，
通过并行计算来
进一步提升

主频难以继续提升

指令级并行带来
的收益逐渐减少

The Free Lunch Is Over — A Fundamental Turn Toward Concurrency in Software. By Herb Sutter 2005

为什么需要并行计算：应用视角

- 科学研究的三类方法

  实验，理论，模拟仿真（计算）

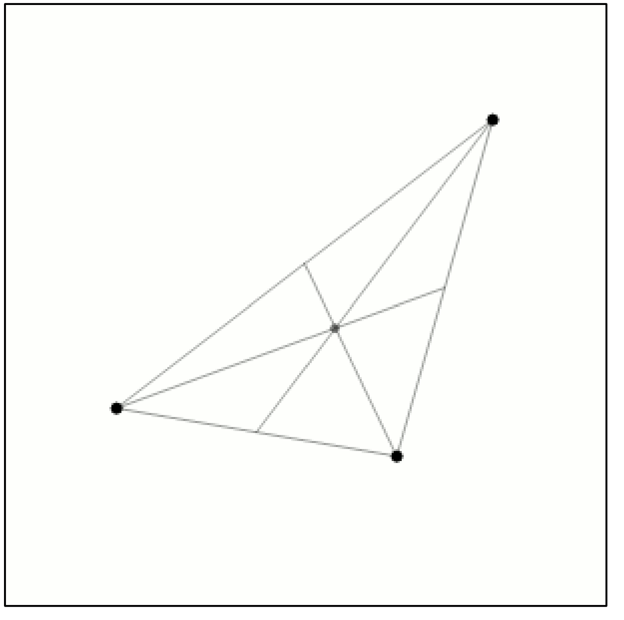
- 在许多情况下
  - 理论模型复杂甚至理论尚未建立
  - 实验开销昂贵甚至无法进行
  - 此时，计算就成了求解科学问题的主要手段

为什么需要并行计算：应用视角

- Example: 三体问题

The mathematical statement of the three-body problem can be given in terms of the Newtonian equations of motion for vector positions $\mathbf{r_i} = (x_i, y_i, z_i)$ of three gravitationally interacting bodies with masses m_i:

$$\ddot{\mathbf{r}}_\mathbf{1} = -Gm_2 \frac{\mathbf{r_1} - \mathbf{r_2}}{|\mathbf{r_1} - \mathbf{r_2}|^3} - Gm_3 \frac{\mathbf{r_1} - \mathbf{r_3}}{|\mathbf{r_1} - \mathbf{r_3}|^3},$$
$$\ddot{\mathbf{r}}_\mathbf{2} = -Gm_3 \frac{\mathbf{r_2} - \mathbf{r_3}}{|\mathbf{r_2} - \mathbf{r_3}|^3} - Gm_1 \frac{\mathbf{r_2} - \mathbf{r_1}}{|\mathbf{r_2} - \mathbf{r_1}|^3},$$
$$\ddot{\mathbf{r}}_\mathbf{3} = -Gm_1 \frac{\mathbf{r_3} - \mathbf{r_1}}{|\mathbf{r_3} - \mathbf{r_1}|^3} - Gm_2 \frac{\mathbf{r_3} - \mathbf{r_2}}{|\mathbf{r_3} - \mathbf{r_2}|^3}.$$



