



计算产业发展简史及趋势



前言

- 以“信息技术”为特征的第三次工业革命，将世界带入了数字化时代；以“智能技术”为特征的第四次工业革命，目前正在席卷全球，将会把人类社会带入智能化时代。
- 本章主要介绍不同种类的处理器芯片以及异构计算的相关知识。





目标

- 学完本课程后，您将能够：
 - 描述计算产业发展简史及趋势
 - 了解处理器芯片基本分类
 - 了解异构计算的基本知识



泰克教育
TECH EDUCATION



目录

1. 计算产业发展简史及趋势

- 什么是计算产业
 - 发展简史
 - 现状及挑战

2. 处理器芯片及异构计算



泰克教育
TECH EDUCATION



计算产业定义

- 计算是使用计算机的任何活动，它包括开发硬件和软件，并使用计算机来管理、处理和通信各种用途的信息。

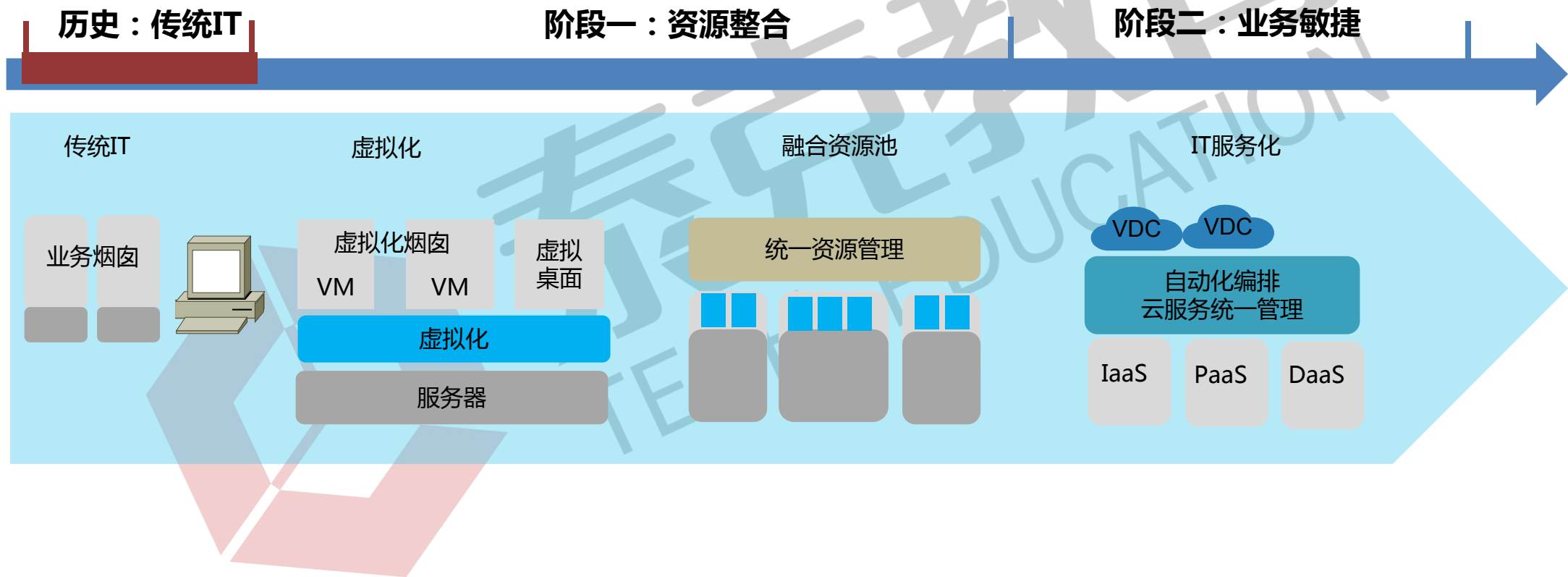
——维基百科

- 计算产业是IT技术的基础，是每一次产业变革的驱动力，从云计算、大数据、人工智能，到区块链、边缘计算、物联网都离不开强大的计算能力的支持。





云计算正在改变IT基础架构





大数据实现数据增值



1960s 大型机



1980s 个人电脑



1990s 互联网



2000s 移动互联网



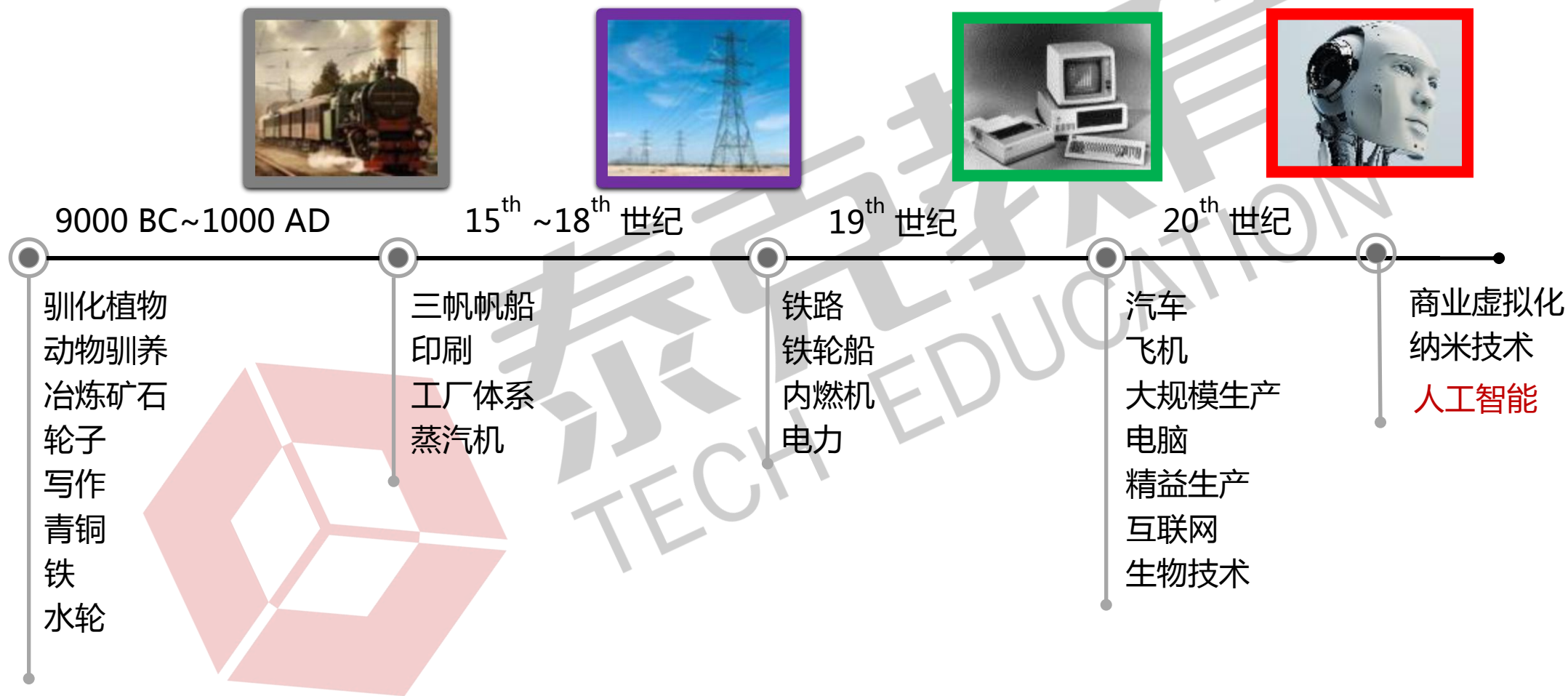
2010s 大数据

一个大规模生产、分享和应用数据的时代正在开启.....

