



ZOMI

# 集群计算 发展历程

# Content



# AllInfra Architecture

AI 系统 + 大模型全栈架构图

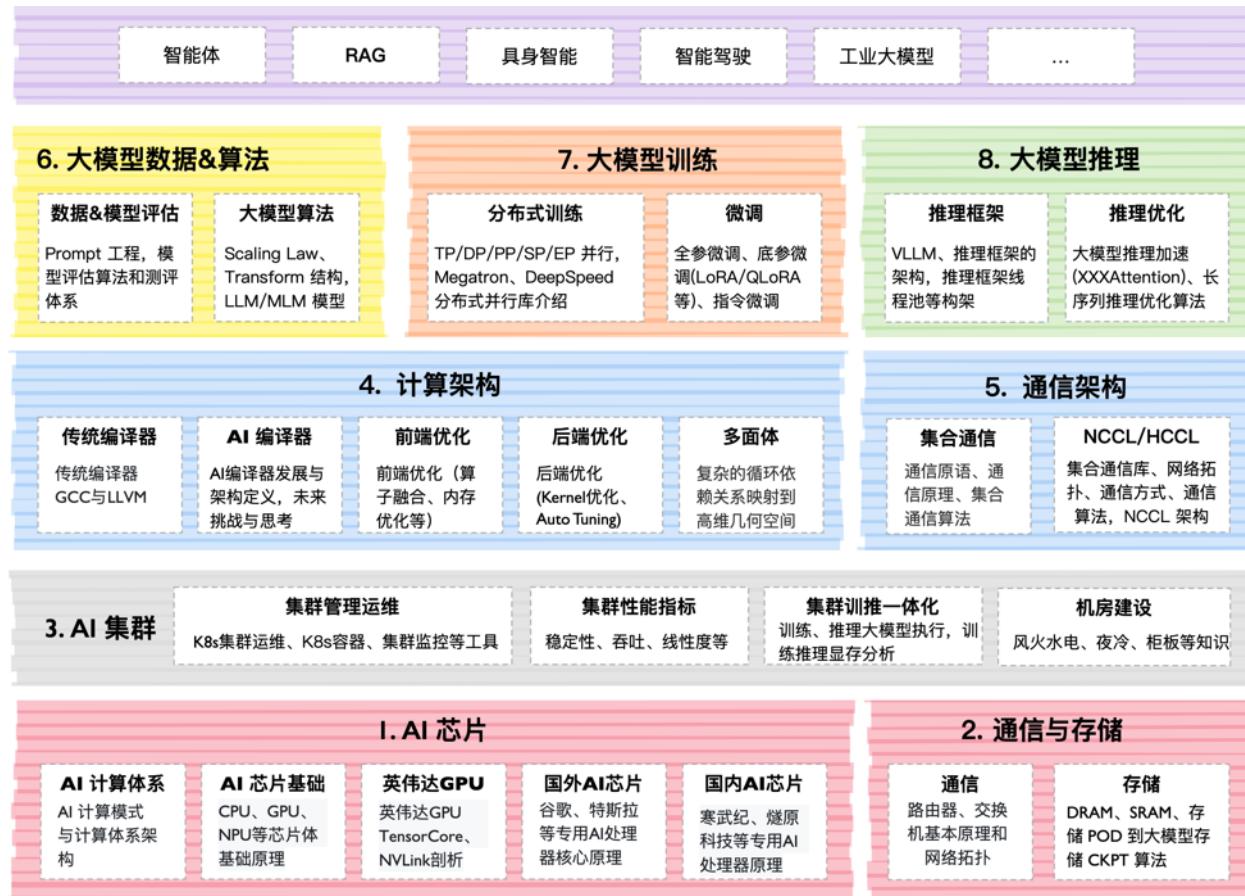


时事  
热点

大模型  
训推

编译  
计算  
架构

硬件  
体系  
结构



# 三类计算集群的主要区别

集群类型	应用目的	数据类型	计算特征	网络特征	存储特征
HPC高性能计算	国防科研	FP64	计算密集, 高并行	密集通信 (每个应用流量特征不同, 部分可隐藏)	密集&复杂 IO
AI人工智能计算	AI训练推理	FP16/BF16/FP8/FP4		密集通信 (部分可隐藏)	密集 IO (按节奏迭代)
云数据中心	互联网云计算	INT32 FP32	通用计算, 高并发	分散通信	分散/密集 IO



# Content

- 计算产业发展历程：
  1. 奠基时代 1945~1970
  2. 通用架构 1971~1995
  3. 云&大数据 1995~2020
  4. 未来计算 2020~2030

