



南開大學  
Nankai University

计算机学院  
并行程序设计实验报告

并行体系结构调研

姓名：林盛森

学号：2312631

专业：计算机科学与技术

2025 年 3 月 15 日

# 目录

<b>1 天河超级计算机发展历史</b>	<b>2</b>
1.1 天河一号	2
1.1.1 基本介绍	2
1.1.2 历史背景	2
1.1.3 系统配置	2
1.1.4 应用领域与典型案例	3
1.2 天河二号	3
1.2.1 基本介绍	3
1.2.2 系统配置	3
1.2.3 技术创新	4
1.2.4 应用领域与典型案例	4
1.3 天河三号	5
1.3.1 基本介绍	5
1.3.2 系统配置	5
1.3.3 技术创新	6
<b>2 “天河一号”并行体系结构分析</b>	<b>6</b>
2.1 系统配置分析	6
2.1.1 性能参数	6
2.1.2 计算节点配置	7
2.1.3 网络与系统软件	7
2.2 并行体系结构分析	7
2.2.1 节点级并行: CPU+GPU 异构节点	7
2.2.2 节点间并行: 7168 个异构节点自主高速互联	8
2.2.3 数据级并行: Lustre 分布式文件系统	8
2.2.4 任务级并行: SLURM 作业调度系统	8
2.2.5 总结	8
<b>3 苹果 M4CPU 分析</b>	<b>8</b>
3.1 功能综述	8
3.2 CPU 分析	9

# 1 天河超级计算机发展历史

## 1.1 天河一号

### 1.1.1 基本介绍

“天河一号”超级计算机是中国自主研发的第一个千万亿次级计算能力的超级计算机。“天河一号”超级计算机从 2008 年开始研制，由国防科学技术大学与滨海新区联合研制，至 2010 年 9 月在国家超级计算天津中心研制成功，并采用 CPU 和 GPU 相结合的异构融合计算体系结构，是我国重要的公共科技服务平台。2010 年 11 月 14 日，国际 TOP500 组织在网站上公布了最新全球超级计算机前 500 强排行榜，中国首台千万亿次超级计算机系统“天河一号”排名全球第一。作为中国超算史上的里程碑，天河一号不仅打破了国外技术垄断，更孕育了自主可控的超算生态体系，为后续天河系列的发展奠定了基础。

### 1.1.2 历史背景

21 世纪初，美国、日本、欧洲通过《出口管制条例》限制向中国出口高性能计算机及相关技术。截至 2008 年，中国 90% 的超算依赖进口，最高性能仅为 30 TFLOPS。面对国际技术封锁，科研团队只能用二手 Intel 芯片拼装出天河一号，但其性能仍跻身全球前五。最让人自豪的是，它首次用上了自主研发的“太赫兹光通信”网络，让千万个个计算节点能像蜂群协作一样高速通信，打破了国外对高速互连技术的垄断。当时，天河一号就能够出神入化地预测台风的路径，把原本 30 公里的误差缩小到 10 公里，让沿海城市能提前 12 小时发布预警。

### 1.1.3 系统配置

主机	TH-1A
性能	峰值性能4700TFlops (LINPACK实测值 2566TFlops)
存储	4PB
近线存储	1.3PB
计算节点	共7168个CPU+GPU异构节点 节点配置：双路12核，2*Intel Xeon X5670 @ 2.93GHz + Tesla M2050，内存24GB/48GB
登陆节点	5个登陆节点，其中2个专用于数据传输
网络系统	天河自主高速互联
操作系统	REHT
文件系统	Lustre
作业调度系统	SLURM

([\[1\]](#) 图片来源：国家超级计算天津中心官网)

“天河一号”超级计算机峰值性能每秒 4700 万亿次、持续性能每秒 2566 万亿次 (LINPACK 实测值)，由 6144 个 CPU 和 5120 个 GPU 装在 140 个机柜组成，占地面积近 700 平方米，总重量达到 155 吨。该机采用了中国自主研发的 3 款大规模集成电路芯片，4 类结点机、2 套网络、15 种印制电路板，及操作系统、编译系统、并程序开发环境与科学计算可视化系统。在异构融合体系结构、64 位多核多