——肯尼思·库克耶《大数据时代》作者



人工智能是一种新的通用技术





目录

1. 计算产业发展简史及趋势

- 什么是计算产业
- 发展简史
- 现状及挑战

2. 处理器芯片及异构计算

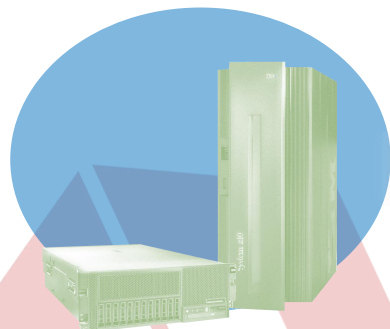


泰克教育
TECH EDUCATION



计算产业在进化：从计算1.0到计算3.0

计算1.0



专用计算

大型机、小型机时代

计算2.0



通用计算

数据中心，X86架构

计算3.0



智能计算

全栈全场景



计算1.0 - 专用计算



System/z



HPE SDX



SPARC



计算2.0 - 通用架构服务器

X86
服务器

关键业务服务器



KunLun

刀片服务器

E9000



半宽 2S



CH121L

半宽 2S



CH121

全宽 2S



CH225

全宽 2S



CH221

全宽 4S



CH242

机架服务器

1U 2S



1288H

2U 2S



2288/2288H

4U 2S



5288

2U 4S



2488H

4U 4S



5885H

高密服务器

2U 4节点



X6000



XH321



XH321L

4U 4节点



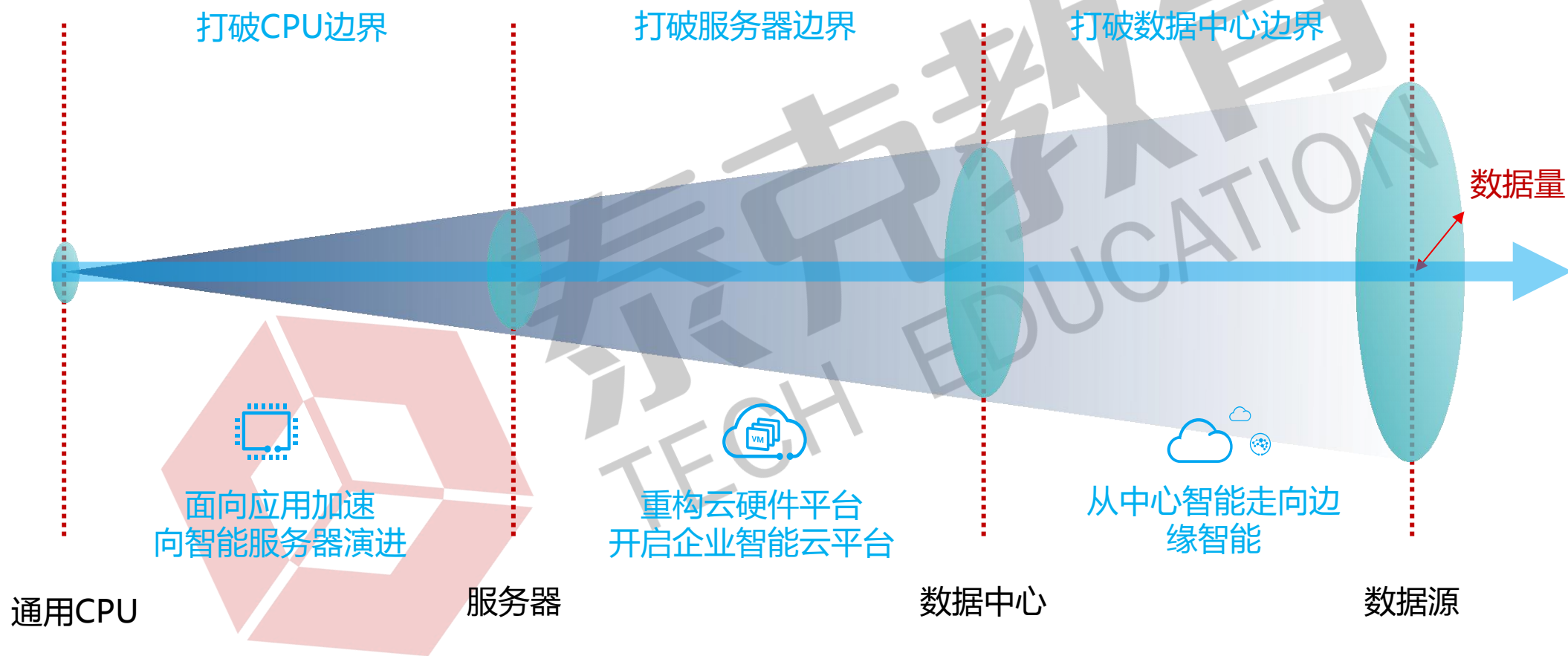
X6800



XH628



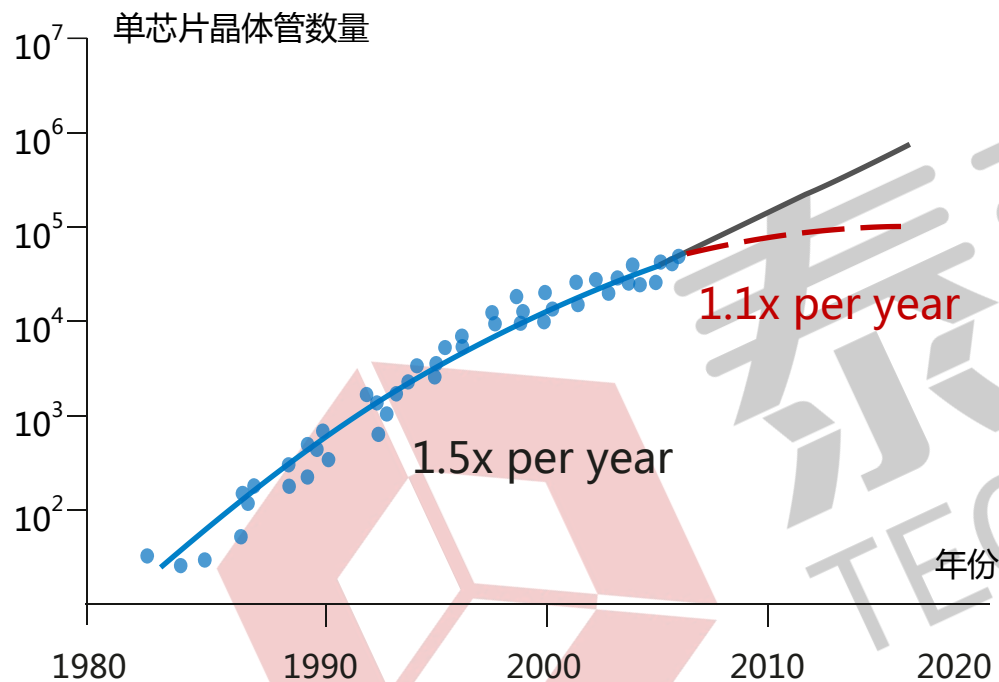
计算3.0 - 智能计算



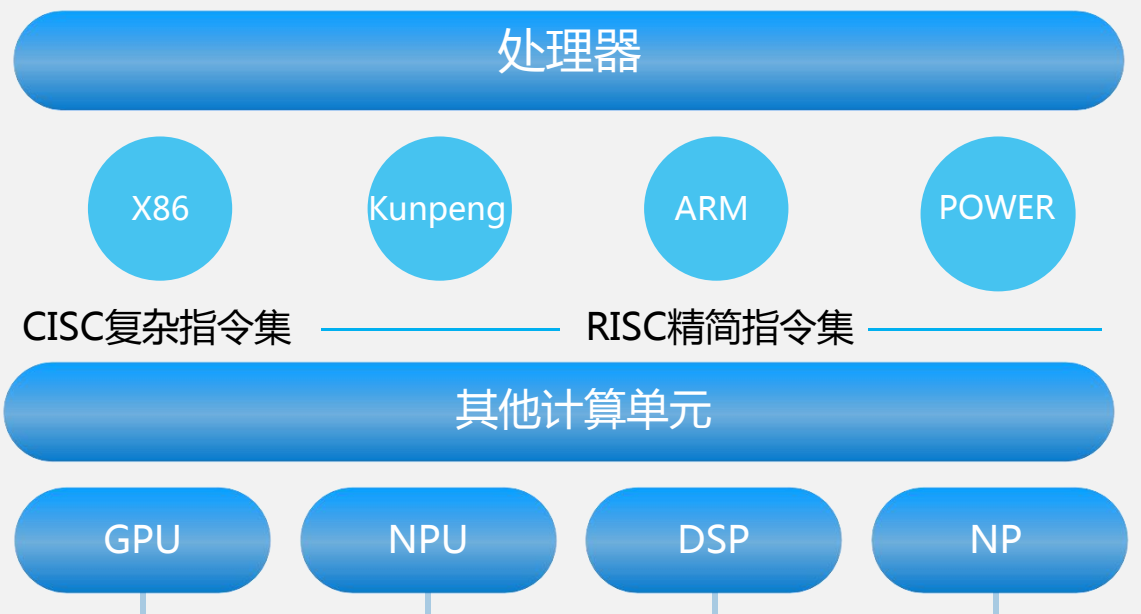


处理器采用多种架构

重构摩尔定律



多种计算架构的组合是最优的解决路径





目录

1. 计算产业发展简史及趋势

- 什么是计算产业
- 发展简史
- 现状及挑战

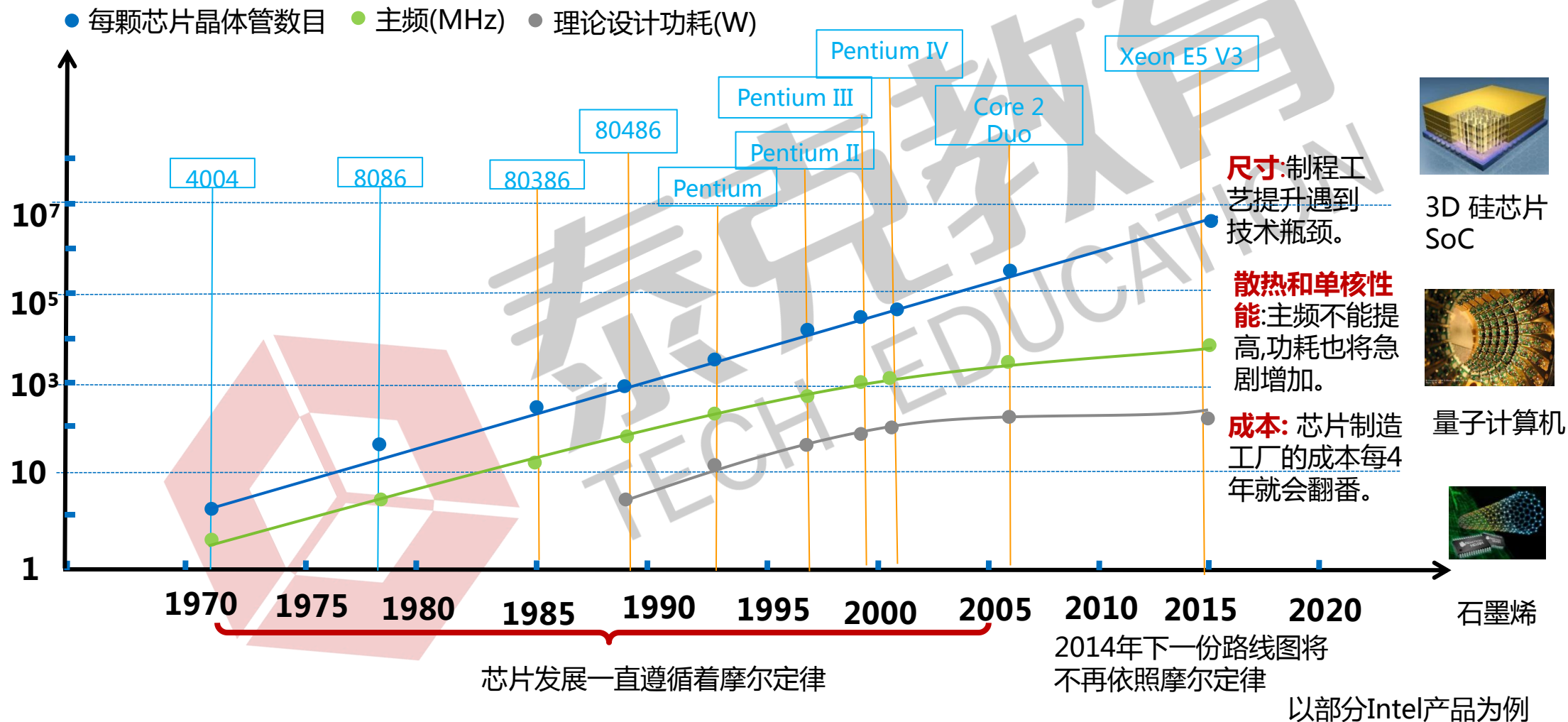
2. 处理器芯片及异构计算



泰克教育
TECH EDUCATION



工艺成本遇瓶颈





人工智能面临巨大挑战

1. 训练需要数日、数月



需要缩短训练时间

2. 算力稀缺且昂贵



需要足够廉价的计算资源

3. AI主要部署在云、少量在边缘



需要将AI部署到所有场景

4. 主要算法诞生于1980年代



需要更高效安全的算法

5. 没有“人工”就没有“智能”



