



LIGHTELLIGENCE

曦智科技





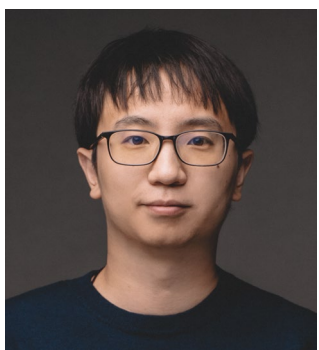
# 光网络赋能全局弹性计算系统

曦智科技创始人&CEO  
沈亦晨博士



“曦智科技光电混合的全新计算范式，致力于满足日益增长的大规模计算需求，为客户提供一系列算力跃迁解决方案，共建更智能、更可持续发展的未来。”

---



沈亦晨 博士

## 曦智科技创始人兼首席执行官

- 麻省理工科技学院 物理学博士学位
- Nature、Science等顶刊学术论文 40+
- 全球专利申请数150+, 已授权数30+
- 麻省理工科技评论 全球“35岁以下科技创新35人”
- 2022达沃斯“全球青年领袖”



LIGHTELLIGENCE  
曦智科技

## 全球光电混合计算领军企业

- 由沈亦晨博士创立于2017年
- 办公室分布于中国、美国、新加坡、比利时等
- 拥有来自全球10余个国家约200名全球员工
- 累计融资超2亿美金

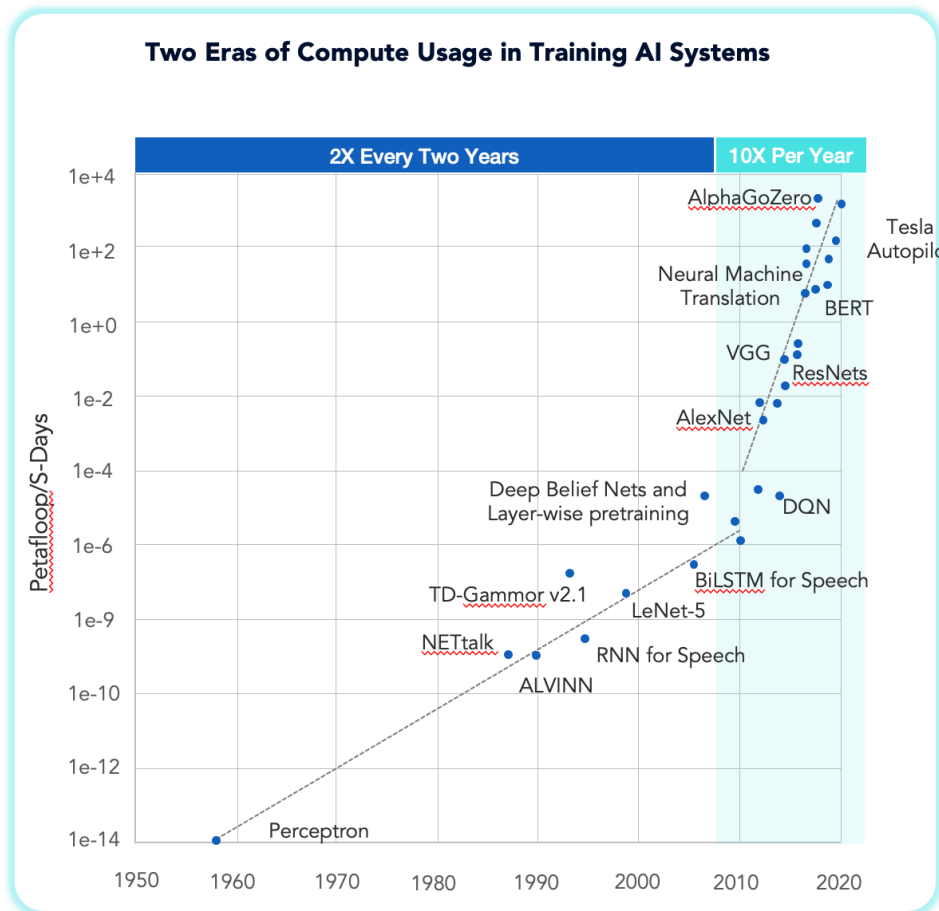


01

算力发展趋势和问题



# 计算需求指数型增长对可持续发展的挑战



AI模型训练时间的指数增长

2020年中国数据中心年耗电量：2045亿度（占社会总用电量2.7%）



功耗比例：  
计算~50%，存储~35%，通信~15%

Source - <http://dc.infosws.cn/20211210/53539.html>



# 未来计算技术的大趋势

- 伴随半导体技术发展60多年的摩尔定律可能在2020s年代结束 (Source: MIT Tech Review)

- 基于领域专用加速器 (DSA) 的异构计算成为解决算力瓶颈的必然趋势

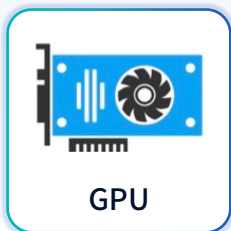
## 基于领域专用加速的异构计算黄金时代

Domain Specific Architecture (DSA)

(John Hennessy & David Patterson: Turing Award Lecture 2018)