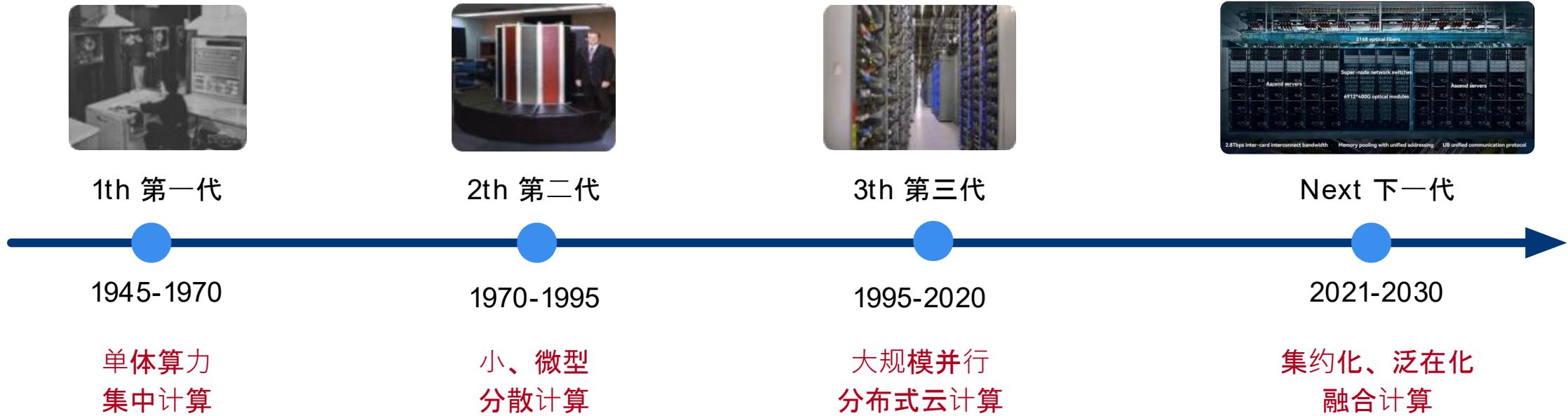


# 01

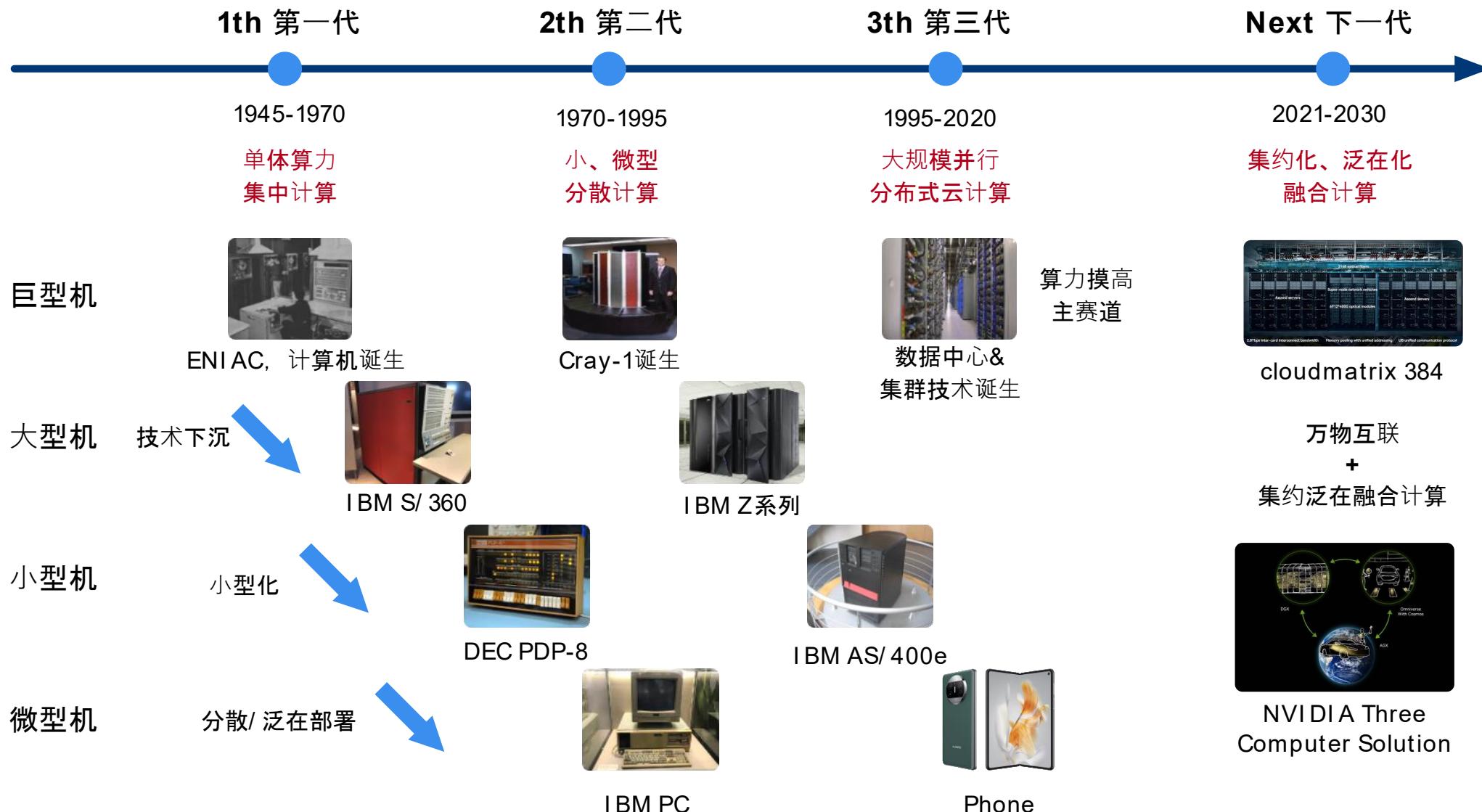
# 计算产业 发展历程

# 计算产业

- 计算需求驱动，从专用计算到通用计算，再到集约泛在融合的未来计算。



# 计算产业



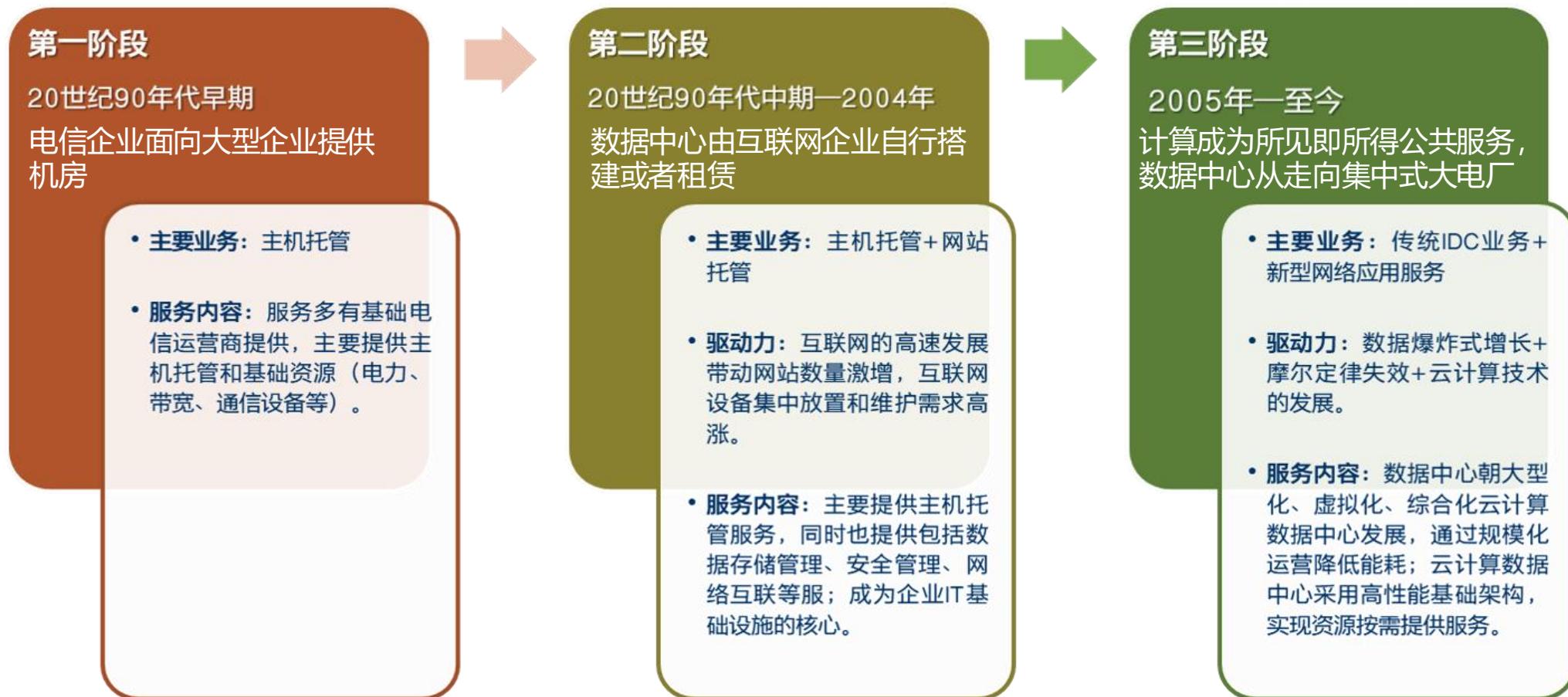
# 04

# 云与大数据

# 1995~2020

# 第三代：1995~，互联网/云计算兴起促使迎来数据中心时代

- 互联网出现，催生更多集中算力需求，传统“机房”升级为新形态“数据中心”，进入数据中心时代算力服务也更灵活多样。



# HPC：算力持续突破，单台大型机演变多服务器集群架构

## 单台大型机：超大内存&存储

单总线互连（全机一套OS），堆CPU增强算力

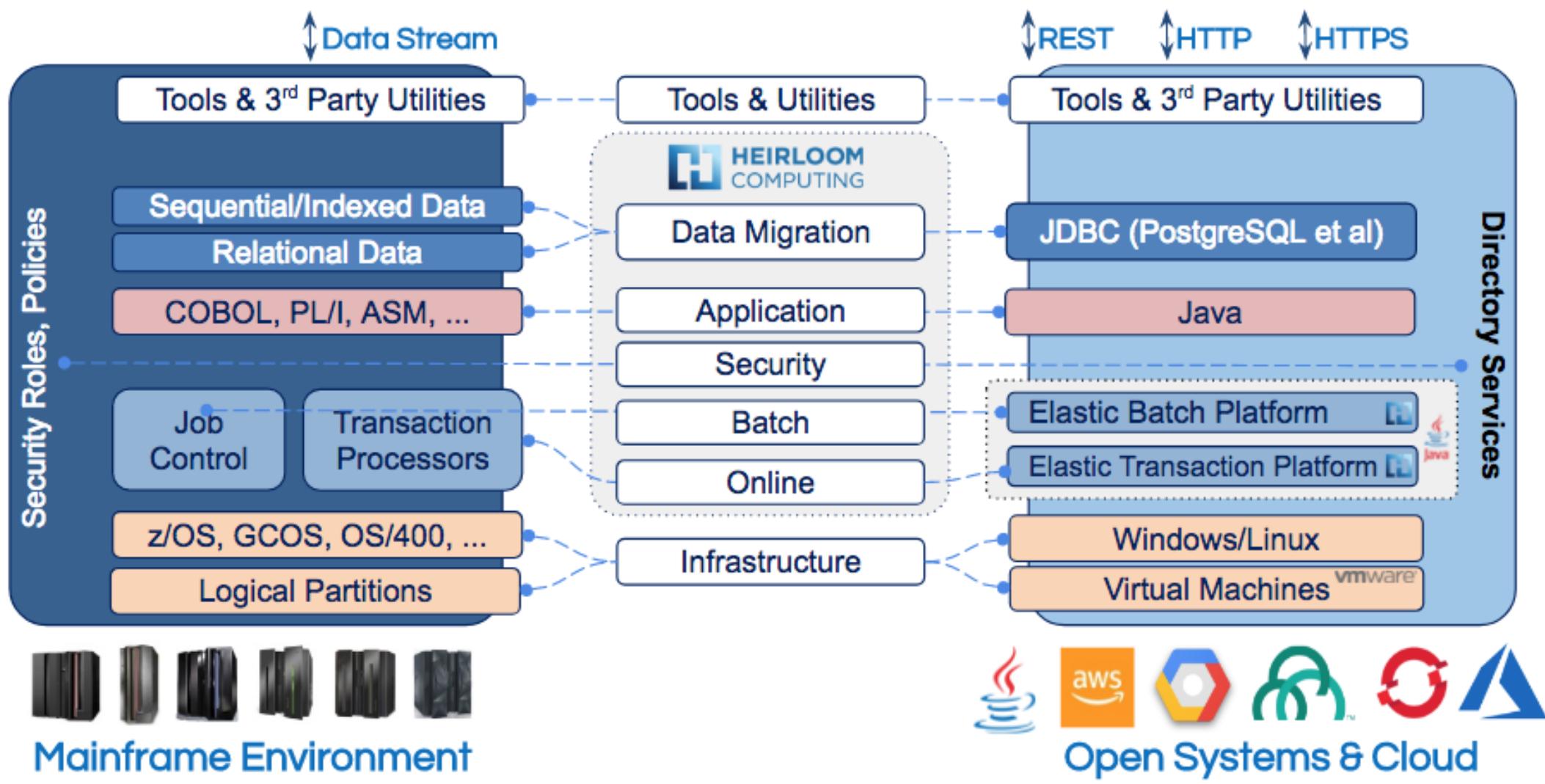


## 集群互连：多独立计算节点(每节点一套OS)

超高速网络大规模互连，实现超强算力+超强扩展



# 单台大型机 Mainframe Architecture



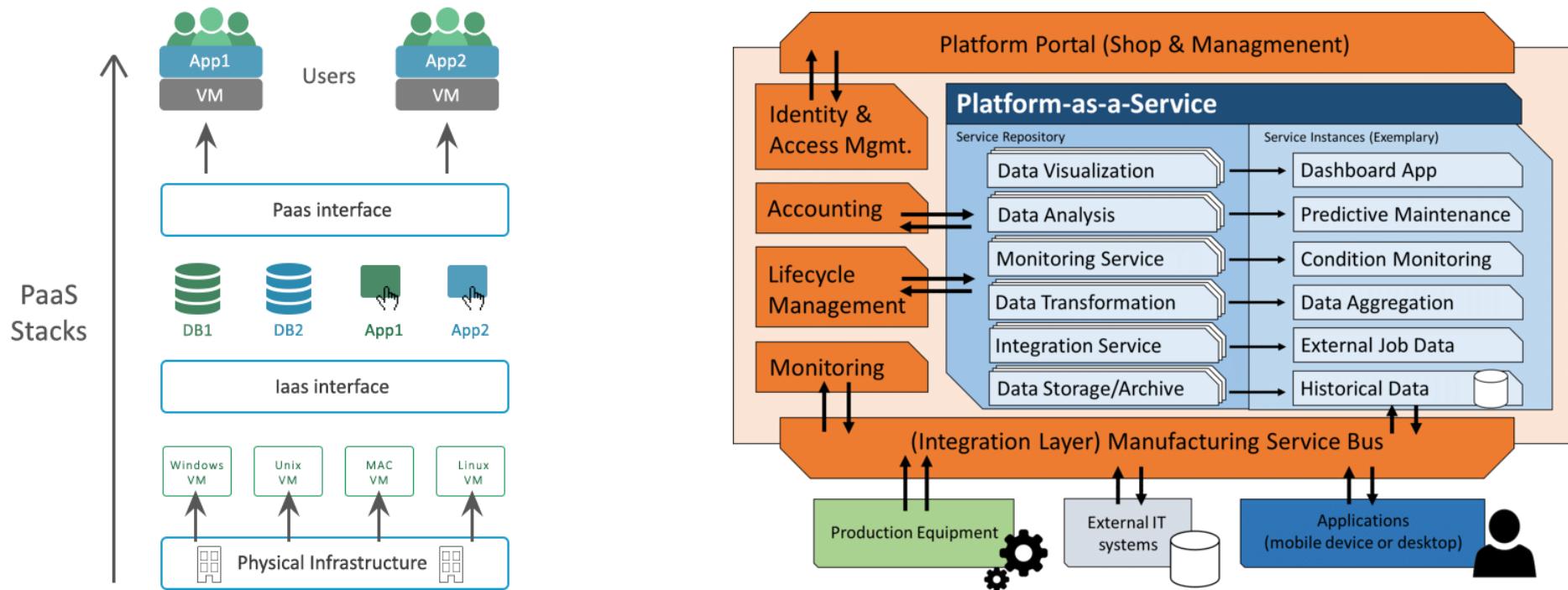
# DC时代 1th：IaaS 基础设施即服务雏形——主机托管

- 90年代随着互联网诞生和蓬勃发展，很多公司开始推行信息化。建设网站、搭建邮件、FTP、OA办公自动化等服务器。
- 云服务提供商（如 AWS、Microsoft Azure、Google Cloud Platform 等）负责提供并维护底层的 IT 基础设施。托管运营商机房，租用运营商场地/电力/网络带宽，代为管理和维护。