需要自动化 / 半自动化的数据标注能力

6. 模型性能“考试”优秀



需要符合实际应用的模型

7. 非实时更新



需要模型实时更新

8. 与其它技术连接不充分



需要协同云、物联网、边缘计算、区块链、大数据、数据库等技术

9. 一项需要高级技能的、专家的工作



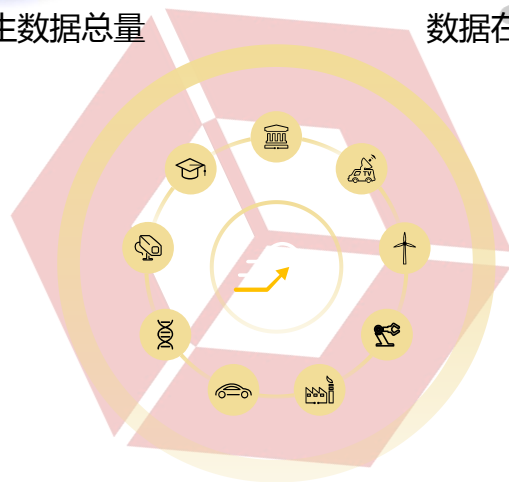
需要一站式平台的支持

.....



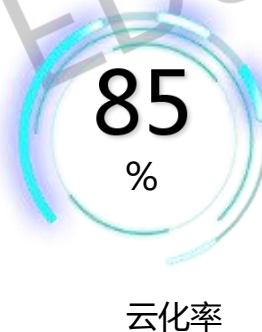
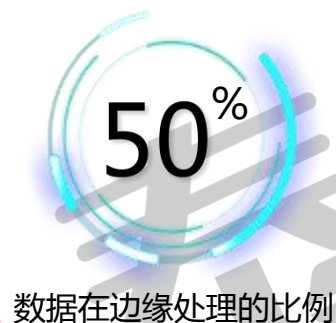
计算将走向边缘化与智能化

2025年数据规模预测



来源: IDC-Data Age 2025

2025年数字化与智能化趋势预测



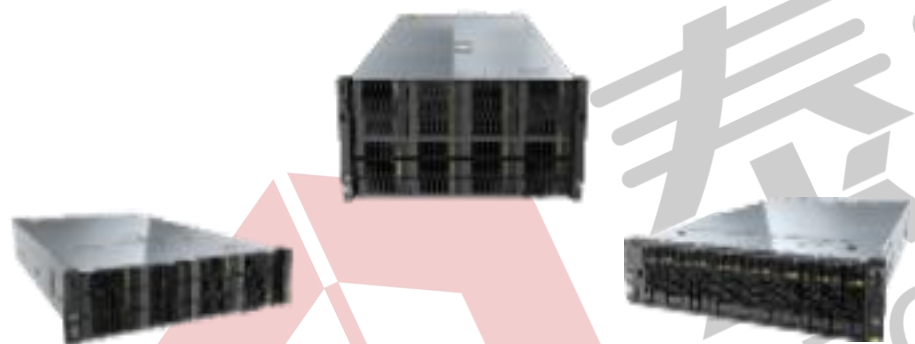
来源: 全球产业展望GIV 2025(Global Industry Vision 2025)