线程自主 CPU、基于高阶路由高速互连通信、多级并行编译优化、高性能虚拟计算域、软硬一体的低功耗控制等方面，实现了一系列新的重大技术突破。

#### 1.1.4 应用领域与典型案例

##### (1) 科学计算

。中科院大气物理所：利用天河一号建立“台风路径预测模型”，将台风“海葵”（2010 年）路径预测误差从 30 公里降至 10 公里，预警时间提前 12 小时，减少直接经济损失超 50 亿元。

。青藏高原冰川模拟：首次实现百万平方公里区域、百年尺度的气候演变模拟，数据精度达  $0.1^{\circ}\text{C}/\text{km}^2$ 。

##### (2) 生命科学

。武汉大学：完成人类基因组全序列分析（30 亿碱基对），比对速度较国际最快水平提升 4 倍，数据存储需求降低 60%。

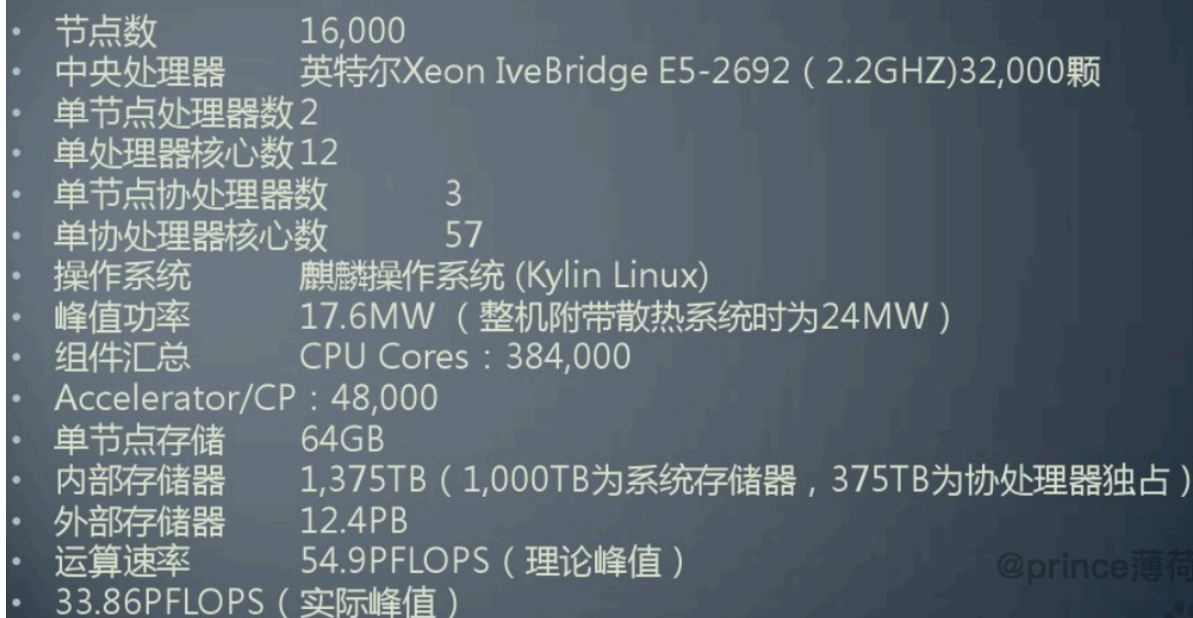
。艾滋病病毒耐药性研究：筛选出 3 种新型抗逆转录病毒药物，将临床试验周期缩短 50%。

## 1.2 天河二号

### 1.2.1 基本介绍

2013 年下半年，“天河二号”超级计算机在国家超级计算广州中心已正式投入运行。2013 年 6 月 17 日，天河二号再次登顶世界超算之巅，从天河一号 4.7 千兆次到天河二号  $5.49 \times 10$  的 16 次方次，实现了超级计算机由千兆（ $10 \times 10$  的 15 次方）次级迈入万兆（ $10 \times 10$  的 16 次方）次级计算速度，并且在生物医药、新材料等前端领域为人民提供便利与服务。