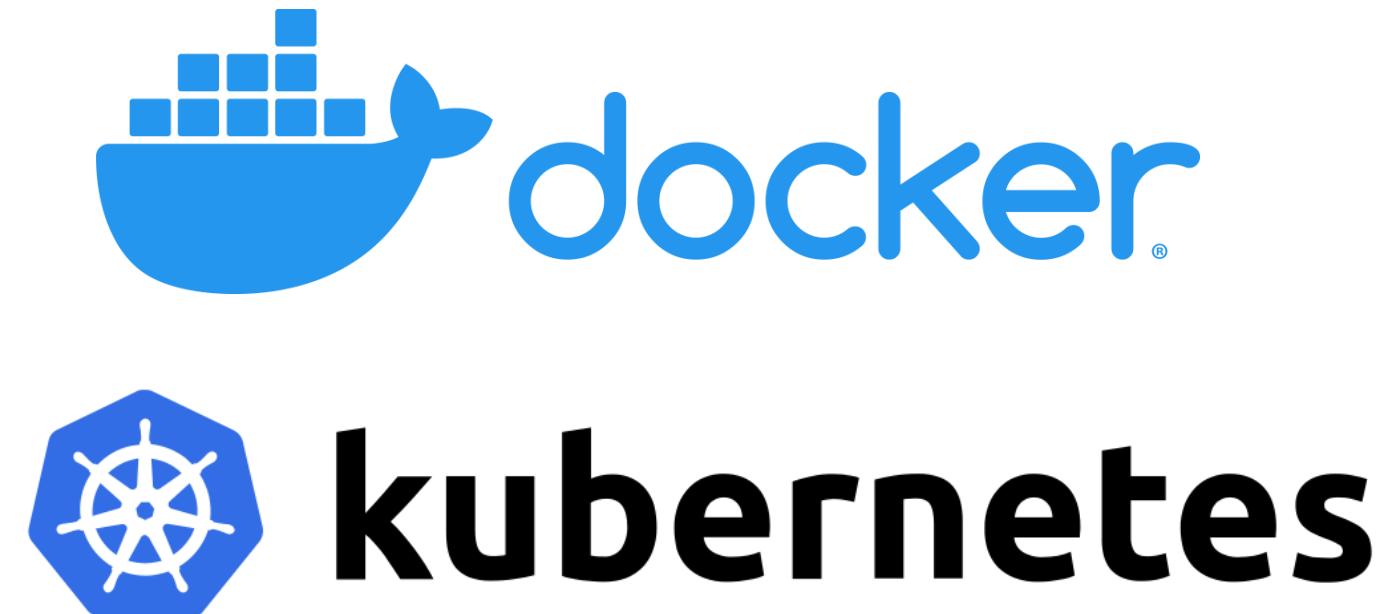
# DC 时代 2th： PaaS 平台即服务——虚拟主机/网站托管

- 某一台服务器上，通过虚拟主机软件，虚拟出N个网站主机，出租给N个客户使用。除网站之外，还出现了数据存储空间租用等多样化的服务。
- **PaaS (Platform as a Service, 平台即服务)** 云计算一种服务模式，为开发者提供完整开发和部署环境，用户可以在这个平台上开发、测试、部署和管理应用程序，而无需关心底层的基础设施。



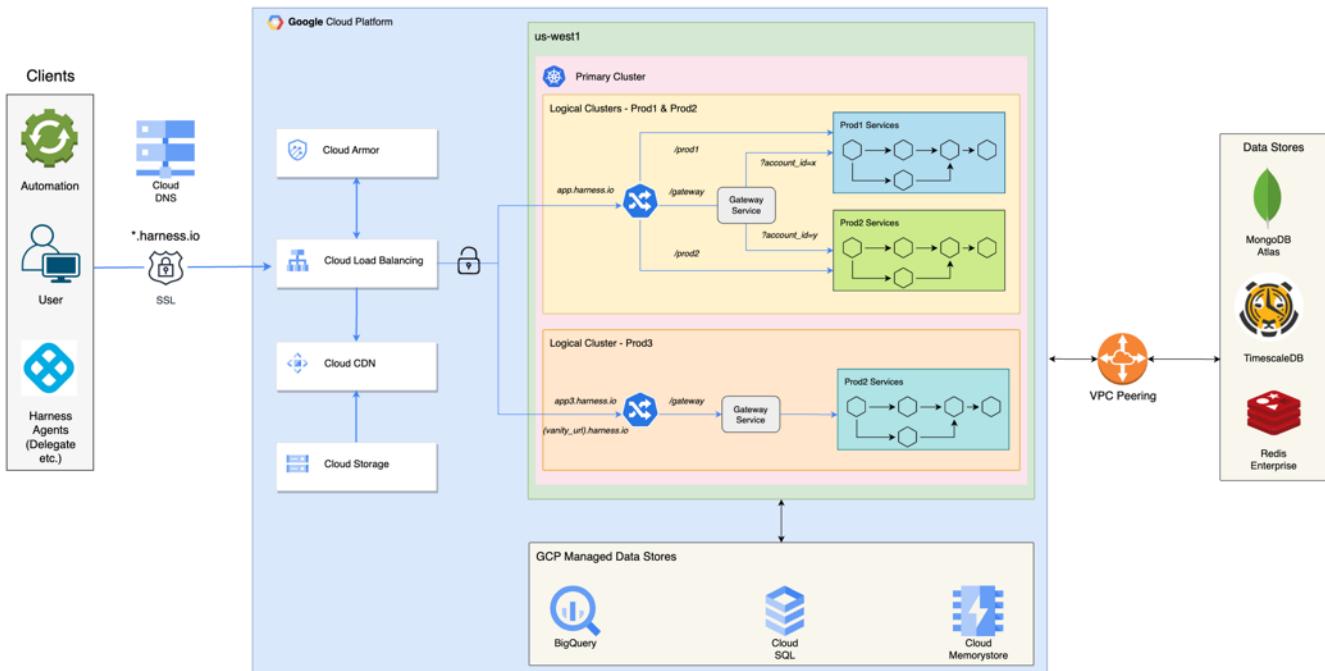
# DC 时代 2th： PaaS 服务雏形——虚拟主机/网站托管

- VMWare 第一个最成功 x86 架构虚拟化公司，开创新技术和产业方向：虚拟化和容器。促进了下一阶段云计算的产生；2022年6月，博通宣布将以610亿美元收购VMWare，并承担其80亿美元的净债务。
- Docker 和 K8s 当前以开源形式成为 PaaS 层跨平台容器化管理和容器编排平台服务的事实标准。



# DC 时代 3th: SaaS 软件即服务雏形——云计算服务

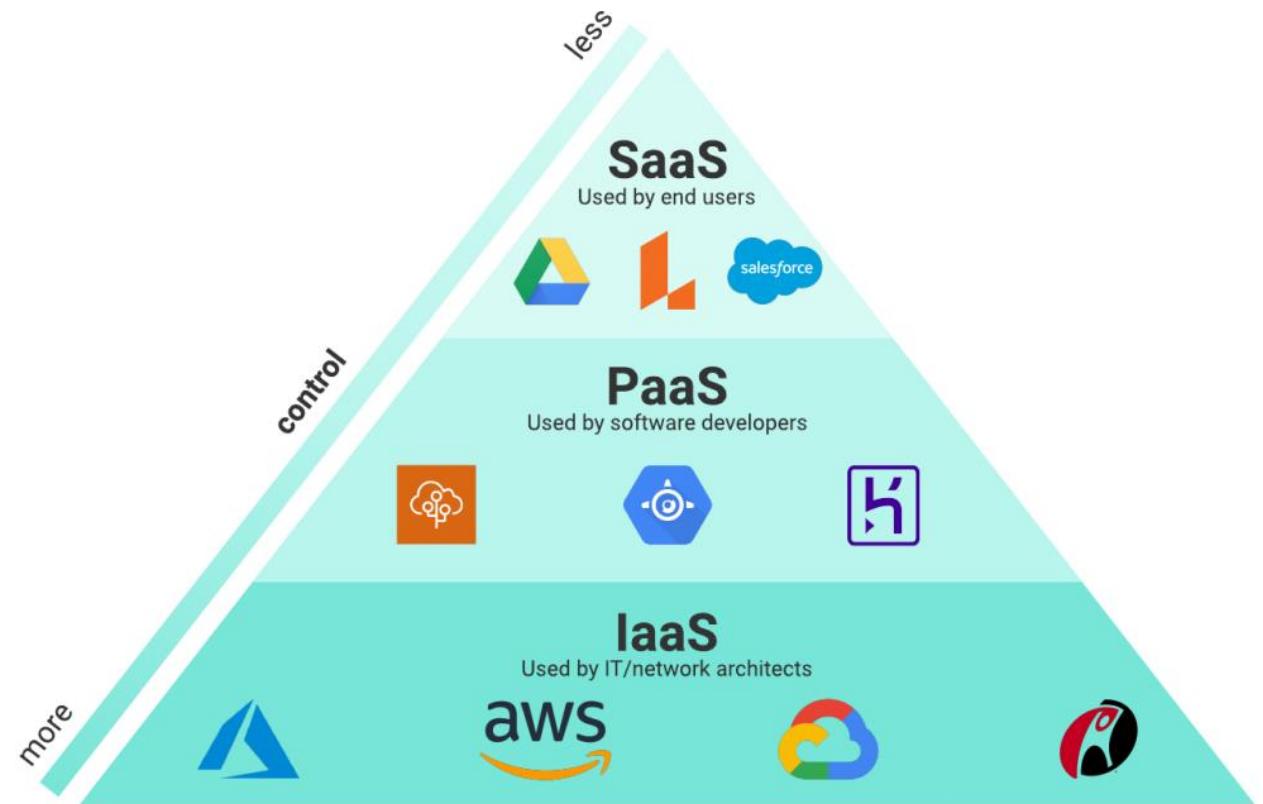
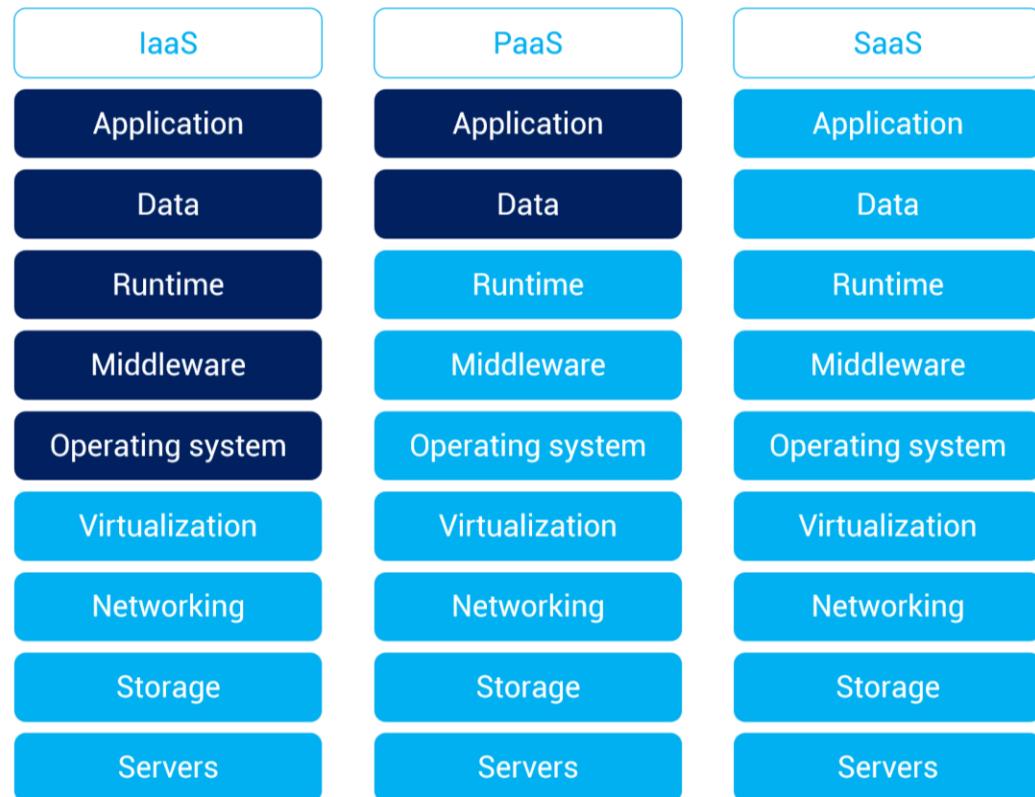
- 云计算阶段，虚拟化技术出现后，可以将同一台物理服务器虚拟成N个独立的虚拟机。既可以将虚拟机作为主机完整租借给客户，也可以将运行于虚拟机上的平台服务和软件服务提供给客户。
- SaaS (Software as a Service, 软件即服务)** 是云计算的一种服务模式，它通过互联网提供完整的应用程序，用户无需安装和维护软件，只需通过浏览器或客户端访问即可使用。