https://en.wikipedia.org/wiki/Three-body_problem

为什么需要并行计算：应用视角

OAK RIDGE NATIONAL LABORATORY

Instituto Nacional de Meteorologia España

Government-Classified Work

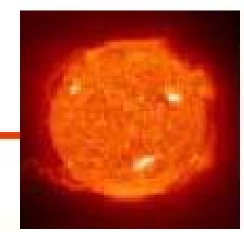
Government - Research

(Severe) Weather Prediction
& Climate Modeling

Automotive Design & Safety

Drug Discovery &
Genomic Research

Aircraft/Spacecraft Design
& Fuel-Efficiency

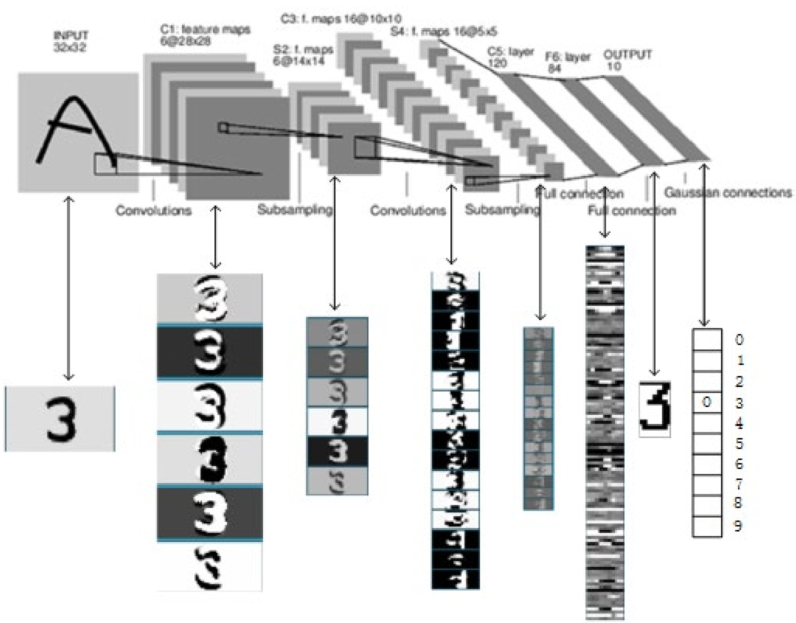


Oil Exploration &
Energy Research

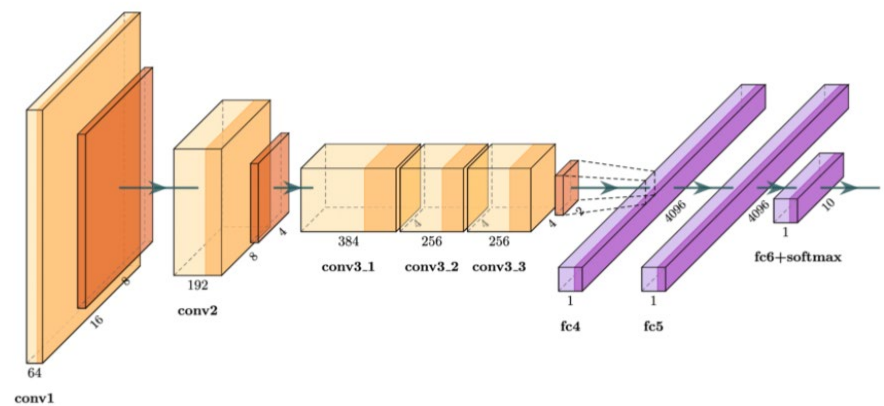
Basic Scientific Research

计算和性能需求是无止境的

- 硬件：满足不断增长的性能需求
- 应用：实现更大规模或更精细、精确的计算



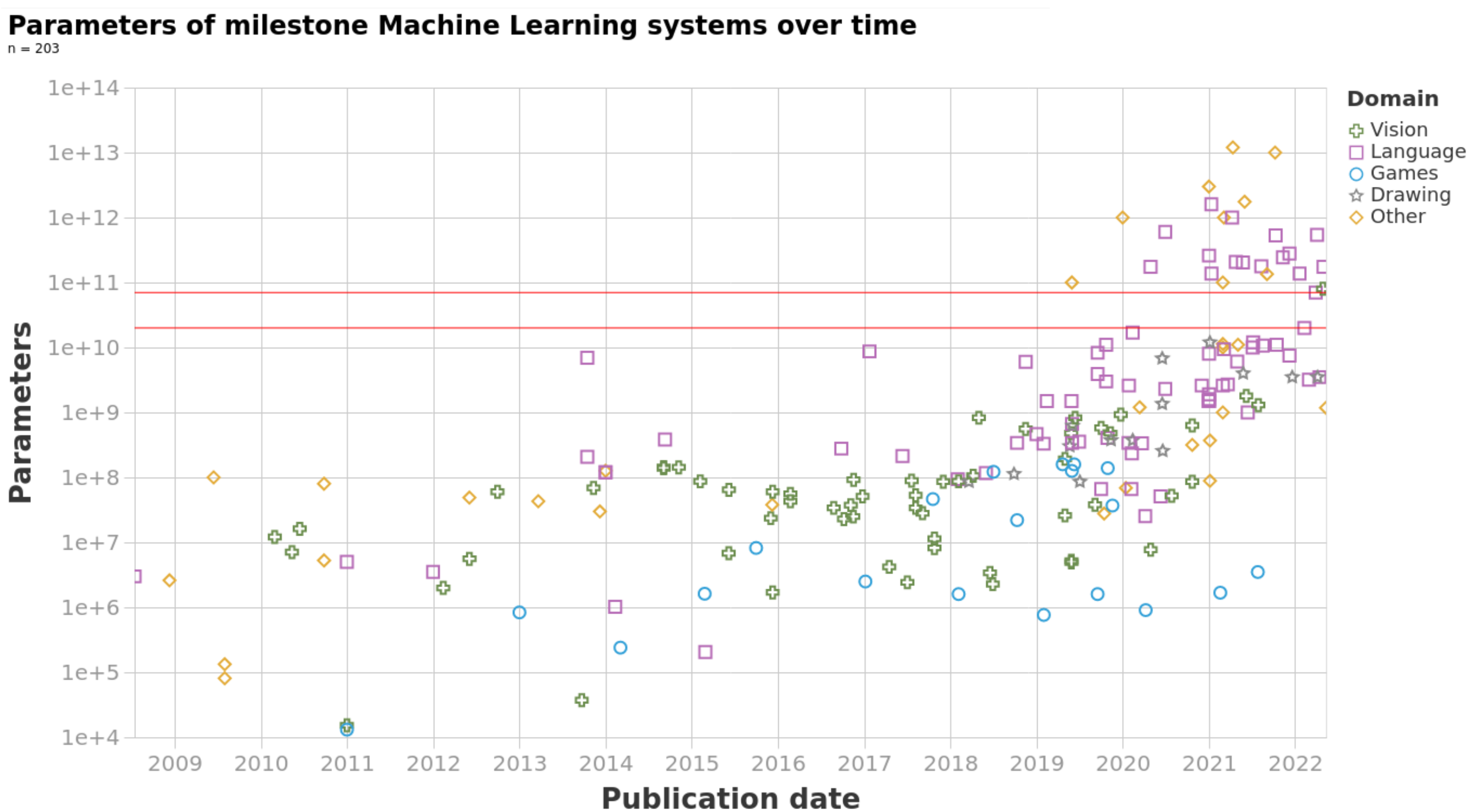
LeNet手写数字识别
60k参数, 1989年



AlexNet在ImageNet上达到15.3% top-5 error
62.3m参数，2012年

计算和性能需求是无止境的

Andy and Bill's law： What Andy Giveth, Bill Taketh Away



*Pablo Villalobos, Jaime Sevilla, Tamay Besiroglu, Lennart Heim, Anson Ho, Marius Hobbhahn: Machine Learning Model Sizes and the Parameter Gap. CoRR abs/2207.02852 (2022) http://www.top500.org

TOP 500
The List.

HOME | LISTS | STATISTICS | RESOURCES | ABOUT | MEDIA KIT

Home »Lists »TOP500 »November 2023

NOVEMBER 2023

The 62nd edition of the TOP500 reveals that the Frontier system retains its top spot and is still the only exascale machine on the list. However, five new or upgraded systems have shaken up the Top 10.

Housed at the Oak Ridge National Laboratory (ORNL) in Tennessee, USA, Frontier leads the pack with an HPL score of 1.194 EFlop/s – unchanged from the June 2023 list. Frontier utilizes AMD EPYC 64C 2GHz processors and is based on the latest HPE Cray EX235a architecture. The system has a total of 8,699,904 combined CPU and GPU cores. Additionally, Frontier has an impressive power efficiency rating of 52.59 GFlops/watt and relies on HPE's Slingshot 11 network for data transfer.

The new Aurora system at the Argonne Leadership Computing Facility in Illinois, USA, entered the list at the No. 2 spot – previously held by Fugaku – with an HPL score of 585.34 PFlop/s. That said, it is important to note that Aurora's numbers were submitted with a measurement on half of the planned final system. Aurora is currently being commissioned and will reportedly exceed Frontier with a peak performance of 2 EFlop/s when finished.