### 1.2.2 系统配置



• 节点数	16,000
• 中央处理器	英特尔Xeon IvyBridge E5-2692 ( 2.2GHZ)32,000颗
• 单节点处理器数	2
• 单处理器核心数	12
• 单节点协处理器数	3
• 单协处理器核心数	57
• 操作系统	麒麟操作系统 (Kylin Linux)
• 峰值功率	17.6MW ( 整机附带散热系统时为24MW )
• 组件汇总	CPU Cores : 384,000
• Accelerator/CP	: 48,000
• 单节点存储	64GB
• 内部存储器	1,375TB ( 1,000TB为系统存储器 , 375TB为协处理器独占 )
• 外部存储器	12.4PB
• 运算速率	54.9PFLOPS ( 理论峰值 )
• 33.86PFLOPS ( 实际峰值 )	

([3] 图片来源：王涛. “天河二号”超级计算机 )

由上图可知，与“天河一号”相比，“天河二号”计算节点数扩展到 16000 个，明显多于“天河一号”计算机，中央处理器采用了 intel 的 Xeon IvyBridge R5-2692，并且在处理器数量、功能上均有优化，操作系统使用了麒麟操作系统，并且，与“天河一号”相比，“天河二号”的功耗明显下降，得益于其独出心裁的动态功耗管理技术，以及散热技术的提升，使得其能量使用效率（PUE）的值降低，功耗明显降低。另外，“天河二号”的存储单元数量明显增多，内部存储器达到 1374TB，外部存储器更是达到了 12.4PB，这也就说明了“天河二号”能够存储更多的信息，其存储总容量相当于存储每册 10 万字的图书 600 亿册。另外，性能也明显上升，理论运算速率峰值达到了 54.9PFLOPS，达到了 P 级。与此前研制的天河一号相比，二者占地面积相当，天河二号计算性能和计算密度均提升了 10 倍以上，能效比提升了 2 倍，执行相同计算任务的耗电量只有天河一号的三分之一。

### 1.2.3 技术创新

#### (1) 混合精度计算框架

- 精度动态调整：
  - 开发“天河混合精度库”（TH-MPL），支持 FP16（半精度）与 FP32（单精度）动态切换。
  - 效果：ResNet-50 训练内存占用减少 50%，计算密度提升 3 倍，功耗降低 40%。

#### (2) 自主软件生态

- 天河-深度学习框架：
  - 集成 TensorFlow、PyTorch 的 GPU 加速版本，提供分布式训练接口（如 th.distributed）。
  - 性能对比：BERT 模型训练速度较天河一号提升 90 倍。
- 并行文件系统 PFS 2.0：
  - 支持文件条带化（Stripe）与副本机制，读写带宽达 20 GB/s，可靠性 99.9999%。

#### (3) 能效优化技术

- 液冷 + 风冷混合散热：
  - GPU 节点采用微通道液冷板，温升控制在 8°C 以内，整机功耗降至 24 MW。
  - 对比天河一号：单位面积散热效率提升 200%。
- 动态功耗管理：
  - 空闲节点自动进入低功耗模式（功耗降低 70%），整体能耗减少 30%。