DSP



GPU



FPGA



ASIC



通用性

计算效率

DSA 架构



# 当前计算模式/生态无法有效挖掘异构计算潜力



## 计算生态

计算架构多样性导致软硬件生态碎片化，软件适配和维护成本高



## 资源调度

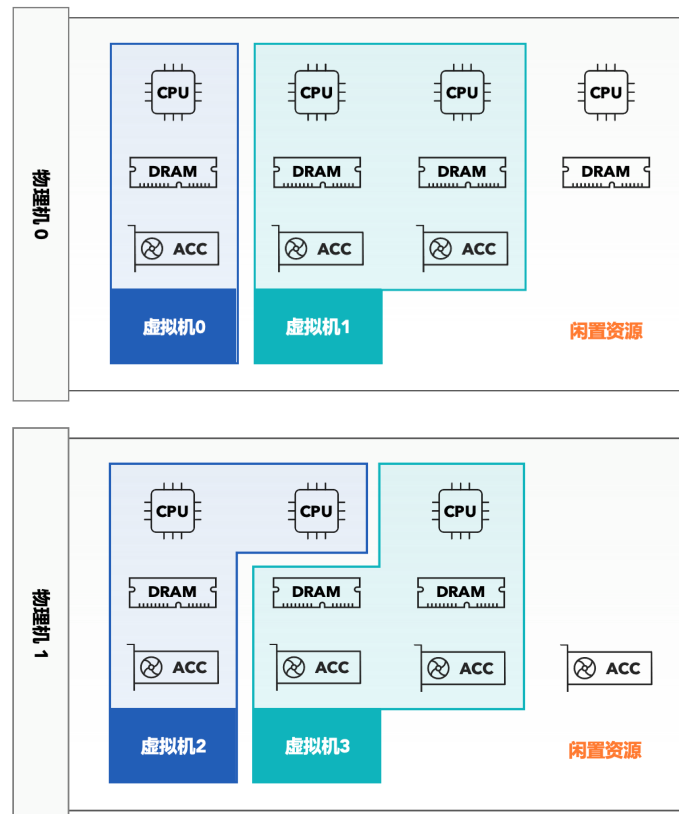