TOP500 Release

THE LIST
PRESS RELEASE
LIST HIGHLIGHTS

Statistics

PERFORMANCE DEVELOPMENT
SUBLIST GENERATOR
LIST STATISTICS
TREE MAPS
HISTORICAL CHARTS

Downloads

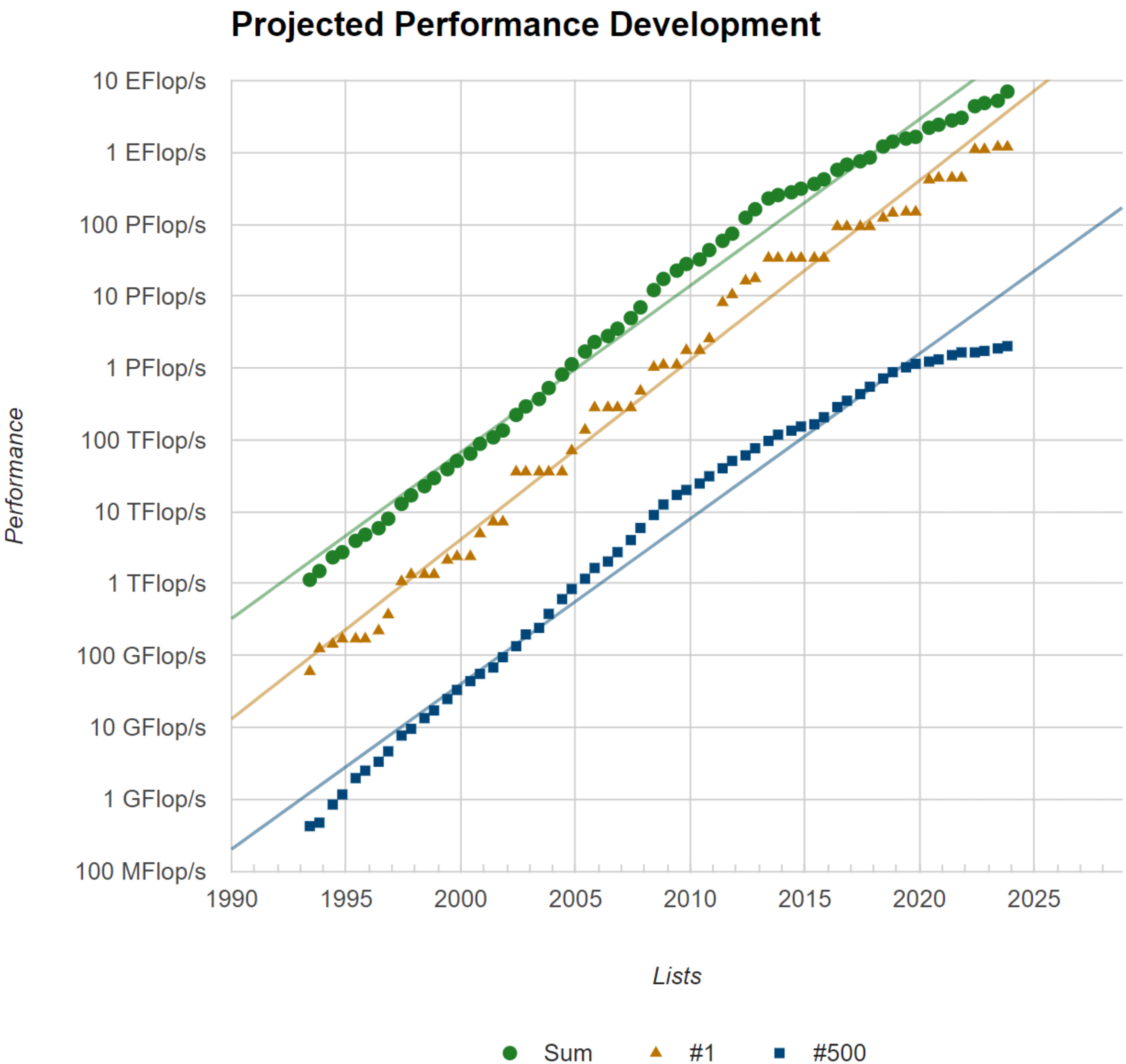
TOP500 LIST (XML)
TOP500 LIST (EXCEL)
GREEN500 LIST (EXCEL)