亚马逊云计算军规：一切接口都要  
呈现为Web Service

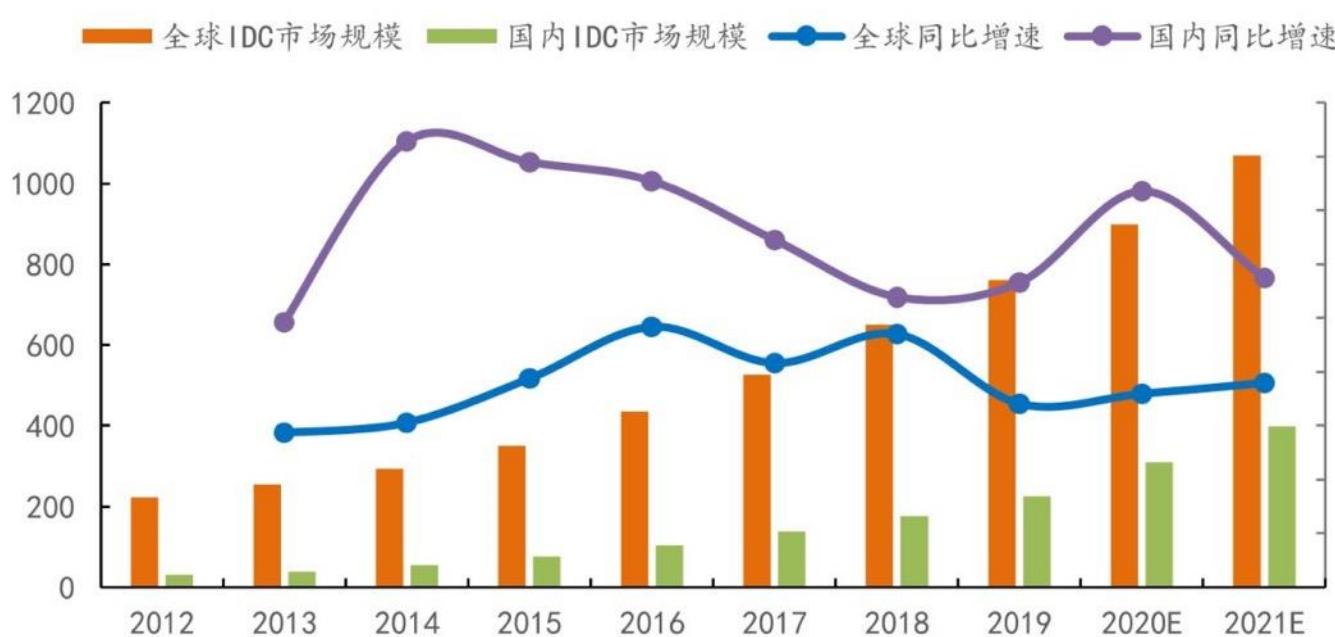
# DC 时代发展：IaaS、PaaS、SaaS

- IaaS、PaaS、SaaS 是云计算发展的三个层次，分别提供基础设施、开发平台和软件服务，推动企业实现资源灵活配置、应用快速开发与高效数字化转型。



# IDC互联网数据中心迅猛发展，算力爆发带来成本快速增加

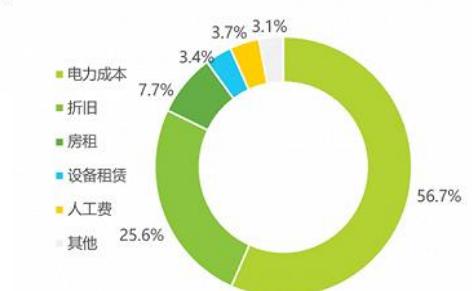
- 大数据时代驱动IDC行业发展，互联网上的数据每年将增长50%（不到两年翻番），目前世界上90%+的数据都是最近几年产生。
- 伴随着IDC快速建设，包含CAPEX、OPEX在内的TCO总成本也快速增加；中国数据中心每年总耗电约2千亿度。



数据中心建设成本拆分 (CAPEX)

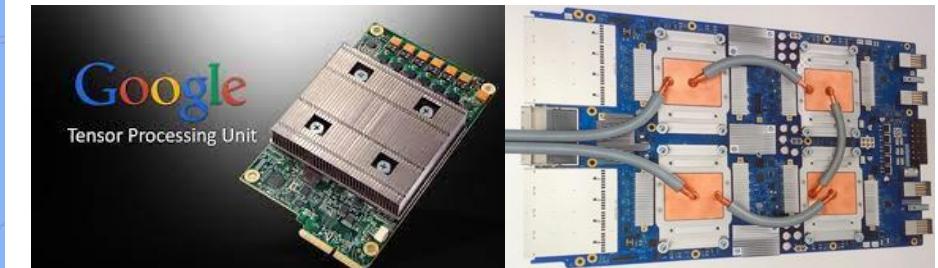


数据中心运营成本拆分 (OPEX)



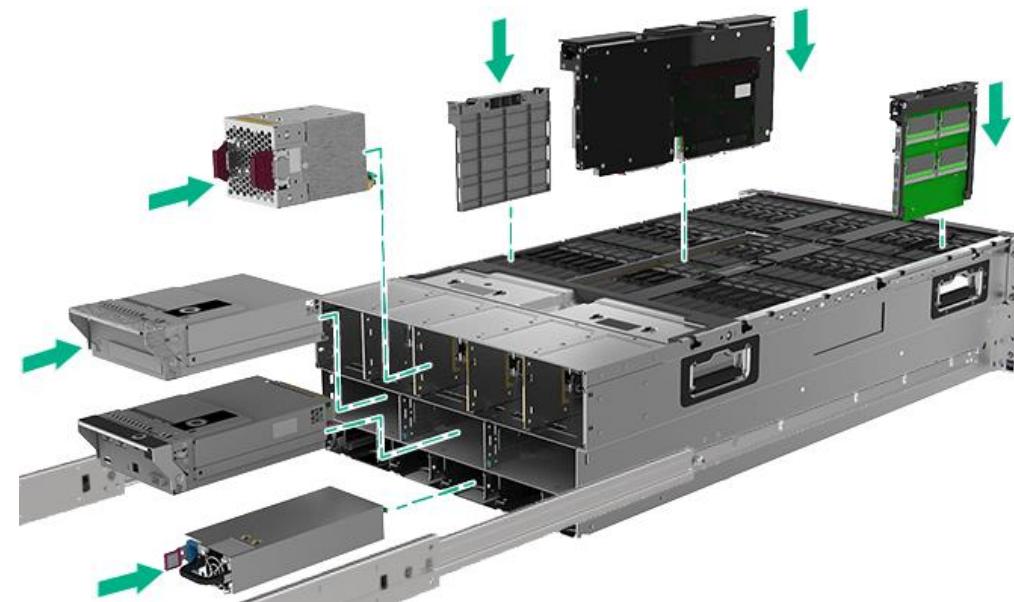
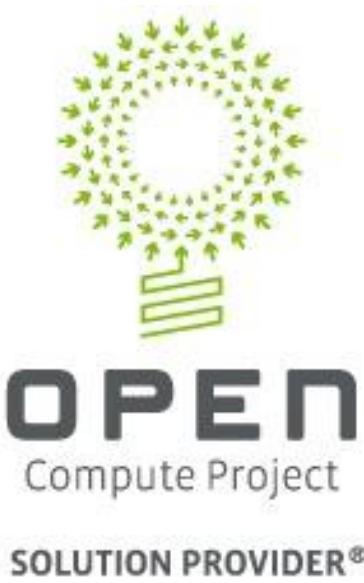
# 谷歌：基于cOTS货架化组件，自建大规模低成本数据中心

- 谷歌在全球自建23个超大规模数据中心，“规模化的共享经济效益” 收获了巨大的成功
- 建设方式：一体化Warehouse-Scale Computer仓储级DC + 节能联合设计，大幅降低TCO成本
- 在WSC仓储级数据中心内开创式的引入TPU异构加速技术，促成现代云计算+AI数据中心的成熟



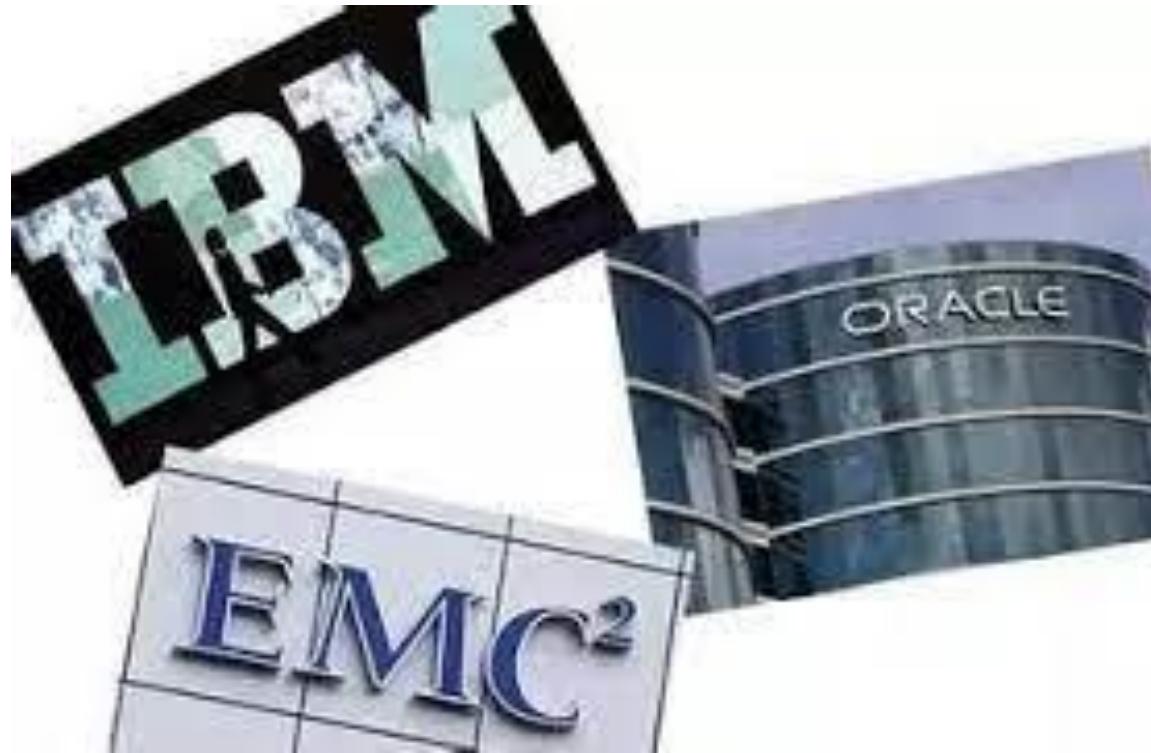
# Facebook OCP：硬件技术开源开放，重塑DC标准体系

- OCP 开放计算项目 Open Compute Project，Facebook 等 2011 年发起非营利组织。宗旨以开源开放方式，重构当前 DC 硬件，发展面向下一代 DC 服务器、存储、网络、基础设施等创新硬件。成员涵盖 Google、微软、百度、腾讯、阿里巴巴等，也涵盖 Intel、ARM、IBM、浪潮等方案供应商。
- 目前，在机架式和整机柜式高密度服务器、下一代 DC 管理架构等领域开放项目已经实现数百万台部署规模，AI、开源交换机、操作系统等新领域标准正在开放测试，是 HPC 最有影响开放技术项目。