华为智能计算战略

TaiShan

Atlas



鲲鹏计算平台（通用计算）



昇腾计算平台（AI计算）

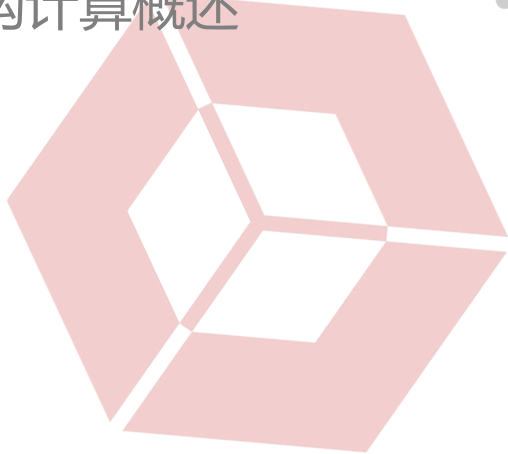


目录

1. 计算产业发展简史及趋势

2. 处理器芯片及异构计算

- 芯片和处理器定义
- 中央处理器概述
- 异构计算概述

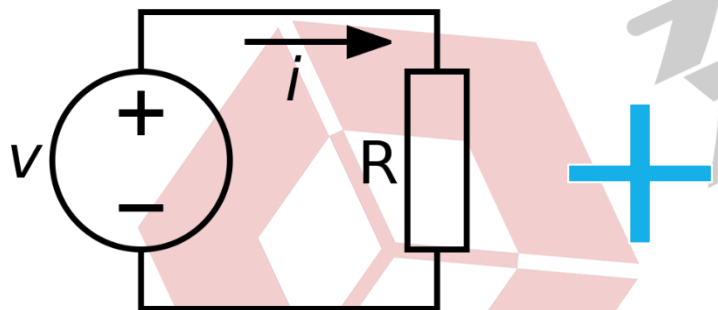


泰克教育
TECH EDUCATION

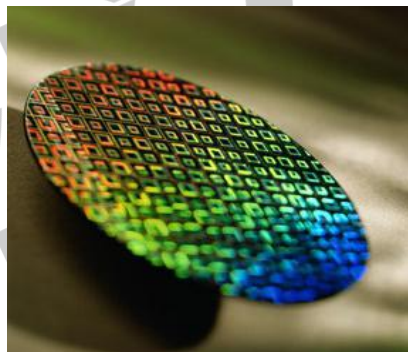


什么是芯片

芯片（chip）是一种把电路（主要包括半导体设备，也包括被动组件等）小型化的方式，并制造在半导体晶圆表面上。



电路



晶圆

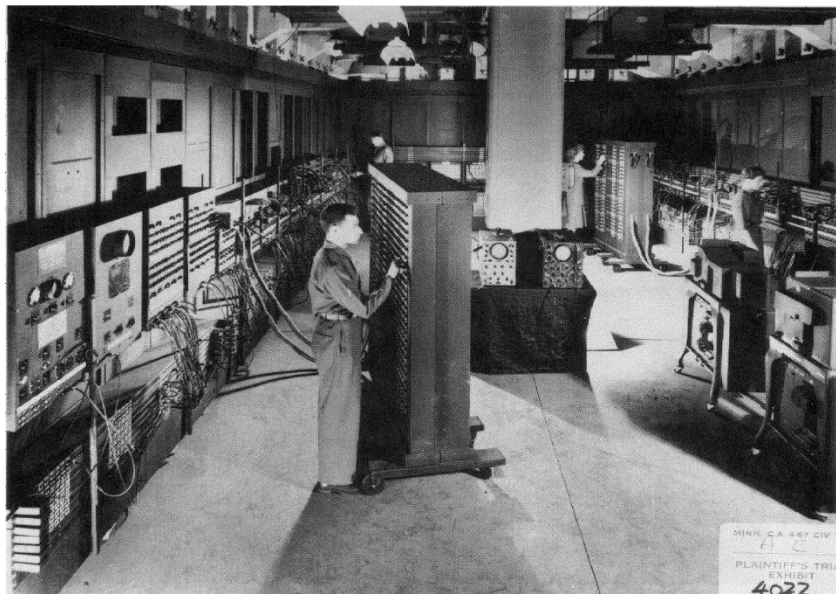


芯片



芯片是信息社会的基石

世界第一台通用电子计算机—ENIAC



1946年诞生，由18,800多个电子管组成，重量30多吨，占地面积170多平方米。

得益于芯片的发展



小型化，高性能



内存



手机



计算机



数码相机



华为服务器芯片产品概览

突破摩尔定律

华为鲲鹏处理器 智能网卡芯片



昇腾AI芯片



智能SSD控制器芯片

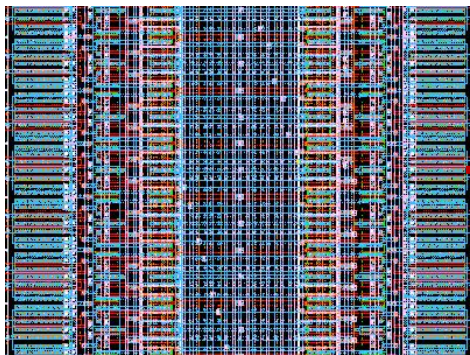


智能管理芯片



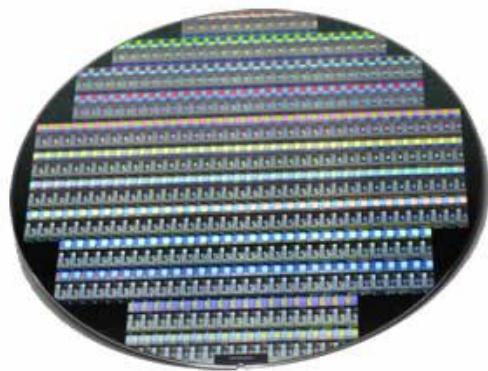


芯片行业产业链



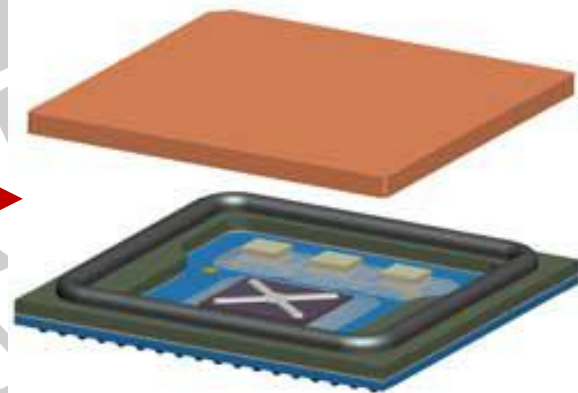
芯片设计

集成电路设计涉及对电子器件和器件间互连线模型的创建。



晶圆加工

晶圆加工是一系列化学处理步骤，使得电子电路逐渐形成在使用纯半导体材料制作的芯片上。



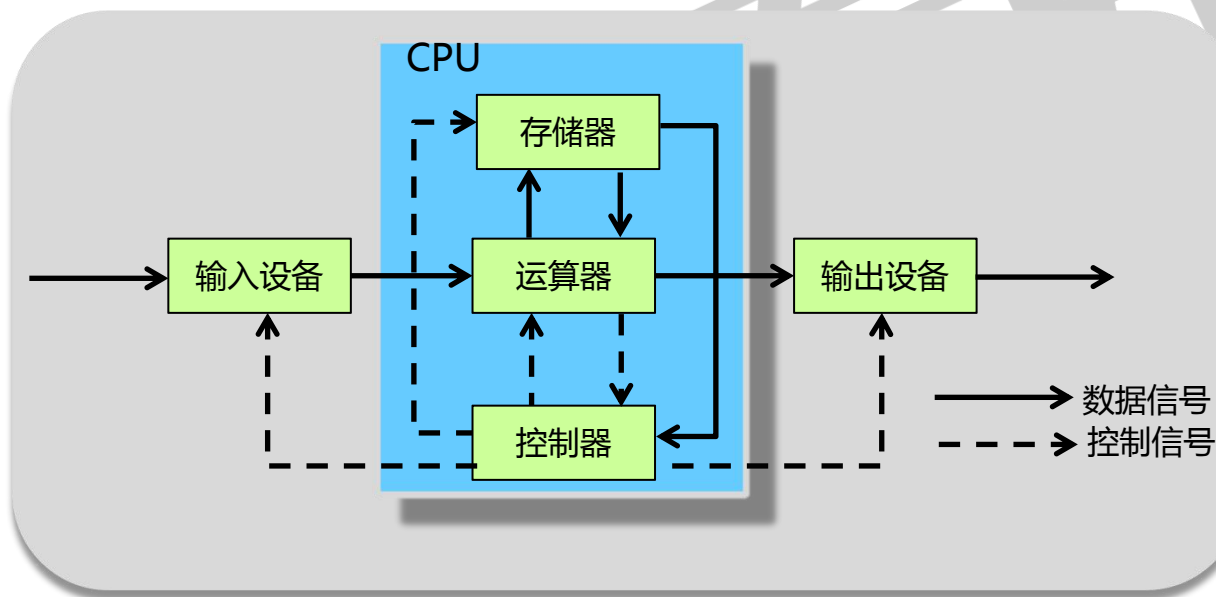
芯片封装测试

封装是将器件的核心晶粒封装在一个支撑物之内的过程，这个封装可以防止物理损坏以及化学腐蚀，并提供对外连接的引脚，之后将进行集成电路性能测试。



什么是中央处理器

- 中央处理器（Central Processing Unit，CPU）是计算机的主要设备之一，功能主要是解释计算机指令以及处理计算机软件中的数据。
- 中央处理器主要包括控制器，运算器，高速缓冲存储器（Cache）及实现它们之间联系的数据、控制及状态的总线。
- 它与主存（Memory）和输入/输出设备合称为电子计算机三大核心部件。



CPU功能

- 解释计算机指令
- 处理计算机软件中的数据

CPU组成

- 控制单元
- 运算单元
- 存储单元

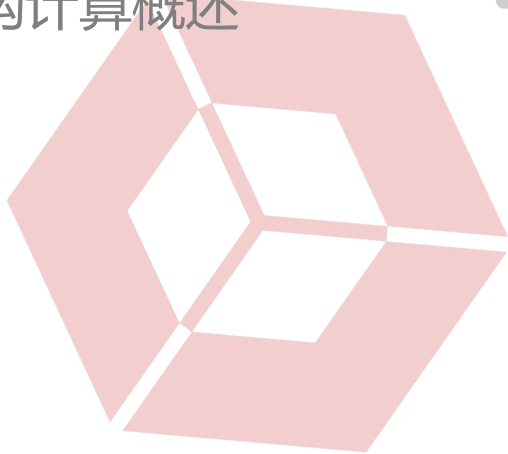


目录

1. 计算产业发展简史及趋势

2. 处理器芯片及异构计算

- 芯片和处理器定义
- 中央处理器概述
- 异构计算概述



泰克教育
TECH EDUCATION

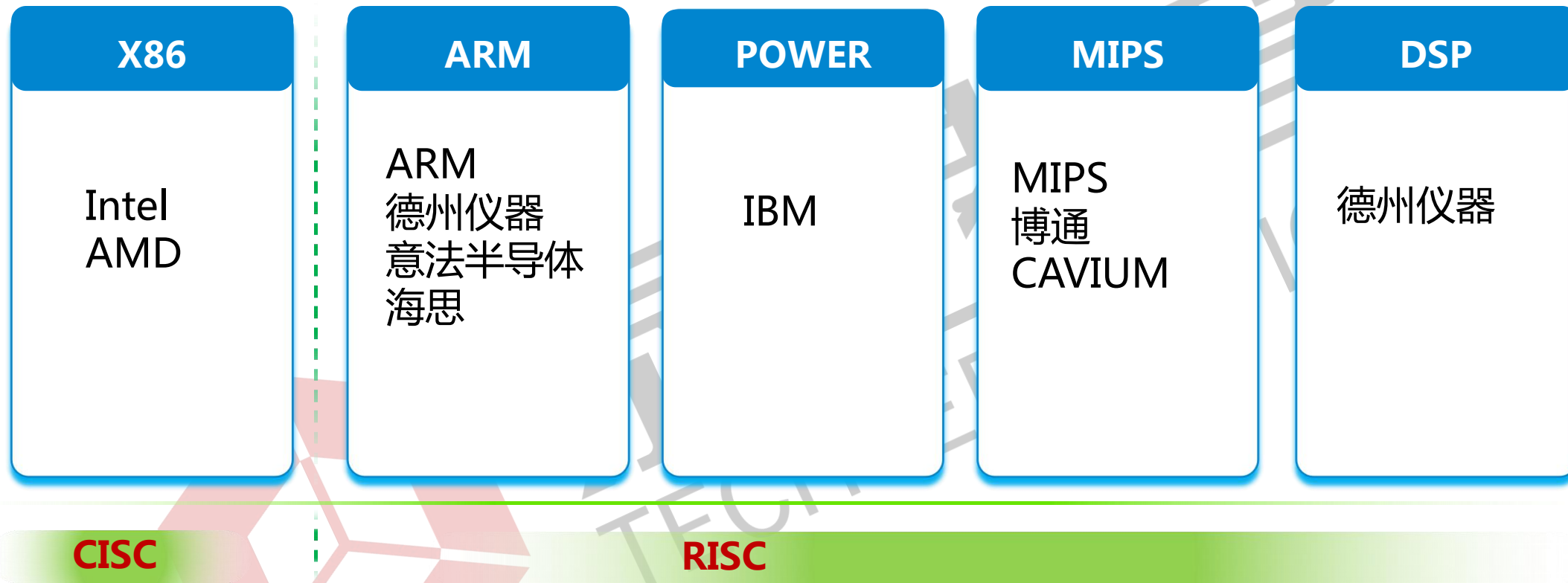


处理器分类

架构类型	架构名称	推出公司	推出时间	主要授权商
CISC	X86	Intel, AMD	1978	海光，兆芯
RISC	ARM	ARM	1985	苹果，三星，英伟达，高通，海思，TI等
	MIPS	MIPS	1981	龙芯，炬力等
	POWER	IBM	1990	IBM



什么是RISC 和 CISC



- CISC(Complex Instruction Set Computer), 复杂指令集。早期的CPU全部是CISC架构, 它的设计目的是要用最少的机器语言指令来完成所需的计算任务。这种架构会增加CPU结构的复杂性和对CPU工艺的要求, 但对于编译器的开发十分有利。
- RISC(Reduced Instruction Set Computer), 精简指令集。RISC架构要求软件来指定各个操作步骤。这种架构可以降低CPU的复杂性以及允许在同样的工艺水平下生产出功能更强大的CPU, 但对于编译器的设计有更高的要求。



RISC vs CISC

对比项	CISC	RISC
指令系统	复杂	精简
存储器操作	多，复杂	少，简单
程序设计	编程效率高	需要大内存空间，不易设计
CPU芯片电路	功能强、面积大、功耗大	面积小，功耗低
应用范围	通用机	专用机