每年更新两次

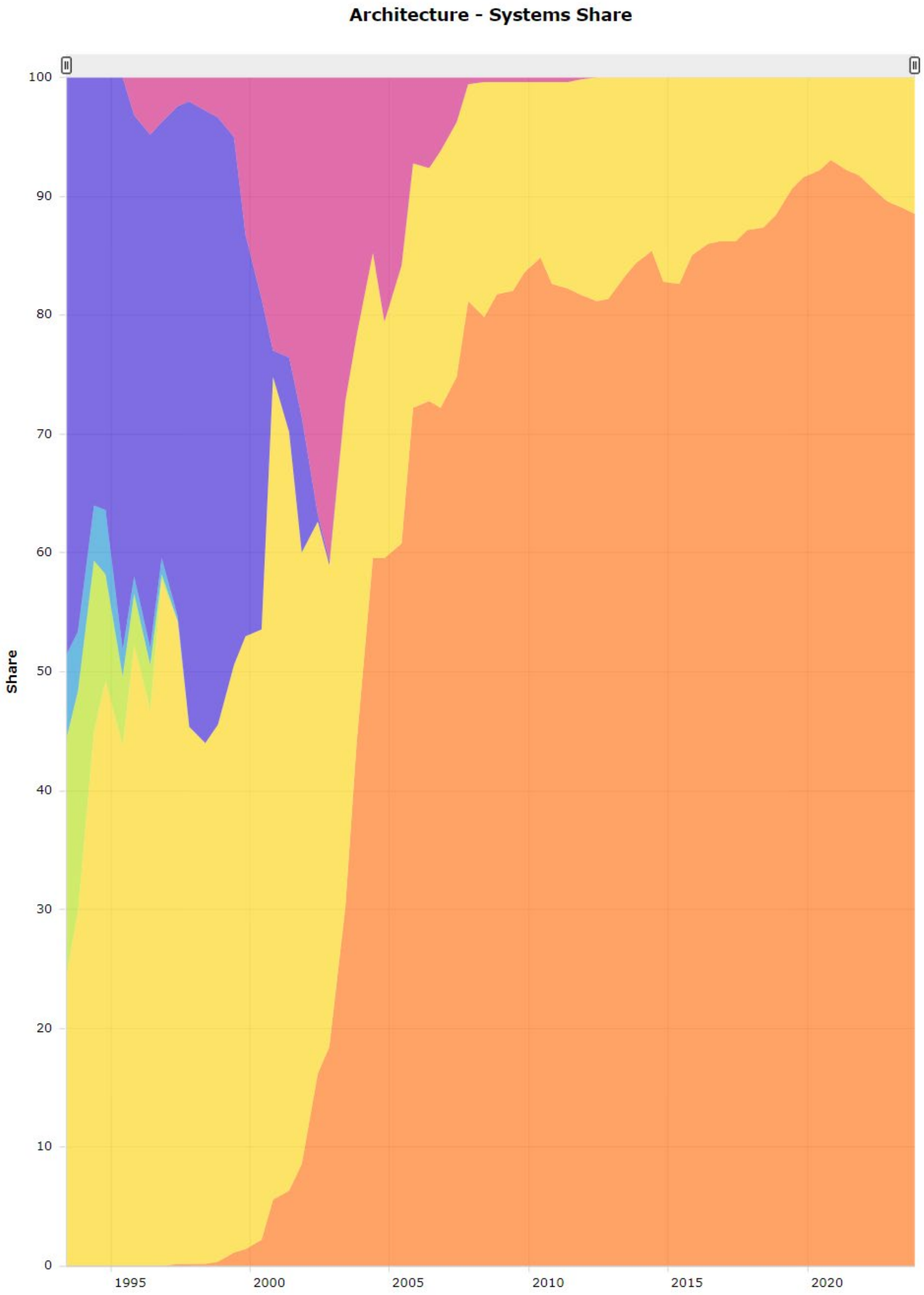
- 十一月：SC（美国）
- 六月：ISC（欧洲）

自1993年至今，2023年11月是第62版

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	**Frontier** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States	8,699,904	1,194.00	1,679.82	22,703
2	**Aurora** - HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11, Intel DOE/SC/Argonne National Laboratory United States	4,742,808	585.34	1,059.33	24,687
3	**Eagle** - Microsoft NDv5, Xeon Platinum 8480C 48C 2GHz, NVIDIA H100, NVIDIA Infiniband NDR, Microsoft Microsoft Azure United States	1,123,200	561.20	846.84	
4	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442.01	537.21	29,899
5	**LUMI** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE EuroHPC/CSC Finland	2,752,704	379.70	531.51	7,107


Projected Performance Development
10 EFlop/s
1 EFlop/s
100 PFlop/s
10 PFlop/s
1 PFlop/s
100 TFlop/s
10 TFlop/s
1 TFlop/s
100 GFlop/s
10 GFlop/s
1 GFlop/s
100 MFlop/s
Performance
1990
1995
2000
2005
2010
2015
2020
2025
Lists
Sum
#1
#500

TOP500 Treemaps (Countries)