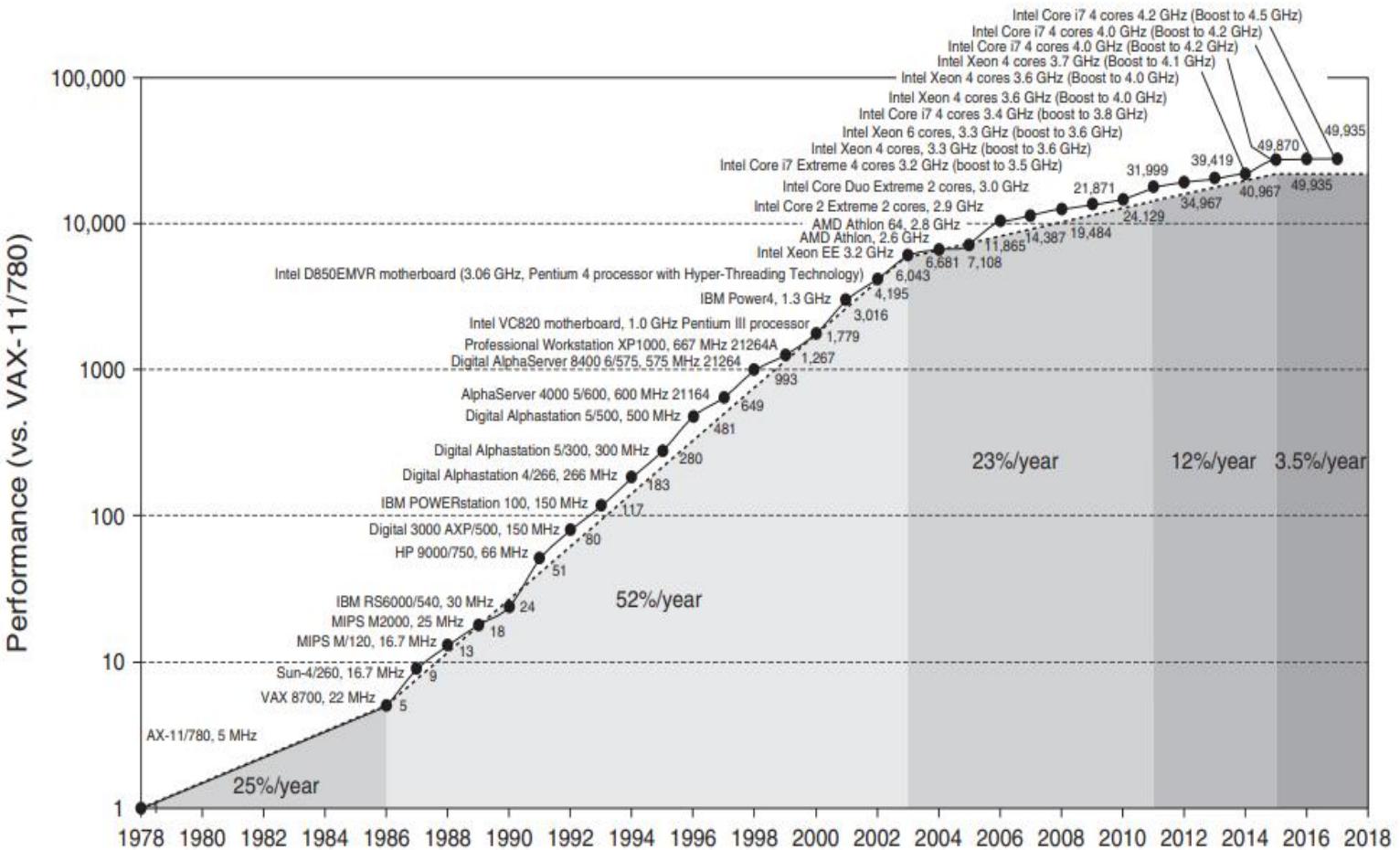
# 阿里：去 IOE，DC 从专有定制到通用开放、自主可用

- 去 IOE 用成本更加低廉软件， MySQL 替代 Oracle、 PC Server 替代 EMC2、 IBM 小型机等设备，以消除 IOE 对集中计算系统垄断。



# 伴随着各厂商加速创新，数据中心走到摩尔定律终结

- 未来计算负荷呈指数趋势增长，如果摩尔定律已经终结，唯一方式将是构建更多数据中心
- 将来趋势是增加数据中心工作负荷，数据中心计算能力将会以 10 倍的速度提升，其结果：
- 软件创新将会突飞猛涨，计算能力将进一步增强，会推动更多软件创新



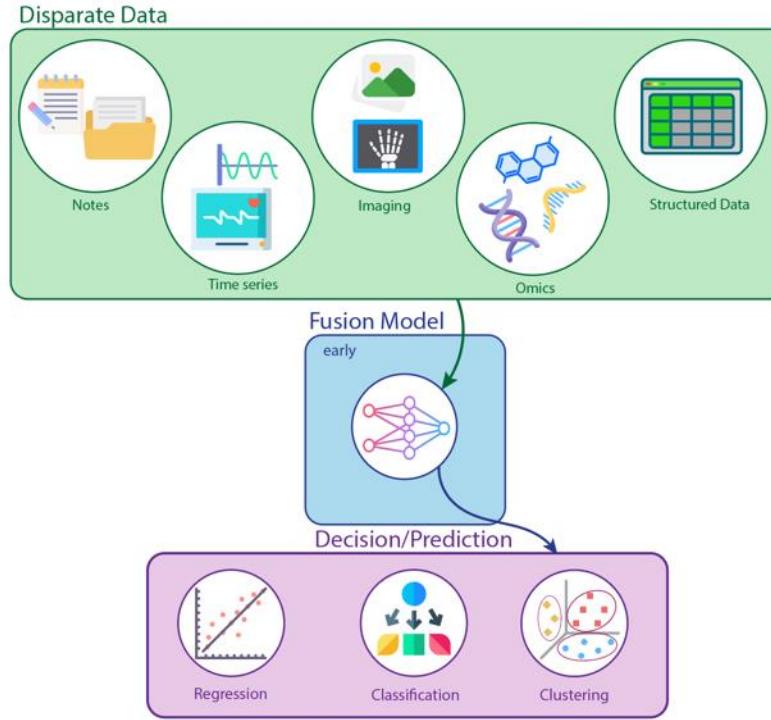
05

# 未来计算

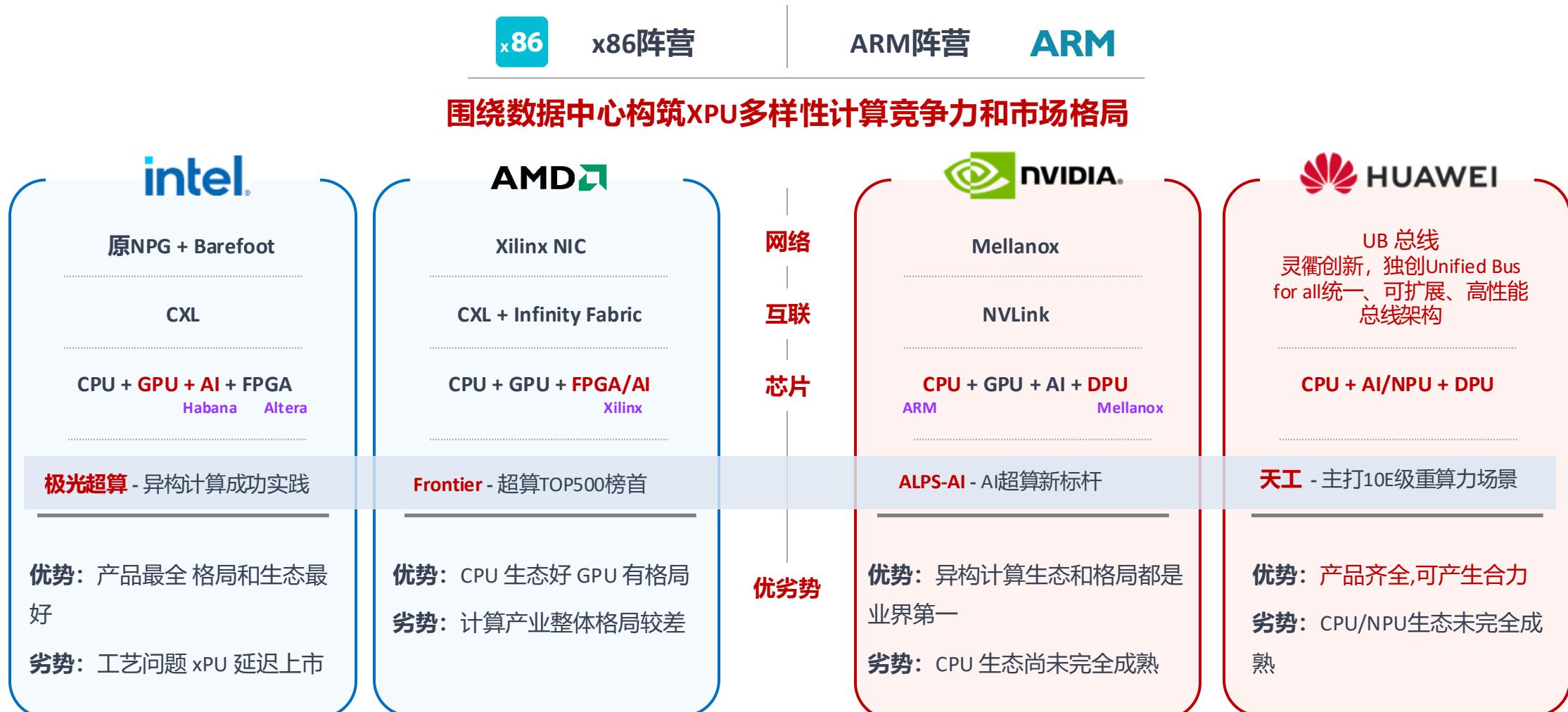
2020~2030

# 下一代：5.5G、AI、元宇宙等兴起驱动多样性计算融合需求

1. 算力形态不断丰富：整数计算、科学计算和渲染 FP64 计算、AI训练推理 FP16/FP8 低精度计算等
2. 多样性算力融合加速：AI 算力需求爆发，促使计算系统同时考虑 System for AI 和 AI for System
3. 多模态超异构融合加速：元宇宙对于通信、智能、渲染多模态超异构融合需求，加速多样性计算超融合