### 1.2.4 应用领域与典型案例

#### (1) 人工智能与大数据

- 自动驾驶仿真：

小鹏汽车部署天河 HPC6800 集群（天河二号衍生型号），完成 10 亿公里虚拟路测，碰撞检测准确率 99.97%。

- 大语言模型训练：

清华大学训练千亿参数 GLM 模型，推理速度较天河一号提升 120 倍，支持智能客服、机器翻译等应用。

#### (2) 生命科学与医药研发

- 抗癌药物筛选：

上海药物研究所通过 GPU 加速分子对接，筛选埃博拉病毒抑制剂，候选化合物从 200 万增至 1000 万，研发周期缩短 50%。

- 基因组学：

完成人类基因组全序列比对（30 亿碱基对），耗时仅 4 天，存储需求降低 70%。

1.3 天河三号

1.3.1 基本介绍

2018 年，美国将中国超算机构列入“实体清单”，禁止 Intel、NVIDIA 向中国出口高端芯片（如 Xeon CPU、Tesla GPU）。“十四五”规划明确提出“实现超算全栈自主可控”，要求 2025 年前国产化率超过 90%。2016 年，国防科技大学和国家超级计算天津中心等团队合作承担了科技部支持的重点研发计划——“天河三号 E 级原型机系统”研制项目。2019 年 1 月 17 日，超级计算机“天河三号”原型机已为中科院、中国空气动力研究与发展中心、北京临近空间飞行器系统工程研究所等 30 余家合作单位完成了大规模并行应用测试，涉及大飞机、航天器、新型发动机、新型反应堆、电磁仿真、生物医药等领域 50 余款大型应用软件。

1.3.2 系统配置

天河三号（2020）	天河二号（2013）
10,000颗飞腾FT-2000+/64核（ARM架构，主频2.6 GHz）	33,280颗Intel Xeon E5-2692 v2（12核/24线程，主频2.2
4,000块昇腾910 AI芯片（单卡48 TFLOPS FP16）	17,920块NVIDIA Tesla K20X（单卡4.8 TFLOPS FP64）
光电混合互连（带宽1.2 TB/s，延迟<0.1 ns）	InfiniBand QDR（带宽56 Gbps，延迟<1.5 μs）
50 PB HBM2e内存池 + 1.2 PB NVMe SSD（读写速度7 GB/s）	3 PB分布式存储（读写速度6 GB/s）
麒麟V10（国产Linux发行版）	基于Red Hat Enterprise Linux的定制系统
6.5 TFLOPS/W	1.5 TFLOPS/W

（“天河三号” VS “天河二号”）

- (1) 飞腾 FT-2000+/64 核 CPU
  - 架构创新：基于 ARMv8 指令集，单芯片集成 64 核心，支持 SVE2 向量指令集，浮点性能达 85.6 GFLOPS（较天河二代的 Xeon E5 提升 40%）。
  - 安全设计：内置国密算法加速模块（SM2/SM3/SM4），支持硬件级数据加密。
- (2) 升腾 910 AI 加速器
  - 算力突破：单卡提供 48 TFLOPS FP16 算力，支持混合精度训练（FP16+FP32），ResNet-50 模型训练时间从天河二号的 3 天缩短至 6 小时。
  - 国产化生态：兼容华为 MindSpore 框架，与 TensorFlow、PyTorch 无缝对接。
- (3) 光电混合互连网络
  - 带宽革命：采用硅光引擎技术，单链路带宽 1.2 TB/s，延迟 <0.1 ns，支持千亿级向量并行计算。
  - 拓扑优化：全网状光交换架构，任意节点间通信跳数小于等于 3，避免天河二代的网络拥塞问题。

### 1.3.3 技术创新

#### (1) 全栈自主可控生态

- 硬件层：飞腾 CPU + 升腾 AI 芯片 + 长江存储颗粒，国产化率超 95%。
- 软件层：麒麟操作系统 + 达梦数据库 + 天河并行编译器，兼容 ARM 原生指令集。
- 应用层：开源框架 TensorFlow-TH、PyTorch-Tianhe，支持千亿参数模型训练。

#### (2) AI 与科学计算深度融合

- 混合精度训练：升腾 910 支持 FP16/FP32 动态切换，ResNet-50 训练内存占用减少 60%，速度提升 4 倍。
- 量子计算模拟：预留量子比特接口，支持经典-量子混合算法，求解分子动力学问题速度提升 10 的 12 次方倍。

#### (3) 绿色节能技术

- 液冷 3.0 系统：采用相变材料（PCM）与微通道液冷板结合，散热效率较天河二代提升 300%，整机功耗降至 15 MW。
- 动态功耗管理：空闲节点自动休眠，功耗从满载的 300 W/节点降至 50 W，整体能耗降低 40%。