TOP500中不同架构占比演变（1993-2023）

Cluster
MPP
SMP
Constellations
Single Processor
SIMD

中国高性能计算机性能TOP100排行榜

首 页 / Home | 排 行 榜 / Lists | 新闻 / News | 工作条例 / | 关于 / About | 联系我们 / Contact

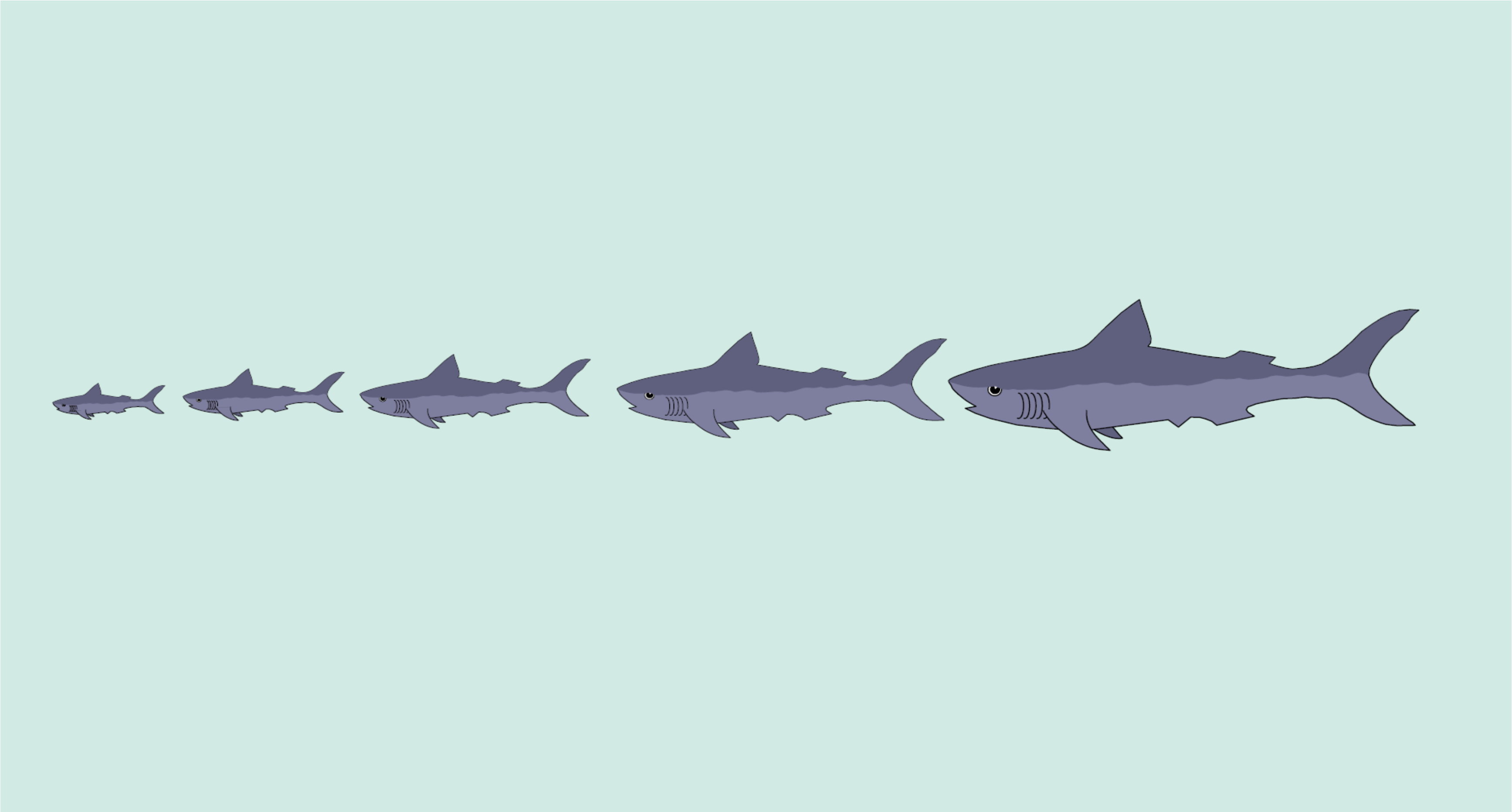
2023年中国高性能计算机性能TOP100排行榜

排名	研制厂商/单位	型号	安装地点	安装年份	应用领域	CPU核数	Linpack值(Pflops)	Linpack来源	Linpack峰值(Pflops)	效率(%)
1	服务器供应商	超算中心主机系统,异构众核处理器	超算中心	2023	算力服务	15974400	487.94	Q	620.0	78.7
2	服务器供应商	网络公司主机系统,CPU+GPU异构众核处理器	网络公司	2022	算力服务	460000	208.26	Q	390.0	53.4
3	服务器供应商	网络公司主机系统,CPU+GPU异构众核处理器	网络公司	2021	算力服务	285000	125.04	Q	240.0	52.1
4	国家并行计算机工程技术研究中心	神威太湖之光,40960*Sunway SW26010 260C 1.45GHz,自主网络	国家超级计算无锡中心	2016	超算中心	10649600	93.015	Q	125.436	74.2
5	服务器供应商	网络公司主机系统,CPU+GPU异构众核处理器	网络公司	2021	算力服务	190000	87.04	Q	160.0	51.2
6	国防科大	天河二号升级系统(Tianhe-2A),TH-IVB-MTX Cluster + 35584*Intel Xeon E5-2692v2 12C 2.2GHz+35584*Matrix-2000,TH Express-2	国家超级计算广州中心	2017	超算中心	427008	61.445	Q	100.679	61.0