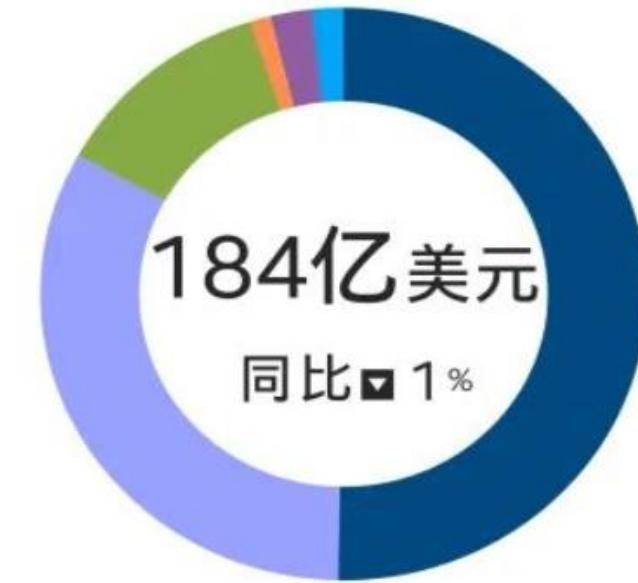
# 面向 DC 业务，各厂商拓展全栈关键技术，市场格局白热化



# Intel 从PC为中心转向DC为中心，翻覆打造新业务格局

- PC 市场渗透率登顶、移动互联网+云计算规模爆发， Intel 将业务布局更多的转向数据增长的主方向（围绕 DC 业务布局），这就是 “数据中心” 的关键转型
- Intel 拥有了完整的从软件到硬件、从通讯到计算到存储的计算架构、存储方案和连接方案。
- 未来十年架构创新将基于六大支柱成为计算创新的关键驱动力：软件、安全、互连、内存&存储、XPU架构、制程&封装

Intel 转型期各业务线收入结构

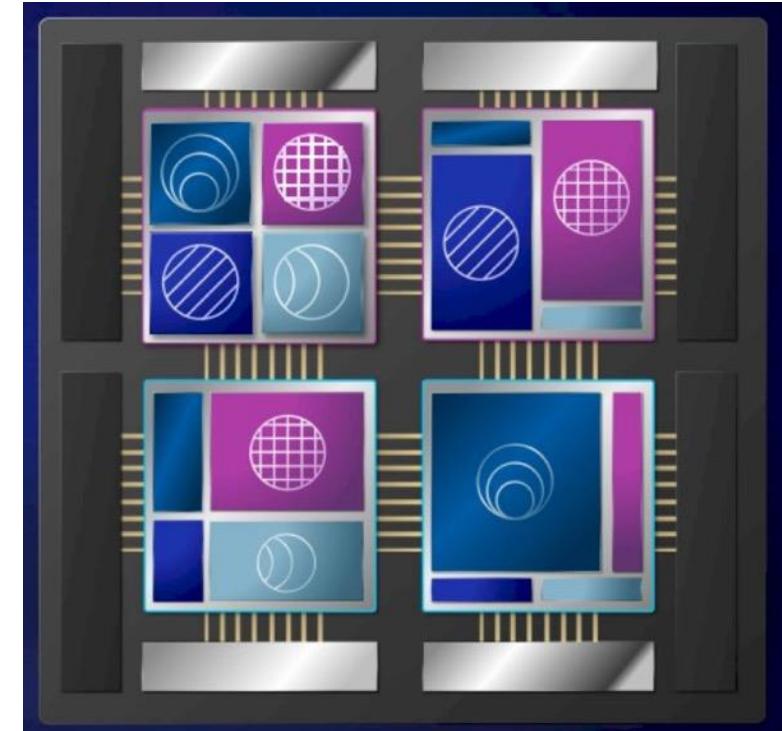


■ 客户端计算事业部 (CCG)	93亿 美元 同比 □ 13%
■ 数据中心和人工智能事业部 (DCAI)	60亿 美元 同比 □ 22%
★ ■ 网络与边缘事业部 (NEX)	22亿 美元 同比 □ 23%
■ 加速计算系统与图形事业部 (AXG)	2.19亿 美元 同比 □ 21%
★ ■ Mobileye	3.94亿 美元 同比 □ 5%
★ ■ 英特尔代工服务 (IFS)	2.83亿 美元 同比 □ 175%



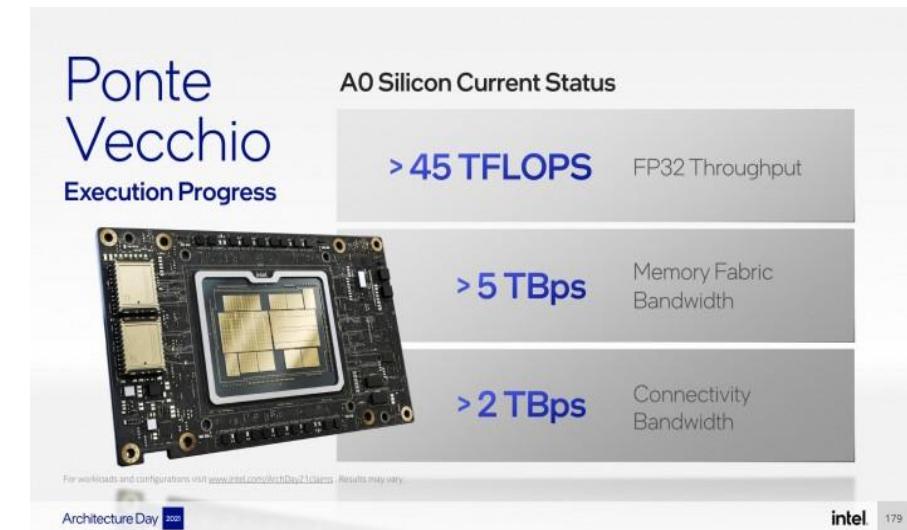
# Intel 面向计算、存储、网络全栈布局，欲计算时代重塑PC辉煌

- PC 时代春秋争霸，到 DC 时代战国争雄，Intel 在前两个时代竞争中脱颖而出，独霸天下
- 大爆发的移动互联网时代，Intel 的 Atom 移动处理器在与 ARM 架构市场竞争中败北
- 算力时代 Intel 增速持续下滑，ARM 和 RISC-V 等新势力将构成越来越大的挑战



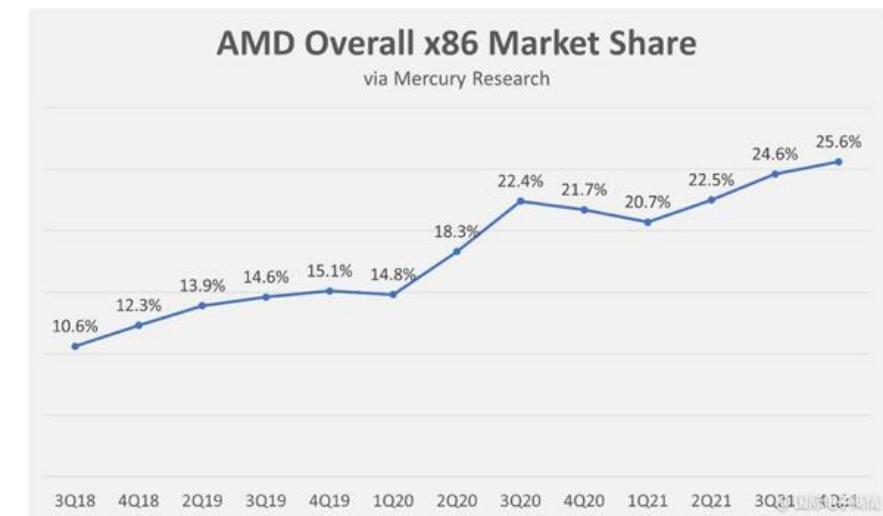
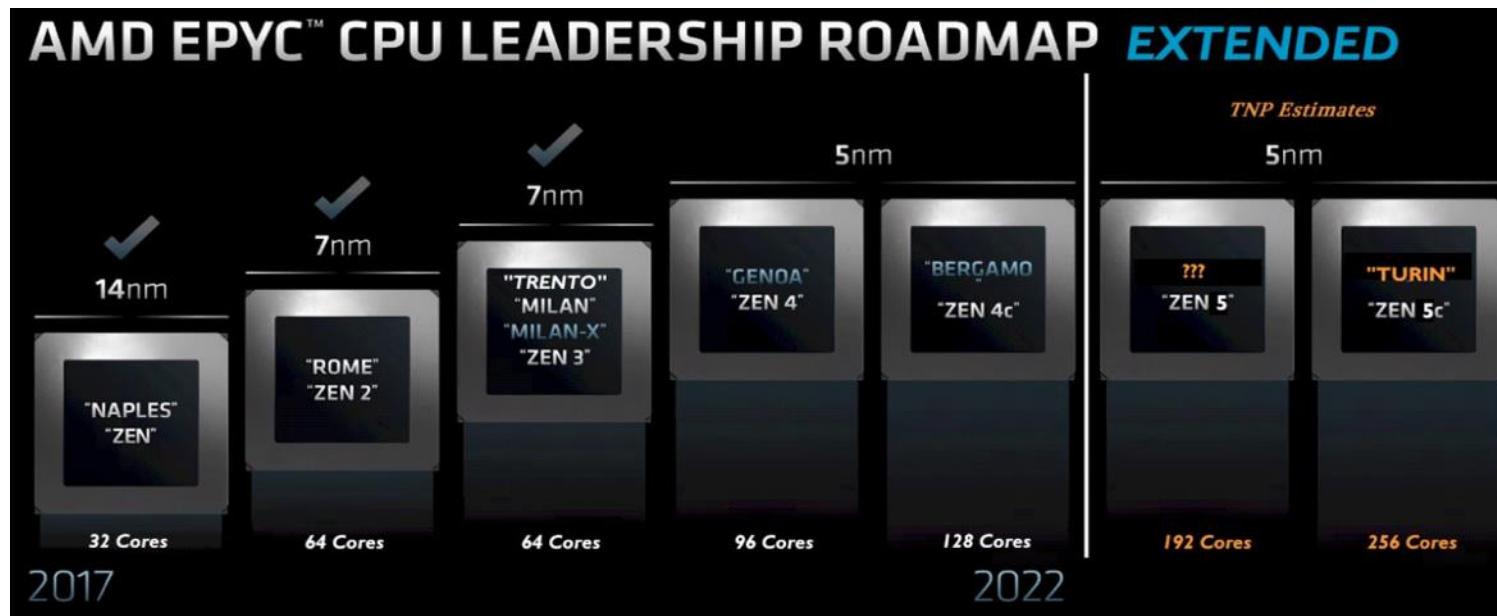
# Intel 异构计算极光超算，峰值算力将突破2E FLOPS