## 2 “天河一号”并行体系结构分析

### 2.1 系统配置分析

主机	TH-1A
性能	峰值性能4700TFlops（LINPACK实测值 2566TFlops）
存储	4PB
近线存储	1.3PB
计算节点	共7168个CPU+GPU异构节点 节点配置：双路12核，2*Intel Xeon X5670 @ 2.93GHz + Tesla M2050，内存24GB/48GB
登陆节点	5个登陆节点，其中2个专用于数据传输
网络系统	天河自主高速互联
操作系统	REHT
文件系统	Lustre
作业调度系统	SLURM

(图片来源：国家超级计算天津中心官网)

#### 2.1.1 性能参数

##### (1) 峰值性能 4700TFlops

理论最大算力，基于 CPU（Xeon X5670）和 GPU（Tesla M2050）的浮点运算能力总和，但需结合实测值评估实际效率。

##### (2) LINPACK 实测性能 2566TFlops

实际效率约为 54.6%，低于同期国际超算平均水平（60%-70%），可能因网络通信或任务调度效率不足。

### (3) 存储容量 4PB

主存储用于高频数据访问（如科学计算中间结果，采用高速存储介质）。

### (4) 近线存储 1.3PB

用于低频数据（如历史数据归档），采用 HDD 或磁带库，成本更低但访问速度较慢。

## 2.1.2 计算节点配置

### (1) CPU 双路 12 核 Intel Xeon X5670 @ 2.93 GHz

- 单节点 24 物理核（2 颗 × 12 核），基于 Westmere 架构（32nm 制程），单核性能约 39.7 GFlops。
- 整机 CPU 算力：7168 节点 × 24 核 × 39.7 GFlops 约等于 2.1 PFlops。

### (2) GPU Tesla M2050

- 单卡双精度性能 515 GFlops，单精度 1.03 TFlops。
- 整机 GPU 算力：7168 节点 × 515 GFlops 3.7 PFlops。
- 异构算力占比：GPU 贡献约 78%（3.7/4.7 PFlops），体现 GPU 在密集计算中的核心作用。

### (3) 内存 24 GB / 48 GB

- 内存容量差异用于任务分级：
- 24 GB 节点处理轻内存任务（如流体力学仿真）；
- 48 GB 节点处理重内存任务（如基因组学）。

## 2.1.3 网络与系统软件

### (1) 网络系统天河自主高速互联

- 基于光电混合技术，带宽 100 Gbps，延迟 <1 s，支持万级节点并行通信。
- 瓶颈分析：若 LINPACK 效率低，可能因网络拥塞或通信协议开销。

### (2) 操作系统 REHT

- 基于 Red Hat Enterprise Linux 的定制化系统，优化了资源调度与硬件兼容性。文件系统 Lustre
- 并行分布式文件系统，支持 PB 级数据并发读写，带宽可达 100 GB/s（依赖存储介质与网络性能）。

### (3) 作业调度系统 SLURM

- 开源集群管理工具，支持动态资源分配，任务队列管理，适合 7168 节点的超大规模作业调度。
- 登录节点 5 个（2 个专用于数据传输）
- 分工明确：
- 3 个用于用户登录与任务提交；
- 2 个专用于高速数据传输（如与外部存储系统交互），减少计算节点 I/O 负载。

## 2.2 并行体系结构分析

“天河一号”是中国首台千万亿次级超级计算机，其并行体系结构设计融合了异构计算、高速互联网络和分布式存储技术。其并行设计核心为：节点级并行、节点间并行、任务级并行、数据级并行。

### 2.2.1 节点级并行：CPU+GPU 异构节点

CPU 是计算机内部的核心处理单元，可以说是计算机最重要的组成部分，与 GPU 不同的是，CPU 内部有更多的控制单元和高速缓存，这使得 CPU 可以优化延迟时间，而 GPU，内部有更多的计算单