计算任务无法在不同架构上平滑迁移，使得带宽和延时阻塞算力的发挥



## 利用率

计算资源配比固化导致使用不均衡、利用率低，算力闲置浪费

## 传统数据中心

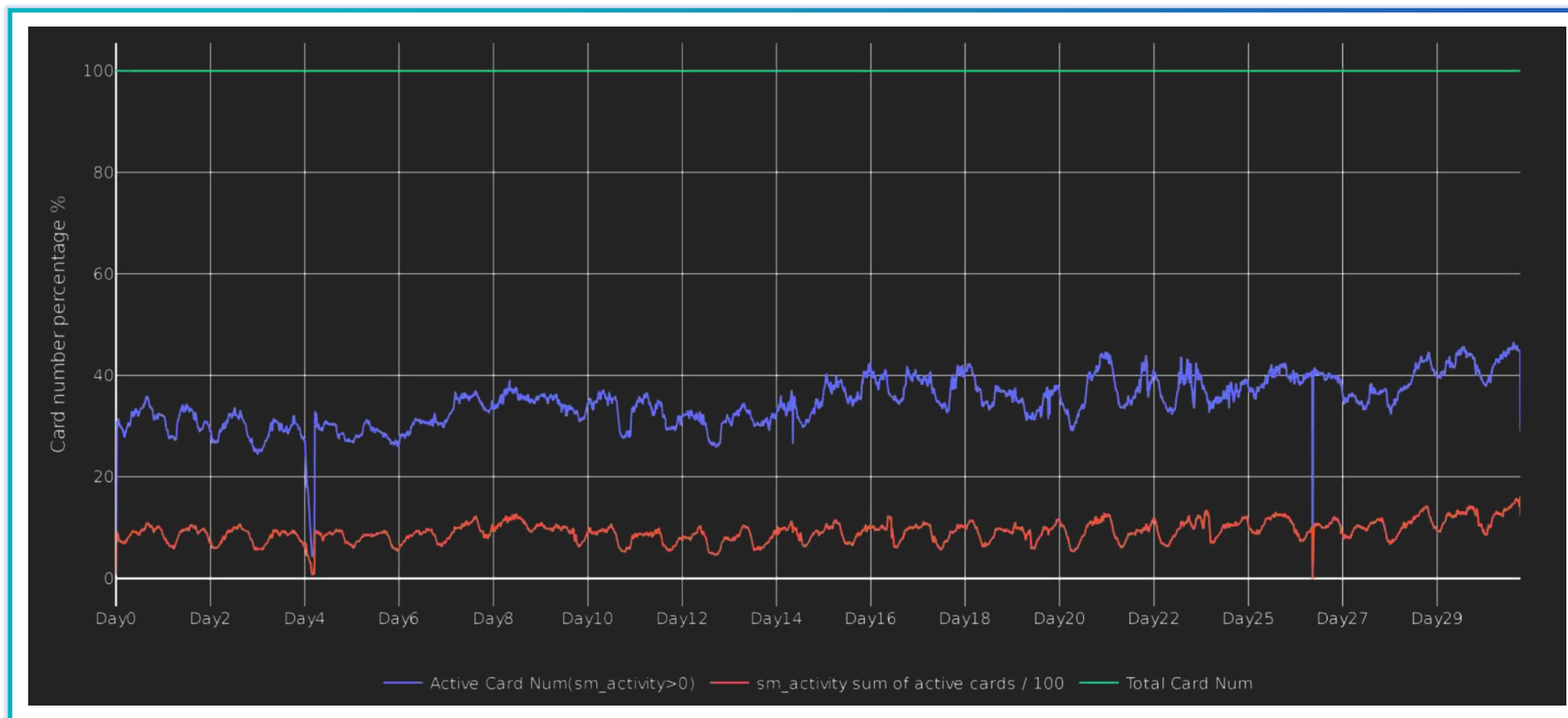