- CPU + GPU + oneAPI 异构计算架构，将使能极光超算峰值算力 突破2E FLOPS



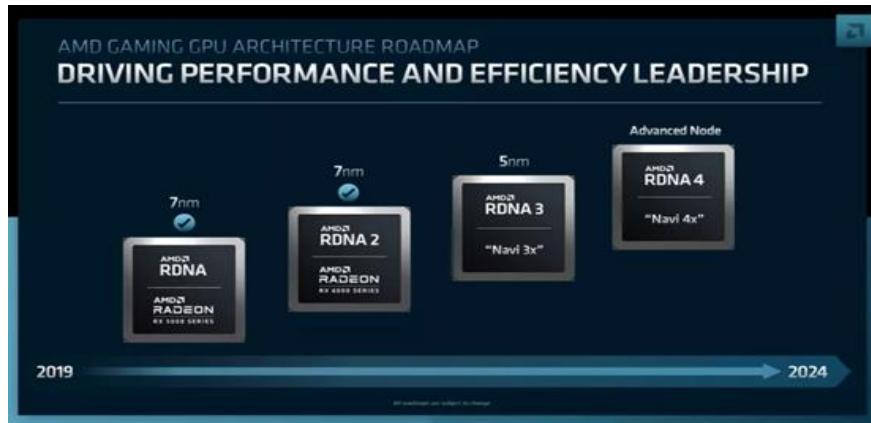
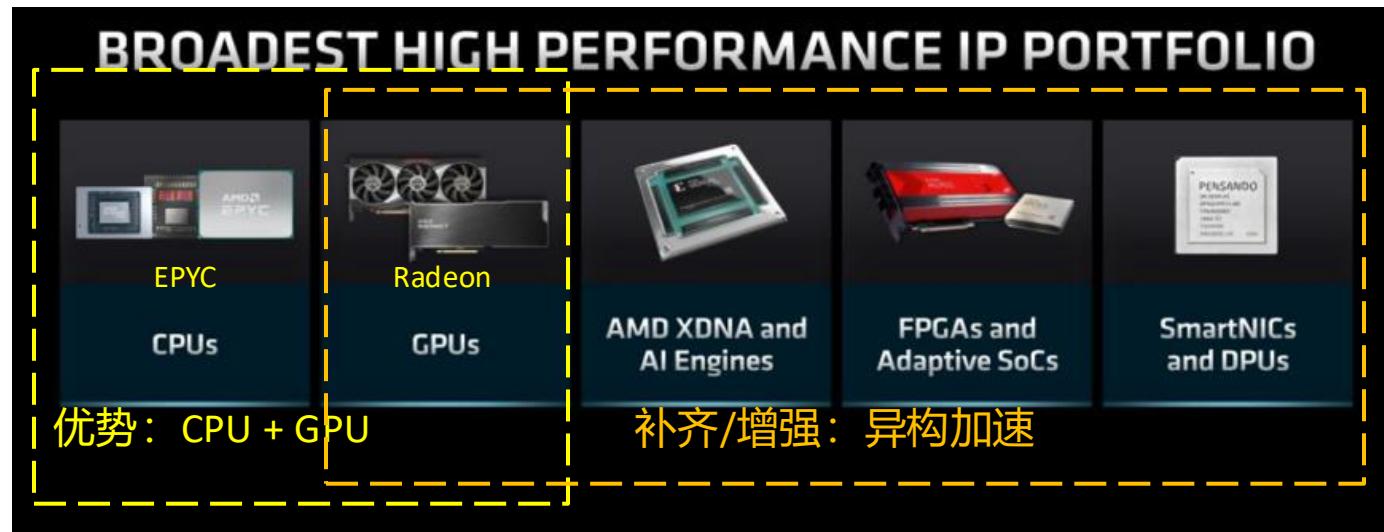
# AMD 霄龙服务器制程优势明显，增长威胁Intel垄断地位

- AMD 服务器演进路线落地顺利，Fabless 制程优势和交付进度持续领先 Intel，伴随市场捷报频传，逐步威胁 Intel 在 DC 市场统治地位
- x86 总体市场 AMD 市场份额涨势相当迅速，从 2018 年仅 10% 市场到 25.6%，近年更有加速上涨趋势

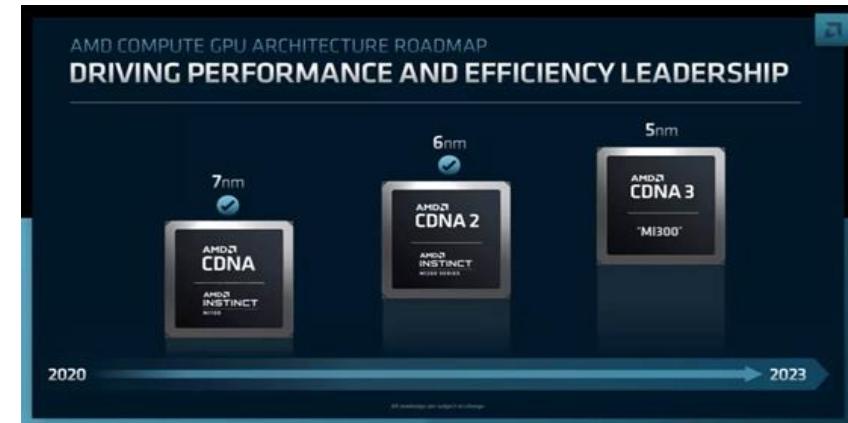


Intel DCG 收入持续下滑 AMD 高歌猛进

# AMD 基于霄龙 CPU 和镭龙 GPU 优势，增强发力异构加速



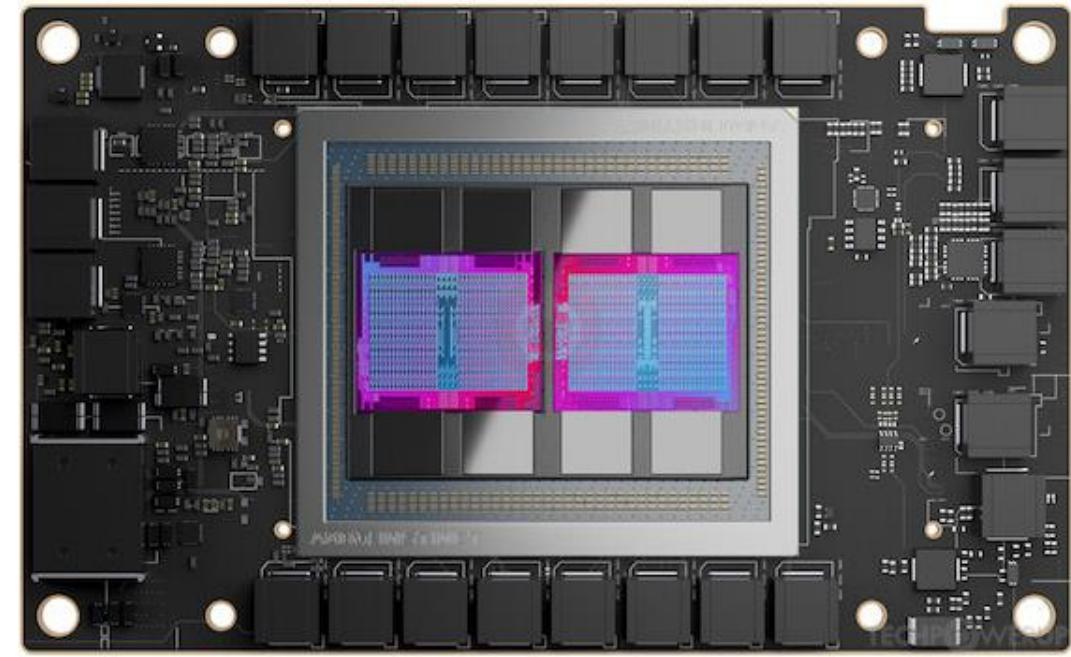
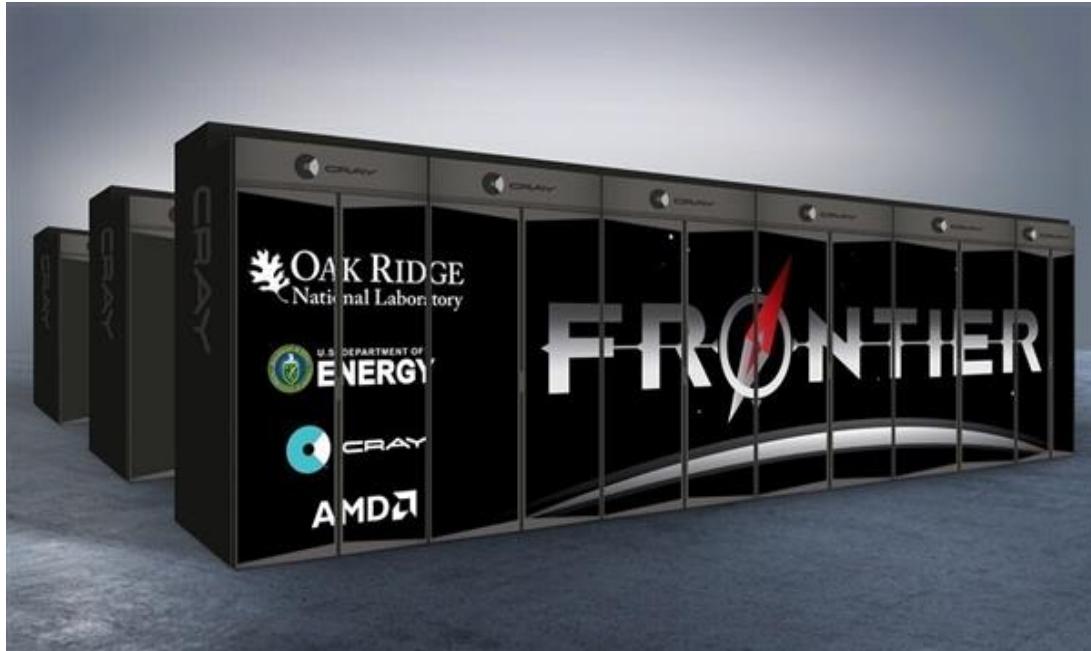
5nm GPU 芯片 RDNA3 将于 2022 推出



5倍性能/瓦特提升 MI300 APU (CPU+GPU 组合封装)

# AMD 融合计算架构助力 Frontier 超算实现“E级算力”

- AMD 与美国能源部、橡树岭国家实验室和 Cray Inc. 公司通力合作，打造新一代超级计算机 Frontier，峰值处理能力为 1.685EFLOPS，拿下超级计算机 TOP500 排行榜 2022 年上半年榜单第一名
- 该系统具有 9408 个CPU+GPU融合的计算节点，每个节点配备一个 64 核 AMD “Trento” CPU 和四个 AMD Radeon InstinctMI250X GPU