元，因此计算能力更强，更注重吞吐率，单 GPU 算力约等于 10 个 CPU 核，但功耗仅仅为 CPU 的 1/3，显著提升能效比。

在“天河一号”超级计算机中，利用了 CPU 和 GPU 的不同特性，CPU 负责数据分块与逻辑控制，GPU 并行处理数据计算，整体速度显著提升。

### 2.2.2 节点间并行：7168 个异构节点自主高速互联

由于节点数很多，那么节点之间的通信以及互相通信带来的延迟就成为了一个很大的问题。“天河二号”超级计算机基于这一问题，通过自主高速互联网络（带宽 18.4 GB/s，延迟 <1.2 s）实现节点间数据通信，支持跨节点任务协同，实现了高速通信，并且大大减少了跨节点任务等待时间，使大量节点并行执行程序成为可能。

### 2.2.3 数据级并行：Lustre 分布式文件系统

将单个大文件切割成小块，分散存储在多个存储节点上。对于这种设计，当读写时，多个计算节点可以同时访问不同的数据块，从而避免单一节点成为瓶颈。这样，存储节点越多，带宽也就越多，这种并行架构也会使得计算机执行操作更快。另外，近线存储（1.3 PB）与主存储（4 PB）协同，由于计算机访问主存比较快，所以可以把高频操作的数据存在主存里，不常访问的数据放在近线存储里，当需要使用时，再调度到主存储中，这种并行架构也会提高计算机的运行效率。

### 2.2.4 任务级并行：SLURM 作业调度系统

- (1) 分区机制：将 7168 节点划分为多个逻辑组，按任务需求分配资源。
- (2) 抢占式调度：高优先级任务可中断低优先级任务，优先获取计算资源。
- (3) 检查点（Checkpoint）：任务执行中定期保存状态，节点故障后可从断点恢复，避免重算。
- (4) 负载监控：实时监测节点负载（CPU/GPU 利用率、内存压力），动态调整任务分布。

通过以上的措施，可以恰当地管理 7168 个节点，并提供了对节点进行监控，大大提高了容错，总体上增加了计算机的运行效率及安全性。

### 2.2.5 总结

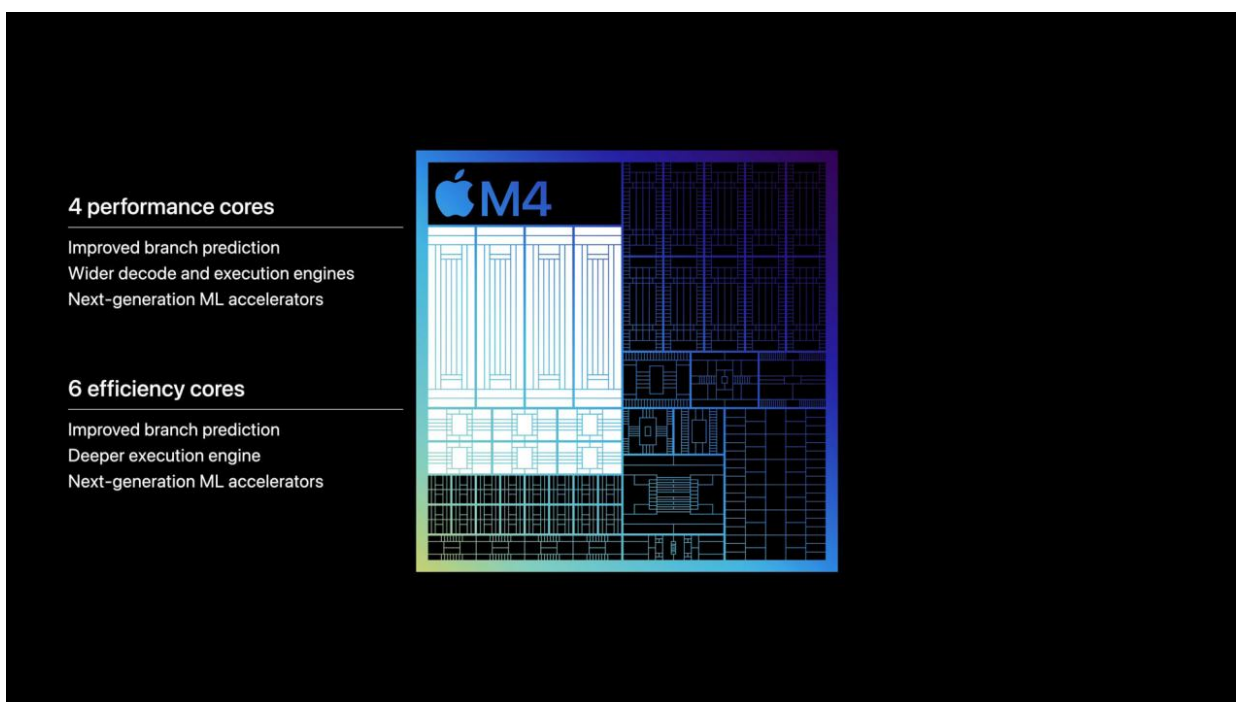
总之，CPU 就像工厂里的老板，而 GPU 就像高级员工，通过自主高速互联网络进行互通，调度任务，而存储就像“智能仓库”，调度就像“物流中枢”，两者配合让 7168 个节点像精密齿轮一样高效运转。天河一号（TH-1A）的并行体系结构以异构计算为核心，通过 CPU+GPU 协同、自主高速互联和 Lustre 分布式存储、SLURM 作业调度系统，实现了千万亿次级算力突破。