# 数据中心资源利用效率案例

## 阿里云数据中心GPU使用率

Source: O' Reilly Conference TensorFlow World 2019



不同客户的应用对计算需求量差别很大，导致目前数据中心利用率不高，资源浪费比较严重

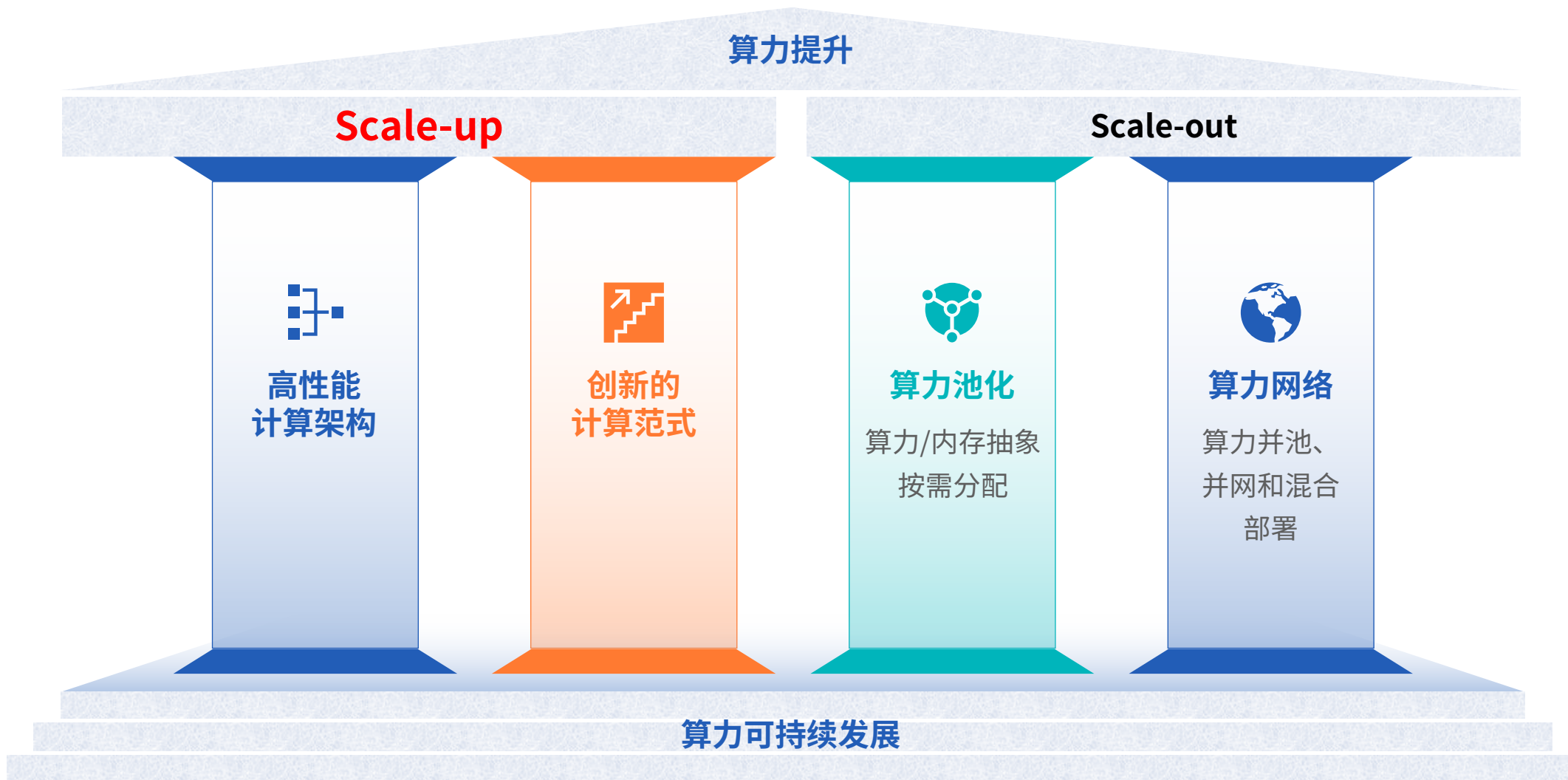


02

愿景和挑战