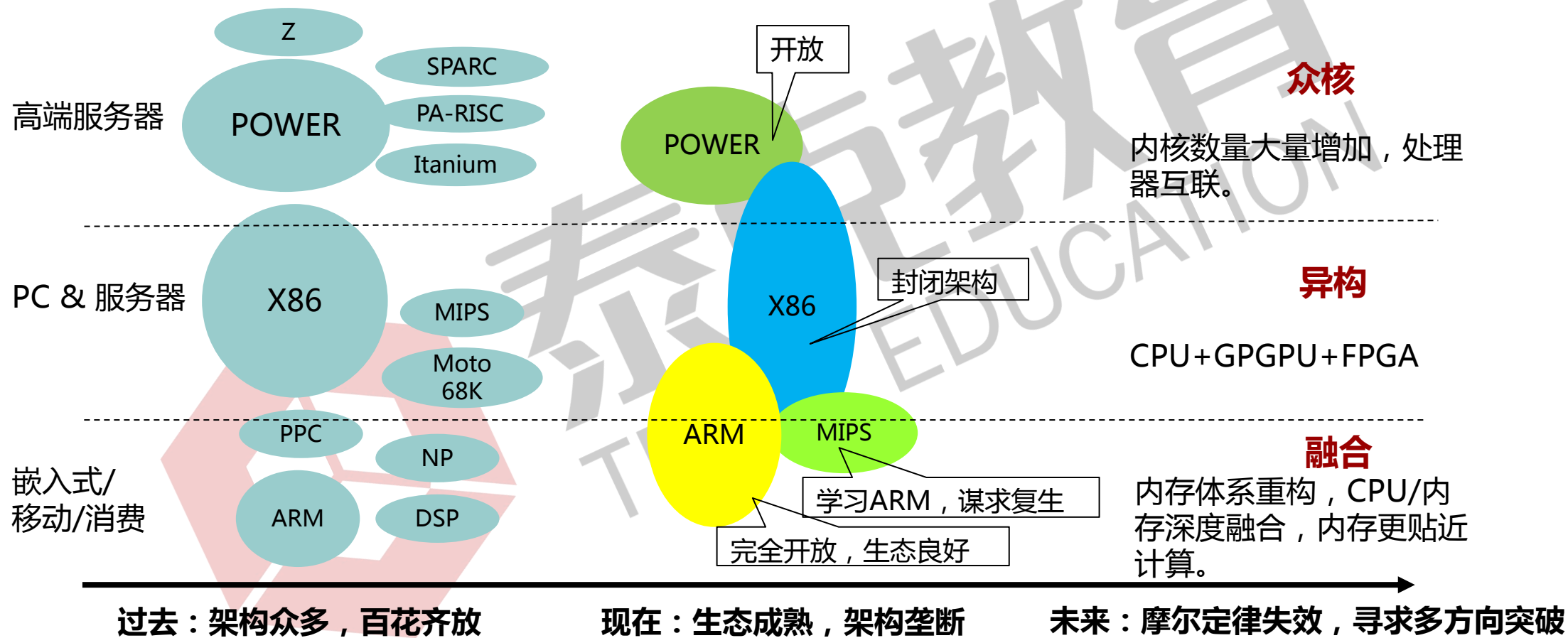


主流计算架构比较

	X86	ARM	POWER
指令集	• CISC, 复杂指令集	• RISC, 精简指令集	• RISC, 精简指令集
架构	• 重核架构, 高性能高功耗	• 多核架构, 均衡的性能功耗比	• 重核架构, 高性能内核
生态	• 生态非常成熟, 通用性强	• 生态正在快速发展与完备	• 聚焦大小型机和HPC
开放性	• 封闭架构, 英特尔及AMD主导	• 开放平台, IP授权的商业模式	• 开放平台 (2019.8), IBM主导

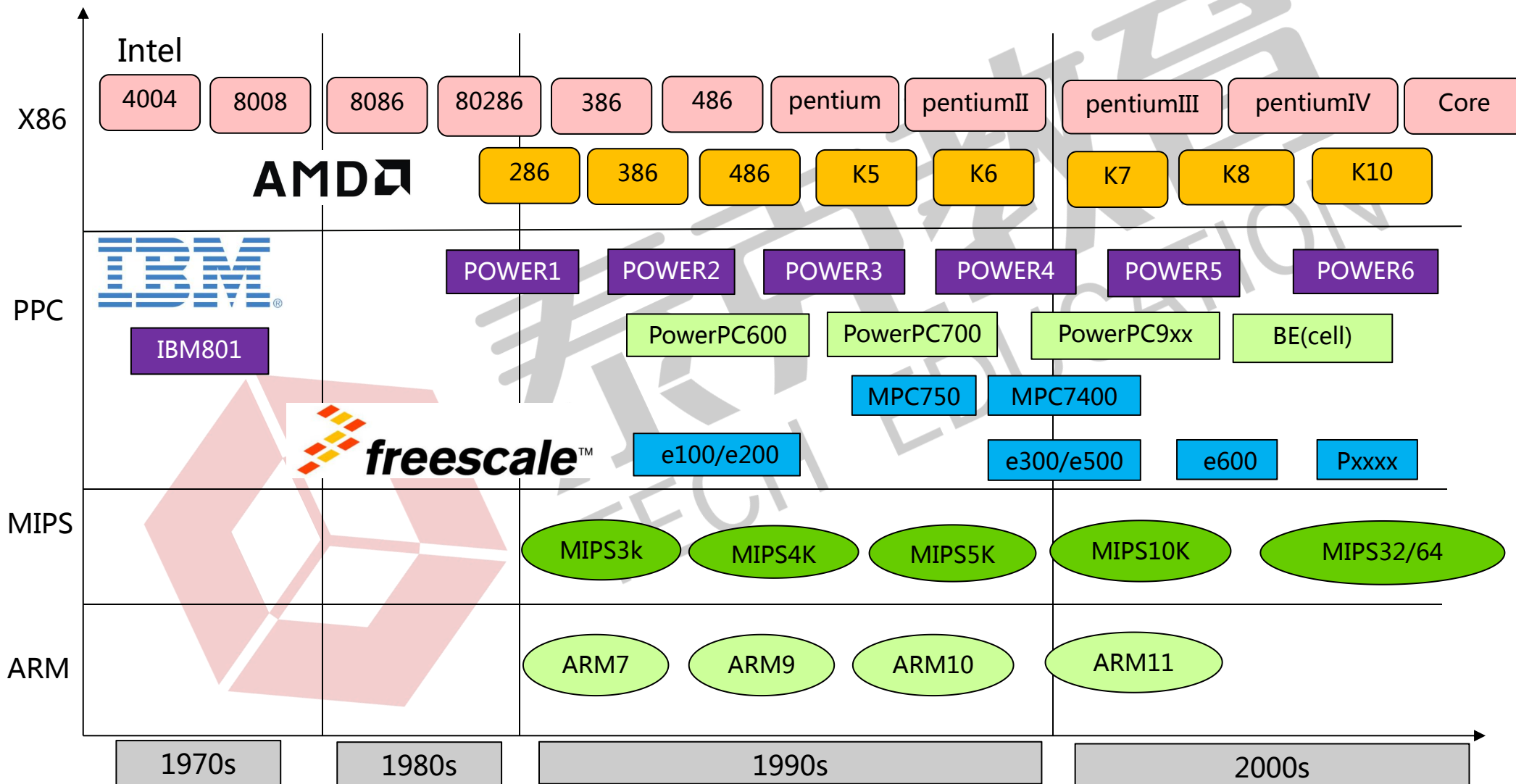


处理器发展趋势





主流CPU发展路径





X86处理器发展简史

X86处理器发展经历50年的历史：

- 字长从4位到64位；
- 时钟频率逐步提升；
- 晶体管集成规模逐渐增大；
- 性能逐渐提高的过程。





什么是ARM

ARM架构，高级精简指令集机器（Advanced RISC Machine，更早称作Acorn精简指令集机器，Acorn RISC Machine），是一个精简指令集（RISC）处理器架构家族，其广泛地使用在许多嵌入式系统设计。