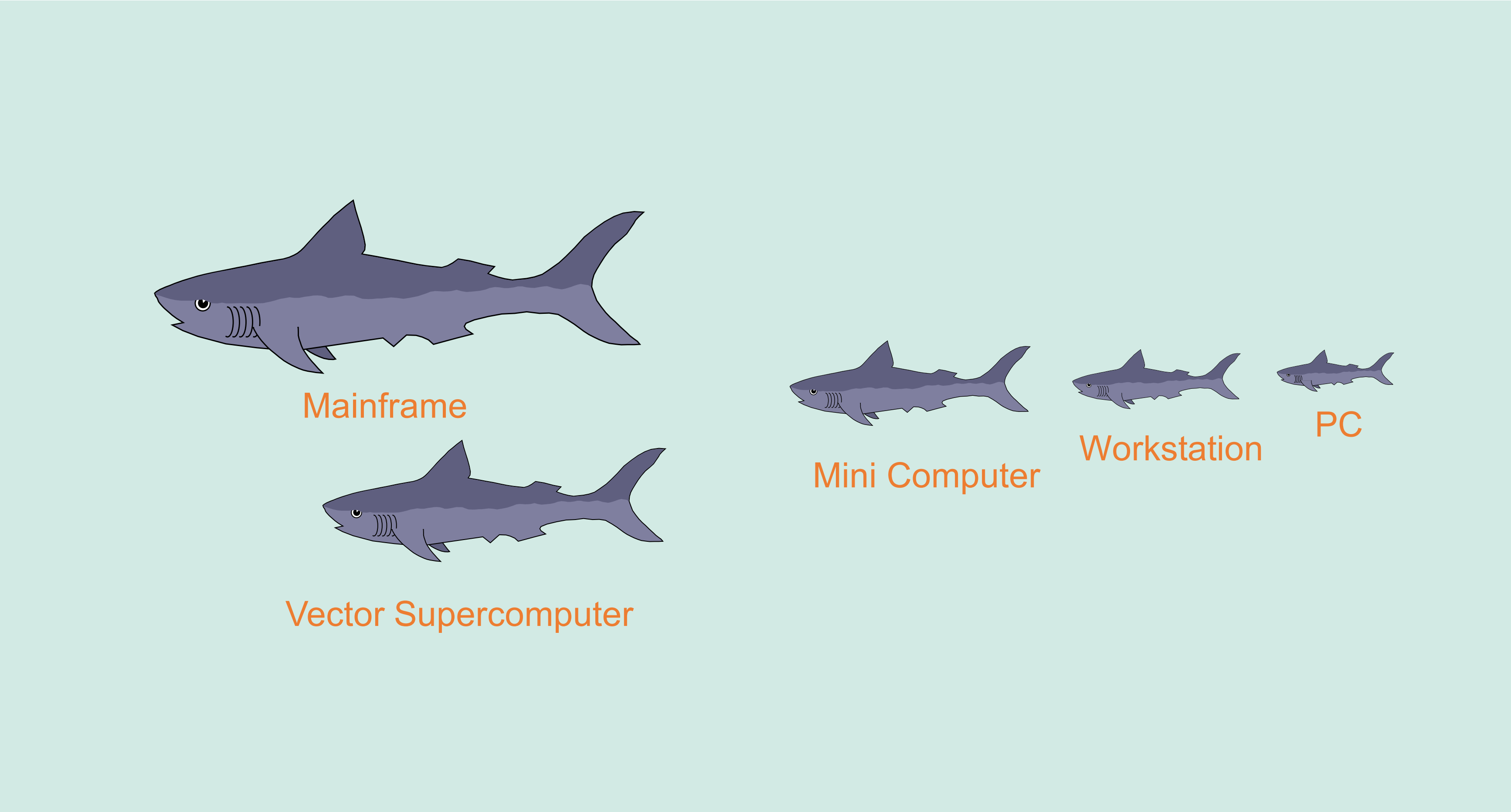
http://www.hpc100.cn/

http://now.cs.berkeley.edu

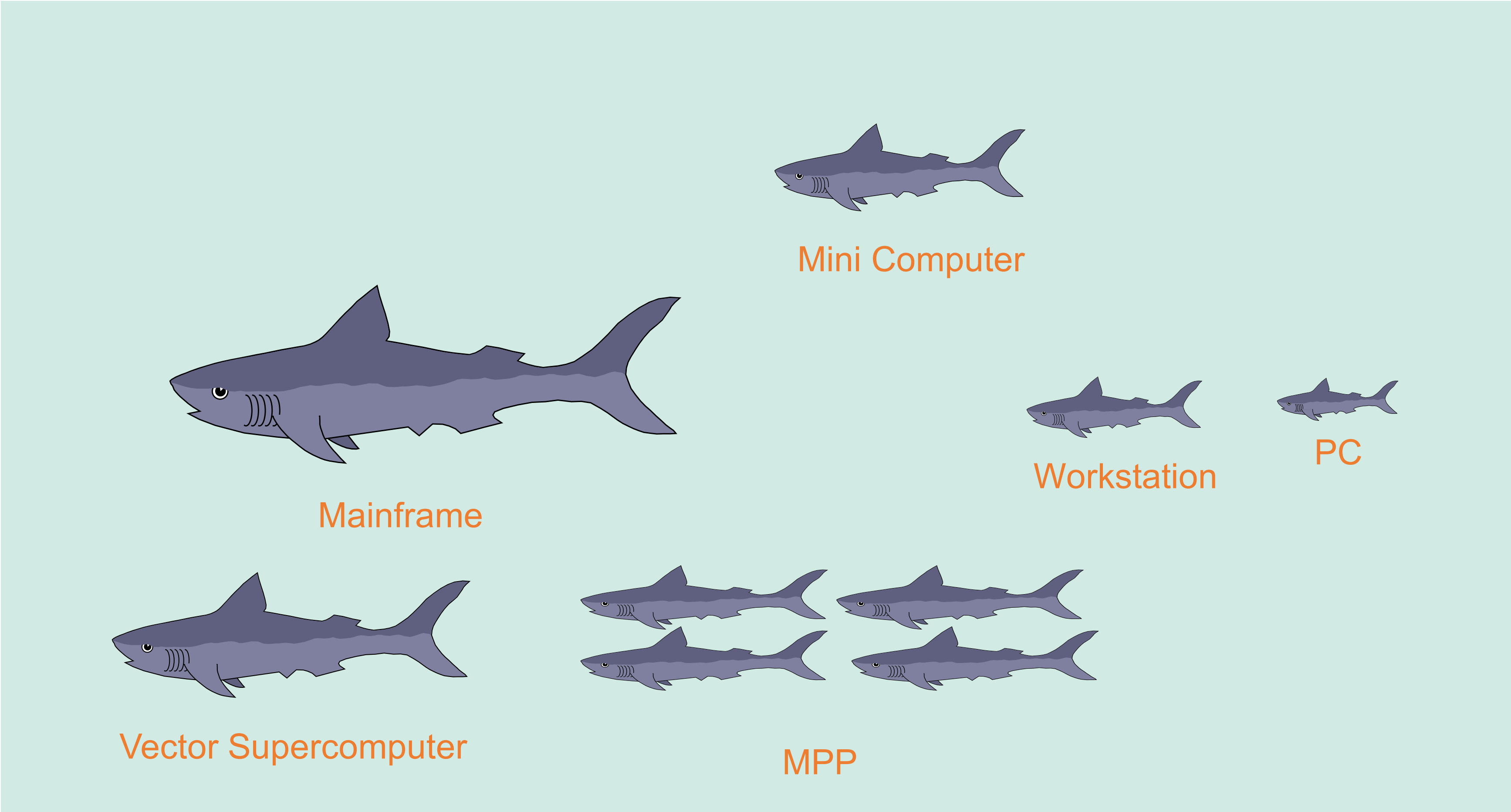
Food chain