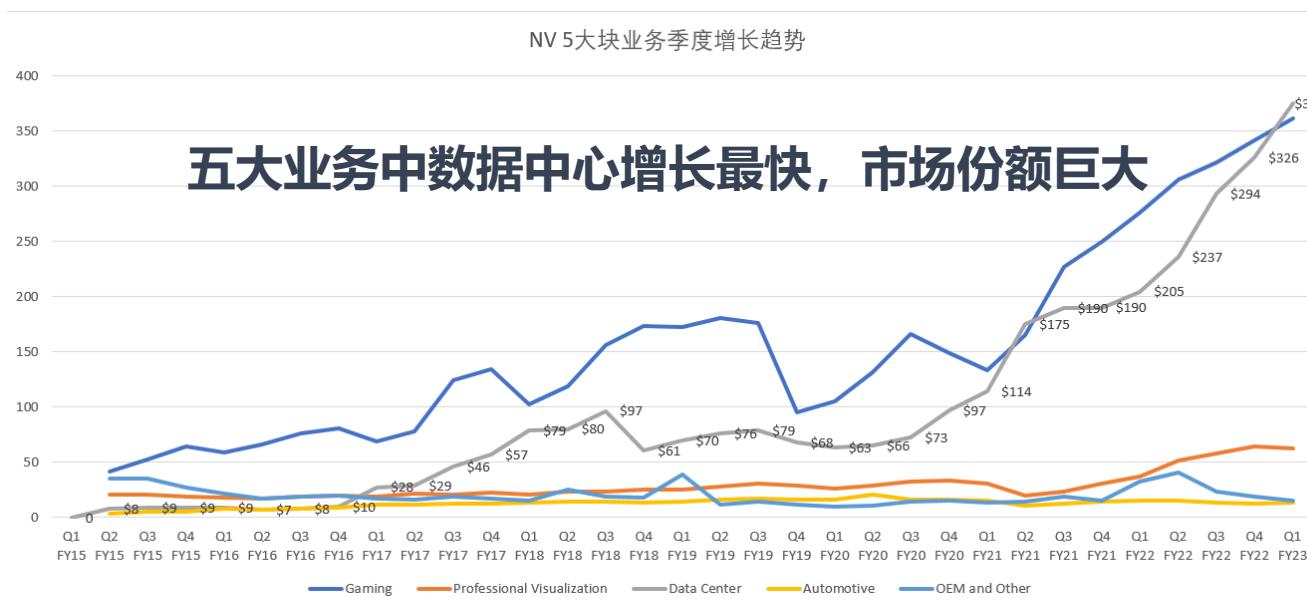
# NVIDIA 数据中心战略成为其最重要发力方向

- NVIDIA 成立于 1993 年，以游戏独立显卡起家，1999 年推出世界第一款替代 CPU 渲染图形处理单元（GPU） GeForce 256，站稳行业龙头地位。
- 尽管 NVIDIA DC 业务 2019 年起步，但增速极快，4 年复合增长率高达 69%，远超游戏业务 18% 复合增长。



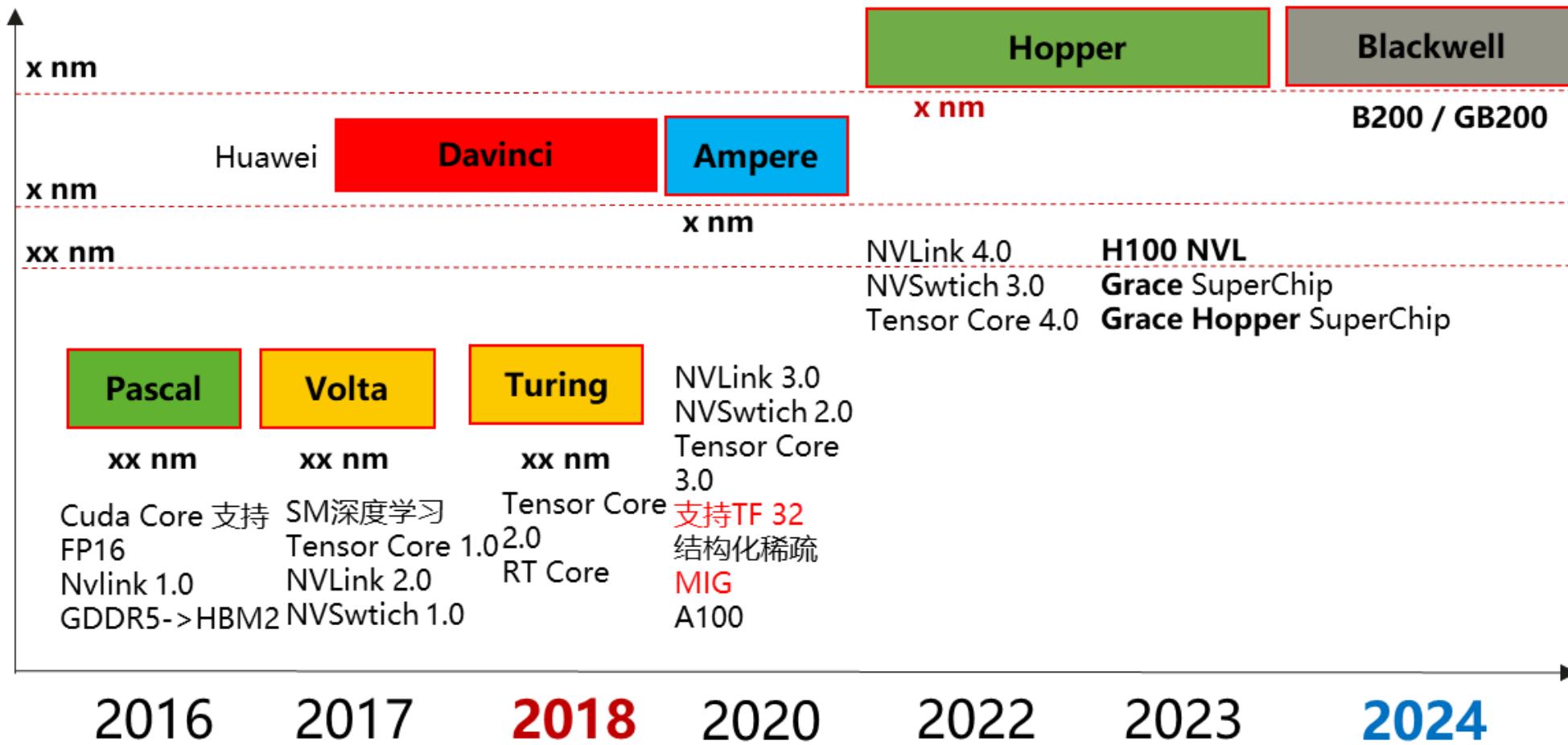
# NV 从游戏入手独辟 GPU 计算，异构并行加速 AI 生态

- 要打造 GPU 计算生态，没有软件应用，没有人从 CPU 转移 GPU 生态。仅仅 1~2 倍加速对 CPU 生态开发者吸引力不大。GPU 提供新计算平台，说服 CPU 计算软件转移到 GPU 生态不容易，其采取策略：
  - 生态合作伙伴 INCEPTION 项目，让 AI 创业公司都参与到这个项目
  - 设立了风险投资的股权基金，投资小部分的AI公司。目标不是财务回报，而是培养生态



# NVIDIA 在人工智能服务器和集群形态的发展

## 英伟达芯片代际演进图



# NVIDIA 在人工智能服务器和集群形态的发展

## DGX-1 服务器, 2016.4

基于Pascal-Volta的8卡服务器  
《规模化AI技术的行业标准》



- 170TFlops(Pascal)、960TFlops(Volta) @ FP16
- 512 GB of DDR4-2133
- 4 x 1.92 TB SSDs
- 2\* 10GE组网
- 3.2KW供电, 3U高度

## DGX-2 服务器, 2018.3

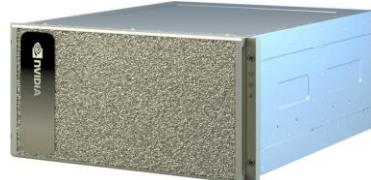
基于Volta的16卡2PFlops服务器  
《突破AI速度和规模的限制》



- 2PFlops @ FP16
- 1.5TB DDR4, 512GB HBM2
- 30.72 TB SSDs
- 8 \* 10Gbps InfiniBand组网
- 10KW供电, 10U高度

## DGX A100 服务器, 2020.5

基于Ampere的8卡第3代服务器  
《统一了所有AI工作负载》



- 5PFlops @ FP16
- 2TB DDR4, 640GB HBM3
- 30 TB U.2 NVMe(8 \* 3.84TB)
- CX-6 / CX-7 InfiniBand / RoCE组网
- 6.5KW供电, 6U高度

## DGX H100 服务器, 2022.3

基于Hopper的8卡第4代服务器  
《AI基础架构的黄金标准》



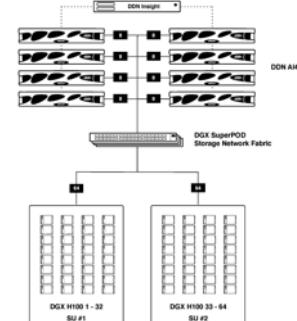
- 32PFlops @ FP8
- 2TB DDR4, 640GB HBM3
- 30.72 TB U.2 NVMe(8 \* 3.84TB)
- CX-7 & BlueField-3 DPU InfiniBand / RoCE组网
- 10.2KW供电, 6U高度

## DGX SuperPoD 集群: A100 → H100 → GB200

基于数十个~上百个 DGX Server互连而成

(900GB 双向带宽将 xxx 张 GPU 互连为超节点), 共 xEFlops@FP8  
《专为满足AI的独特需求而打造》

Eos 超级集群, 32000GPU, 80EFlops@FP16, 64MW



# ALPS-AI 超算新标杆，20E 级 AI 算力集群

- 2024 年全球超级计算机 500 榜单中，有将近 2/3 HPC 配备 NVIDIA GPU 或 Mellanox 系统。三年前，使用英伟达技术的超级计算机还不到 TOP 一半。



ARM CPU (Grace) + 新一代 GPU (Hopper)



**ANNOUNCING**  
THE WORLD'S FASTEST  
SUPERCOMPUTER FOR AI

20 Exaflops of AI

Powered by NVIDIA Grace CPU and  
Next Generation NVIDIA GPU

HPC and AI for Scientific and Commercial Apps  
Advance Weather, Climate, and Material Science



瑞士国家计算机中心 (CSCS)

# 总结与思考

# 计算产业发展的四个时代

1. 奠基时代1945~1970：单体算力集中，器件以电子管/晶体管为主，体积&耗电大
2. 通用架构时代1971~1995：大规模集成电路爆发，巨型机算力摸高，小/微机快速普及
3. 互联网云计算时代1995~2020：“数据中心”出现，应用催生更多算力服务形态和需求
4. 未来计算时代：2020~2030：计算面向集约&泛在&融合演进，引领K型算力需求爆发式发展



# 计算产业发展的四个时代

- 计算需求驱动，产业发展从专用计算到通用计算，再到集约泛在的未来计算
- Intel/AMD/NVIDIA 转型发力数据中心业务，加速构建全栈计算技术，市场竞争白热化





# Thank you

把AI系统带入每个开发者、每个家庭、  
每个组织，构建万物互联的智能世界

Bring AI System to every person, home and  
organization for a fully connected,  
intelligent world.

Copyright © 2024 XXX Technologies Co., Ltd.  
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. XXX may change the information at any time without notice.



**ZOMI**

GitHub <https://github.com/chenzomi12/Allinfra>

# 引用与参考

- PPT 开源在: <https://github.com/chenzomi12/AllInfra>