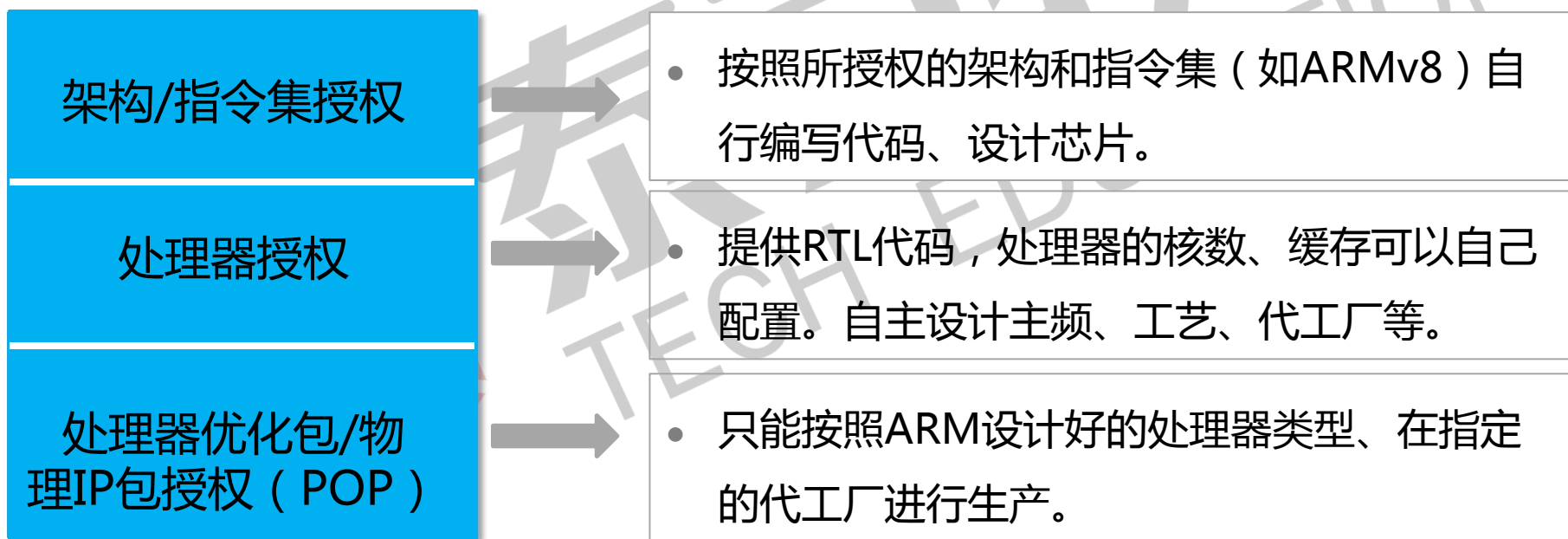
ARM Ltd. headquarters





ARM公司授权体系

- ARM目前在全球拥有大约1000个授权合作商、320家伙伴，但是购买架构授权的厂家不超过20家，中国有华为、飞腾获得了架构授权。





ARM vs X86



扩展性

重核，多核，多线程，高主频

轻核、众核

指令集

CISC, 通用指令集

RISC, 根据负载优化

供应商

只有两家CPU供应商(Intel , AMD)
Intel处于垄断地位

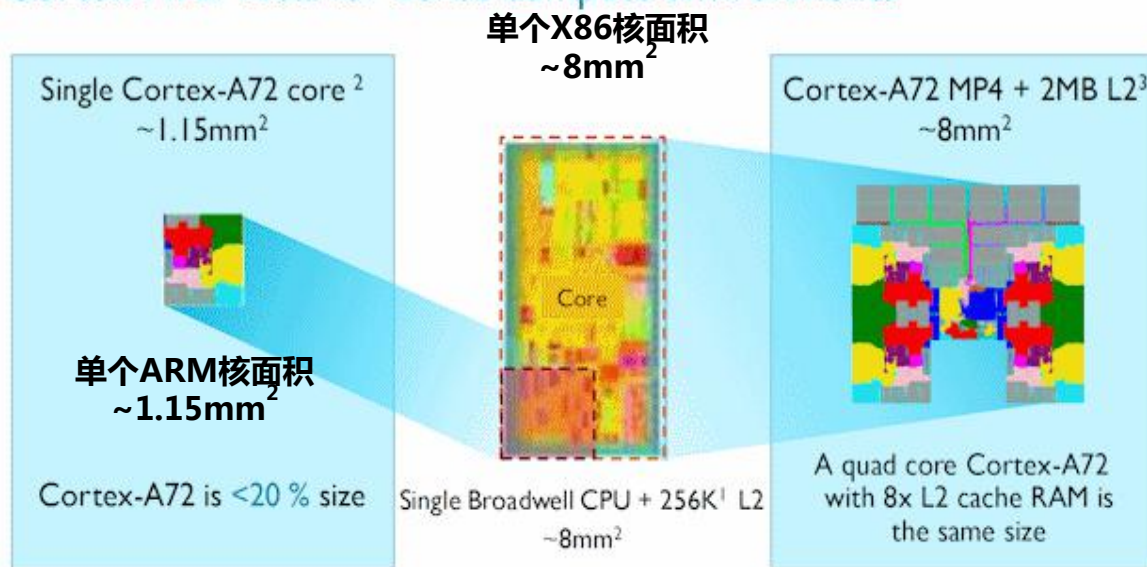
开放的授权策略
众多供应商



ARM提供更多计算核心

- 工艺、主频遇到瓶颈后，开始通过增加核数的方式来提升性能；
- 芯片的物理尺寸有限制，不能无限制的增加；
- ARM的众核横向扩展空间优势明显。

Cortex-A72: Ideal for dense compute environments





ARM服务器级别处理器一览



Hi1612

32C, 2.1GHz
16nm

Hi1616

32C, 2.4GHz
16nm

Hi1620

48C, 3.0GHz
7nm

CAVIUM

Thunder-X

48C, 2.5GHz
28nm

Thunder-X2

32-54C, 3.0GHz
14nm

高通

2400

48Cores, 2.2-2.6GHz
14nm

AMPERE

X-Gene3

32C, 3.3GHz
16nm

飞腾

FT1500

16 Cores, 1.6GHz
28nm

FT2000+

64Cores, 2.3GHz
16nm

横轴代表性能



华为鲲鹏处理器



1991
第一颗传输网
络的ASIC芯片

2005
第一颗基于ARM
的无线基站芯片

2009
第一颗基于ARM的移
动端处理器



K3

2014
Hi1612
第一颗基于ARM的
64位处理器



Hi1612

2016
业界第一颗支持
多路ARM 处理器



华为鲲鹏916

2019
业界第一颗7nm
数据中心处理器



华为鲲鹏920



华为鲲鹏930



华为鲲鹏950

2023

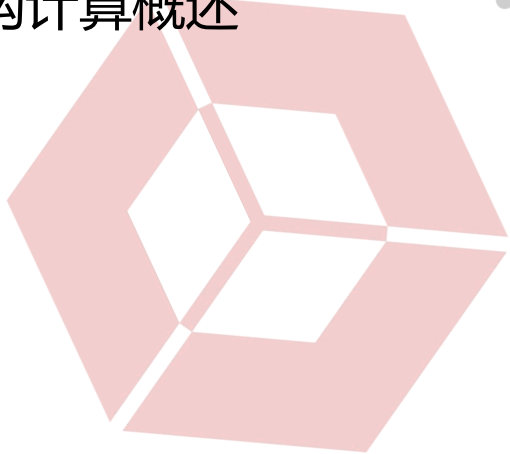


目录

1. 计算产业发展简史及趋势

2. 处理器芯片及异构计算

- 什么是芯片
- 中央处理器概述
- 异构计算概述



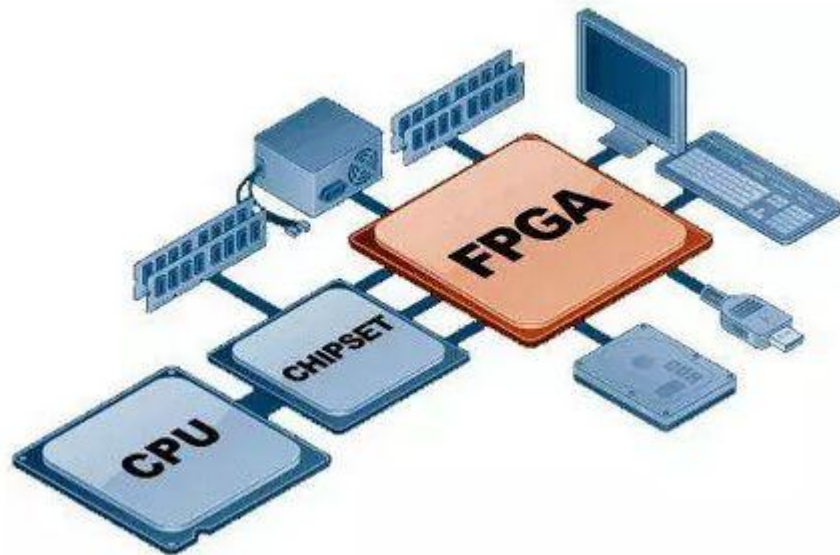
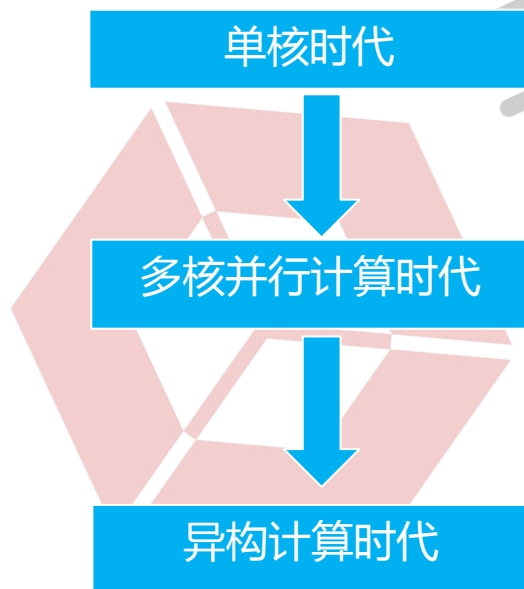
泰克教育
TECH EDUCATION



什么是异构计算

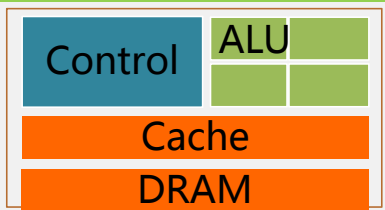
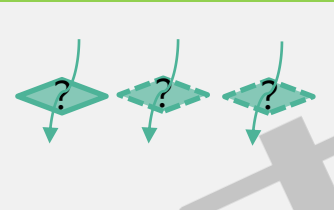
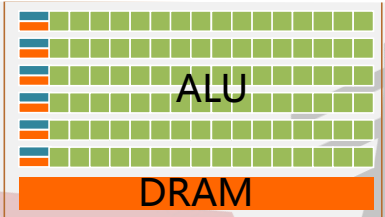
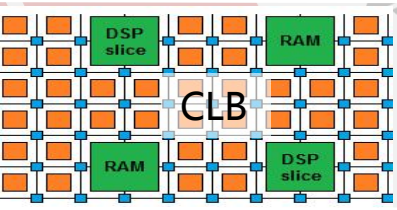
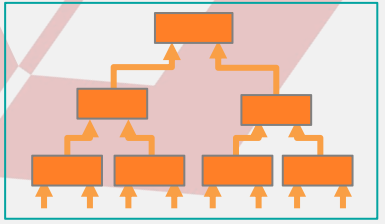
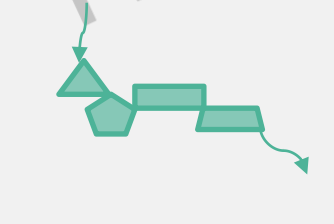
异构计算（Heterogeneous Computing），又称为异质运算，主要是指使用不同类型指令集和体系架构的计算单元组成系统的计算方式。常见的计算单元类别包括CPU、GPU、DSP、ASIC、FPGA等。