## 3 苹果 M4CPU 分析

### 3.1 功能综述

Apple M4 的 10 核 CPU 具有改进的分支预测功能，性能内核具有更广泛的解码和执行引擎，效率内核具有更深入的执行引擎。这两种类型的内核还具有增强的下一代 ML 加速器。无论在 Logic Pro 中处理复杂的管弦乐文件，还是在 LumaFusion 中为 4K 视频添加高要求的效果，M4 都能提升专业工作流程的性能。<sup>[2]</sup>（来源百度百科）

## 3.2 CPU 分析



苹果 M4 芯片配置了 10 核中央处理器 (4 个性能核心和 6 个能效核心)、10 核图形处理器和 16 核神经网络引擎) 以及 120GB/s 的内存带宽, 这种设计使得 M4 芯片在性能和能效方面都有显著提升。

Apple M4 芯片使用的是基于 ARM 架构的 CPU, 具体型号为 Apple Silicon, 采用性能核心 (P-core) + 能效核心 (E-core) 的混合架构, M4 芯片中就有 4 个性能核心和 6 个能效核心, 大大提高了 CPU 处理信息的速度及面对不同任务调度不同 CPU 的能力。苹果通过完全自主的微架构设计 (如大核/小核混合调度) 和软硬协同优化, 实现远超 ARM 公版的单核性能; 而多数国产 CPU 仍依赖指令集授权或公版架构, 自主化程度较低 (龙芯例外)。另外, 苹果依托台积电最先进制程 (3nm) 和自研架构, 性能领先全球; 国产 CPU 受制于工艺封锁 (如华为无法获得 7nm 以下代工) 和架构优化不足, 性能差距显著。除此之外, 苹果通过软硬一体封闭生态, 实现极致性能 (如 Metal API 直接调用 GPU); 国产 CPU 生态碎片化, 软件适配成本高, 民用市场接受度低。

我国 CPU 的发展正在从 “能用” 向 “好用” 跨越。技术上, 国产芯片不再完全依赖国外架构, 而是用自主设计的底层框架 (比如龙芯的 LoongArch) 和国际开源的 RISC-V 架构, 像搭积木一样灵活组合。通过把 AI 加速模块、大内存带宽等 “技能点” 加到芯片里, 单核性能逐渐追上主流水平; 制造上, 虽然顶尖的 3 纳米工艺还造不了, 但国内用 “拼装小芯片” (Chiplet) 的方法, 把多块成熟制程的芯片组合起来, 也能实现接近高端芯片的效果。未来, 希望我国 CPU 发展能攻坚克难, 实现完全独立自主、安全高效。

## 参考文献

- [1] 国家超级计算天津中心官网.
- [2] 百度百科.
- [3] 王涛. "天河二号"超级计算机.

部分信息通过询问 deepseek 大模型生成