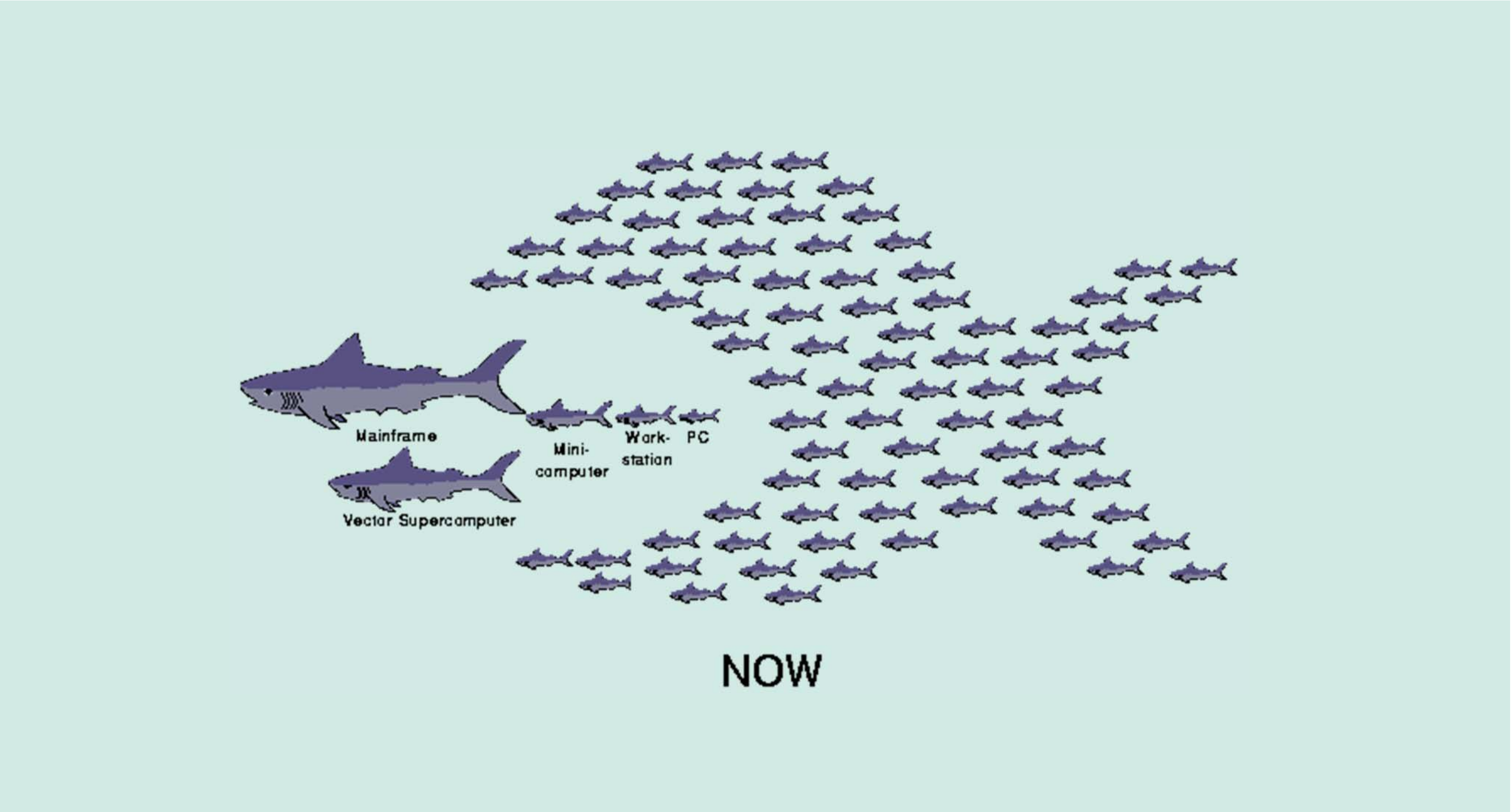
1984 computer food chain



1994 computer food chain



NOW and future



NOW: Networks of Workstations， 是一种典型的Cluster of Workstations (COW)