——维基百科





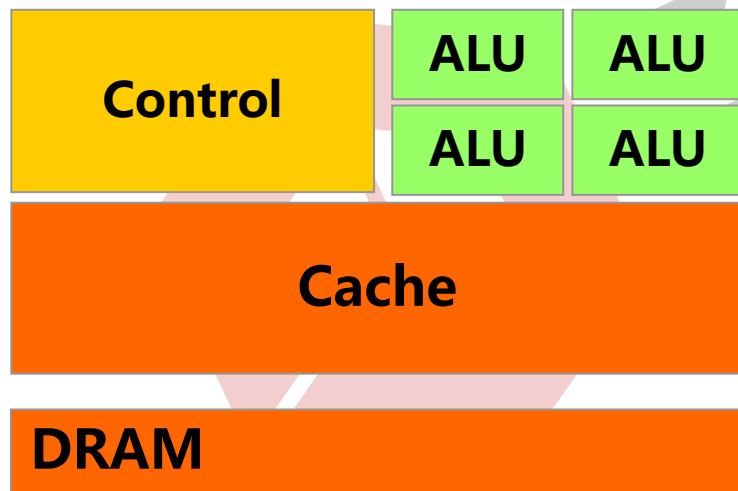
计算单元特性对比

	硬件抽象	计算模式	特性	Perf/W
灵活性 ↑ CPU			通用编程，上市时间短。	1x
GPU			适用大批量数据并行计算，高性能，高能耗。	10x
FPGA			适用于不规则并行计算，实时性好，能效比高。	20x
效能 ↓ ASIC			专用电路，高性能，低能耗。	> 20x

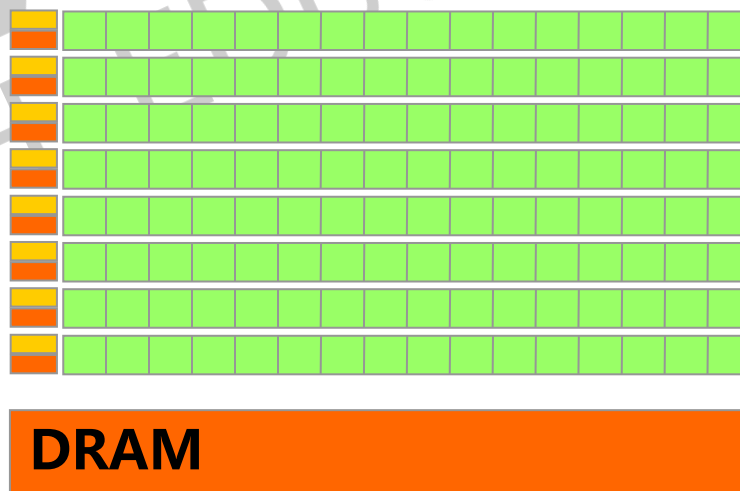


什么是GPU

- GPU (Graphics Processing Unit) 是图形处理器又称显示核心、视觉处理器、显示芯片，是一种专门在个人电脑、工作站、游戏机和一些移动设备上运行图像运算工作的微处理器。
- GPGPU (General-Purpose computing on Graphics Processing Units) 即图形处理单元上的通用计算，利用处理图形任务的图形处理器来计算原本由中央处理器处理的通用计算任务。
- GPU与CPU的硬件逻辑架构对比



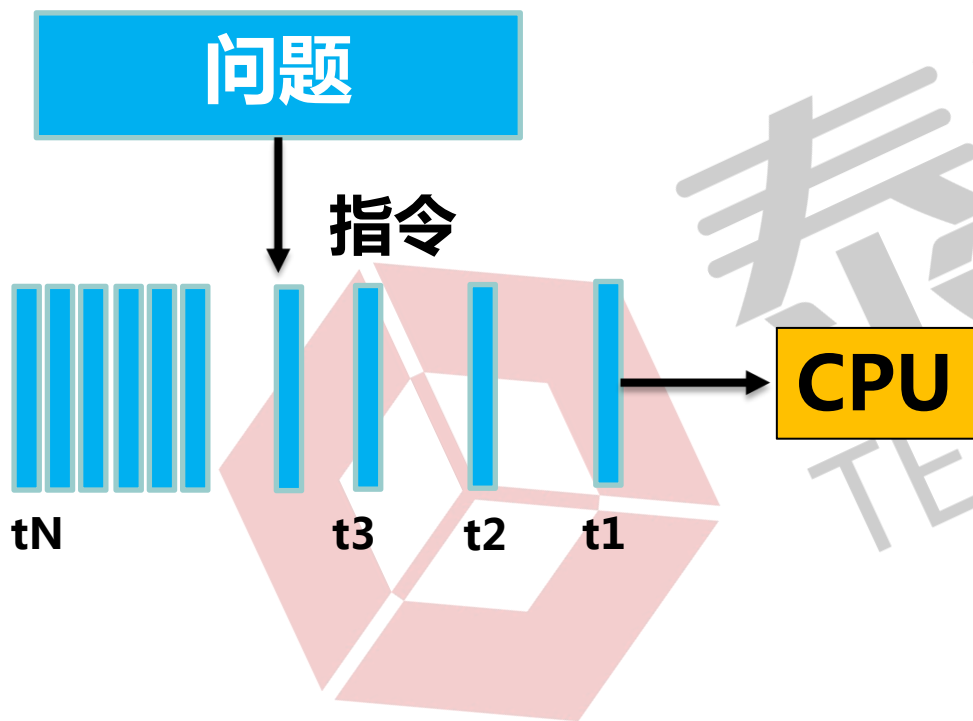
CPU



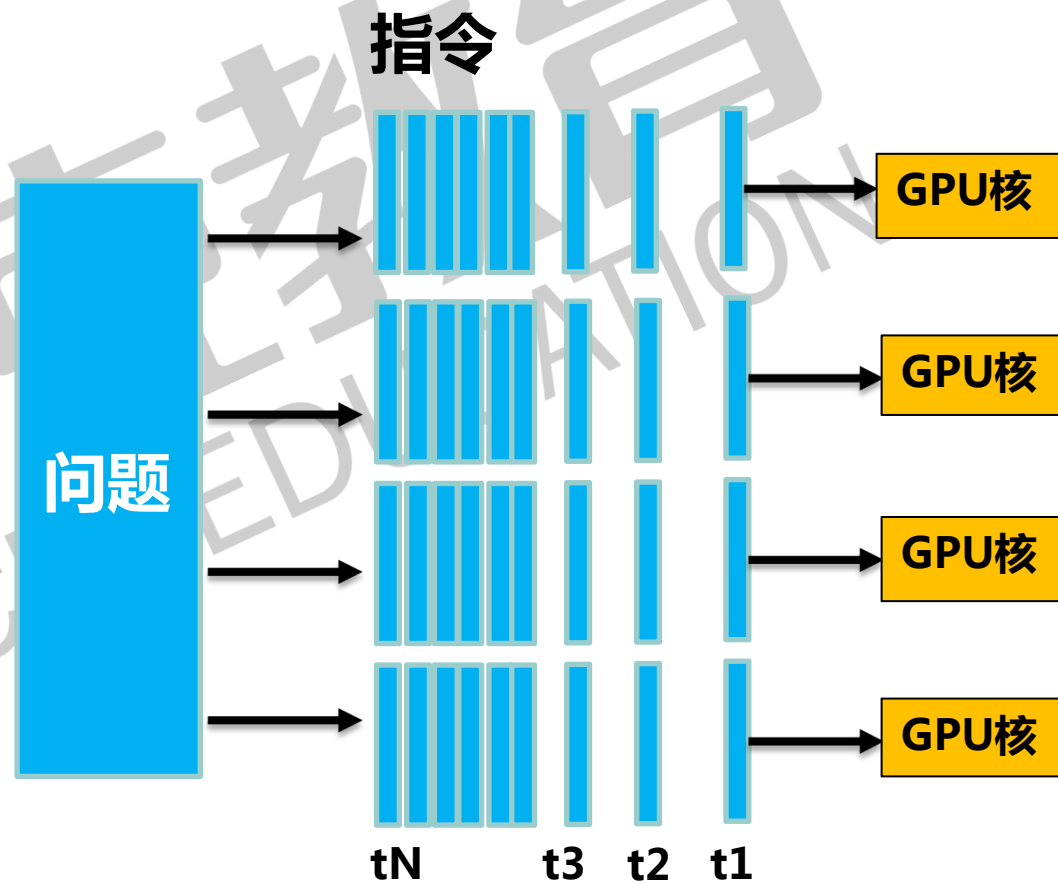
GPU



GPU指令执行模型



CPU指令执行模型



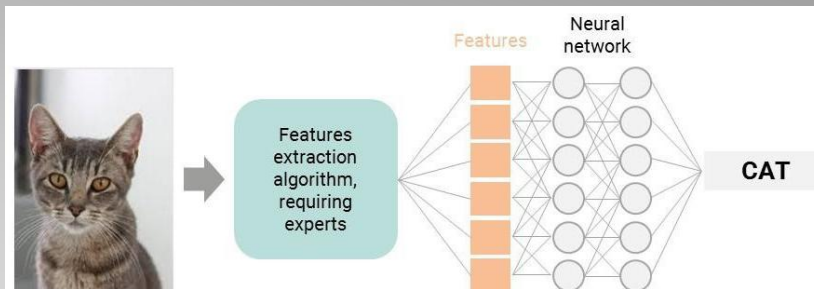
GPU指令执行模型



GPU 加速应用场景



2006年，多伦多大学教授
Geoffrey Hinton发表论文开启
深度学习。



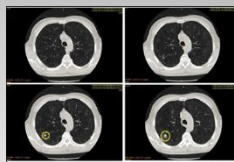
2012年6月，谷歌Brain项目用16,000个CPU
从10万个Youtube视频截取图片通过深度学
习学会什么是“猫”。



2016年3月，谷歌AlphaGo以
4：1战胜李世石。



智能金融



智能医疗



平安城市



智慧交通



机器翻译



语音识别

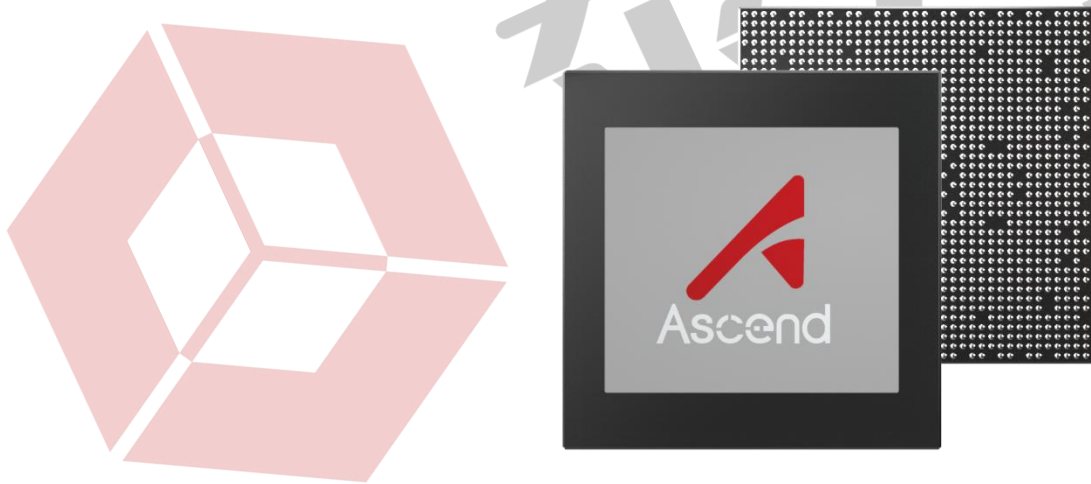


自动驾驶



什么是ASIC

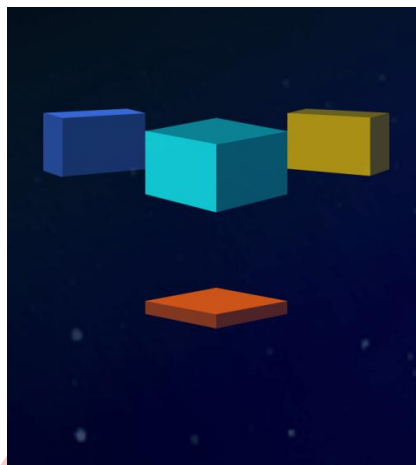
- ASIC是一种专用芯片，与传统的通用芯片有一定的差异。是为了某种特定的需求而专门定制芯片。ASIC芯片的计算能力和计算效率都可以根据算法需要进行定制。
- 优点：体积小、功耗低、计算性能高、计算效率高、芯片出货量越大成本越低。
- 缺点：算法是固定的，一旦算法变化就可能无法使用。





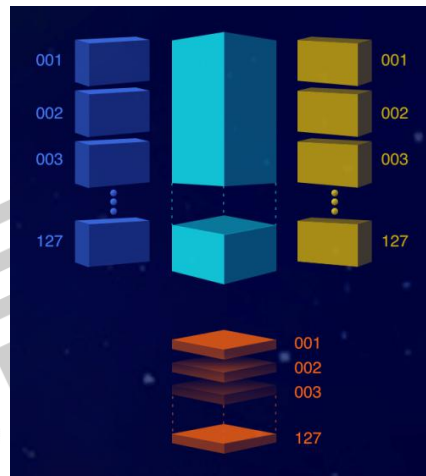
华为达芬奇架构

传统处理器：标量计算



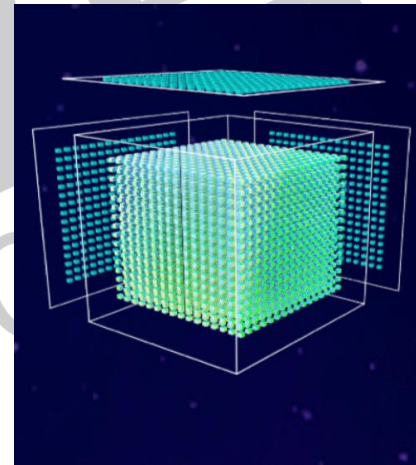
0.003 TOPS/W

传统加速卡：向量计算



0.3 TOPS/W

达芬奇架构：张量计算

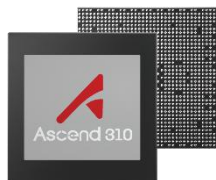


3 TOPS/W

华为独创 3D Cube
16*16*16 三维弹性立方体



昇腾系列AI处理器



Ascend 310

Ascend-Mini

架构: 达芬奇

半精度 (FP16) : 8 TeraFLOPS

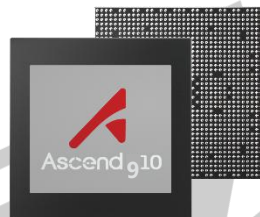
整数精度 (INT8) : 16 TeraOPS

16 通道全高清视频解码器 – H.264/265

1 通道全高清视频编码器 – H.264/265

功耗 : 8W

12nm



Ascend 910

Ascend-Max

架构: 达芬奇

半精度 (FP16): 256 TeraFLOPS

整数精度 (INT8): 512 TeraOPS

128 通道全高清视频解码器 – H.264/265

最大功耗: 350W

7nm



异构计算总结



CPU

- 主频高，核数有限；
- 逻辑控制&算术运算单元、大量缓存；
- 管理和调度任务。



GPU

- 高并发：几千核并行；
- 强浮点能力；
- 高显存带宽：芯片内置HBM存储。



FPGA

- 主频低但集成大量计算单元；
- 流水线并行和数据并行；
- 硬件编程和加速，特定应用IP核。

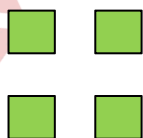


ASIC

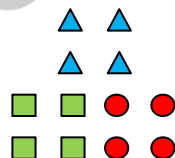
- 针对某一场景优化的专用处理单元；
- 硬件基本不可编程；
- 多个IP集成，高性价比和能效比。



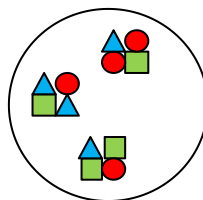
单核



多核/众核



异构



资源池化



思考题

多选题

1. 相比其他的处理器，FPGA具有的优势包括（ ）。

- A. 并行能力强
- B. 通用编程
- C. 比ASIC能耗低
- D. 实时性好

判断题

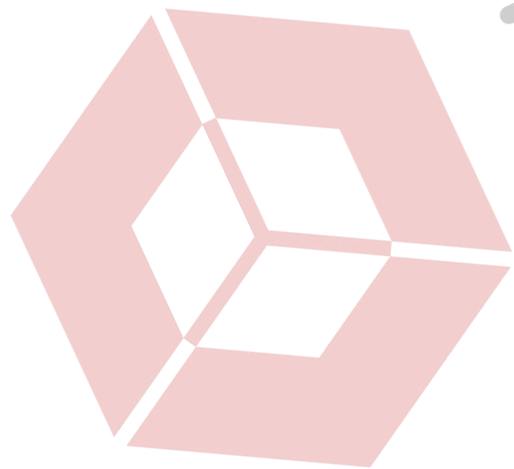
2. 复杂指令集（CISC）面积小，功耗低。



本章总结

学完本章之后，您将能够了解

- 处理器芯片的概念以及发展历程；
- 处理器的分类；
- 异构计算的分类和特征。

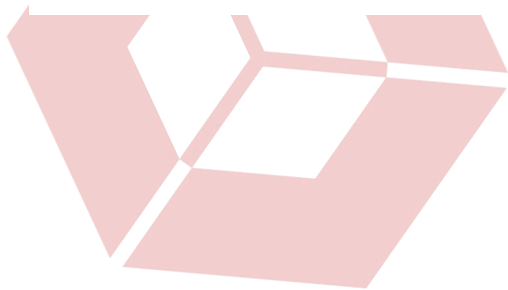


泰克教育
TECH EDUCATION



更多信息

- 更多信息请参考如下二维码





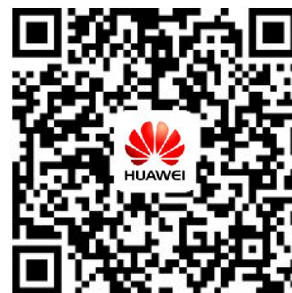
学习推荐

- 华为官方网站

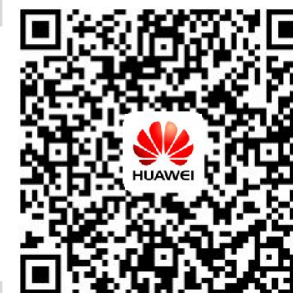
- 在线学习：<https://e.huawei.com/cn/talent/#/home>
- 企业业务：<http://enterprise.huawei.com/cn/>
- 技术支持：<http://support.huawei.com/enterprise/>

- 热门工具

- HedEx Lite
- 网络资料工具中心
- 信息查询助手



技术支持



热门工具

The background of the slide features a blue-tinted image of several business professionals in a modern office. They are standing on a highly reflective floor, and their silhouettes are clearly visible. In the background, a city skyline with various skyscrapers is visible through large windows. The overall atmosphere is professional and corporate.

谢谢

www.huawei.com