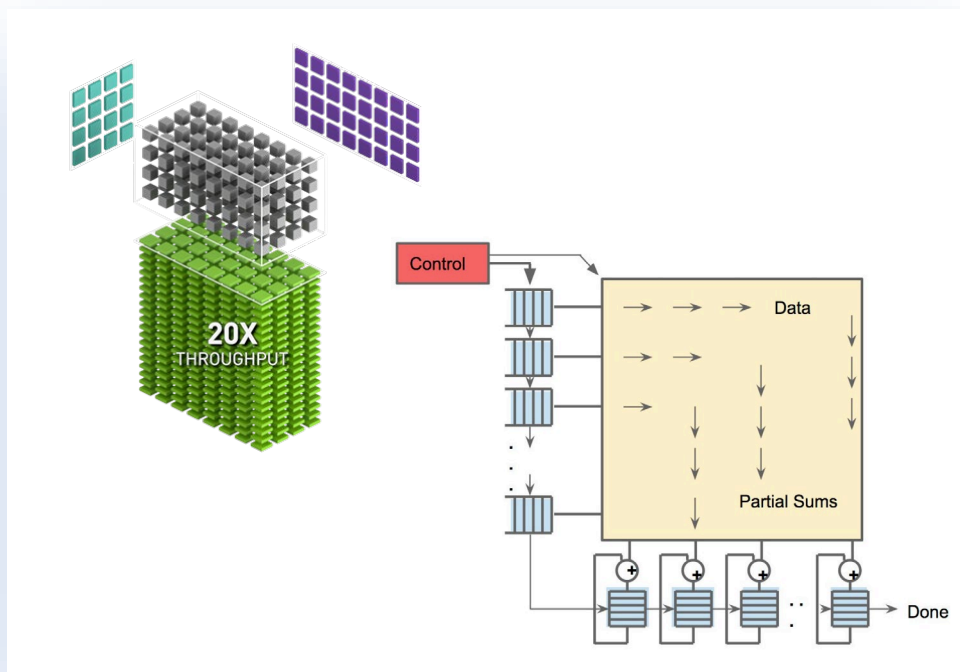
# 异构计算和资源池化是算力提升的必要条件





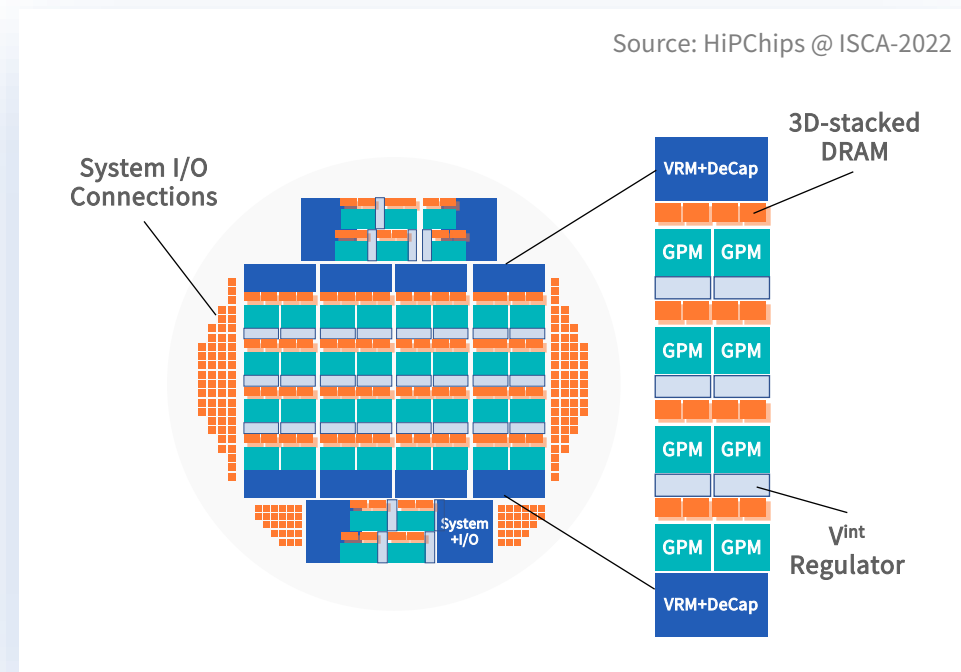
# 计算架构创新 – scale-up 范例

## 异构的张量计算加速单元



张量加速

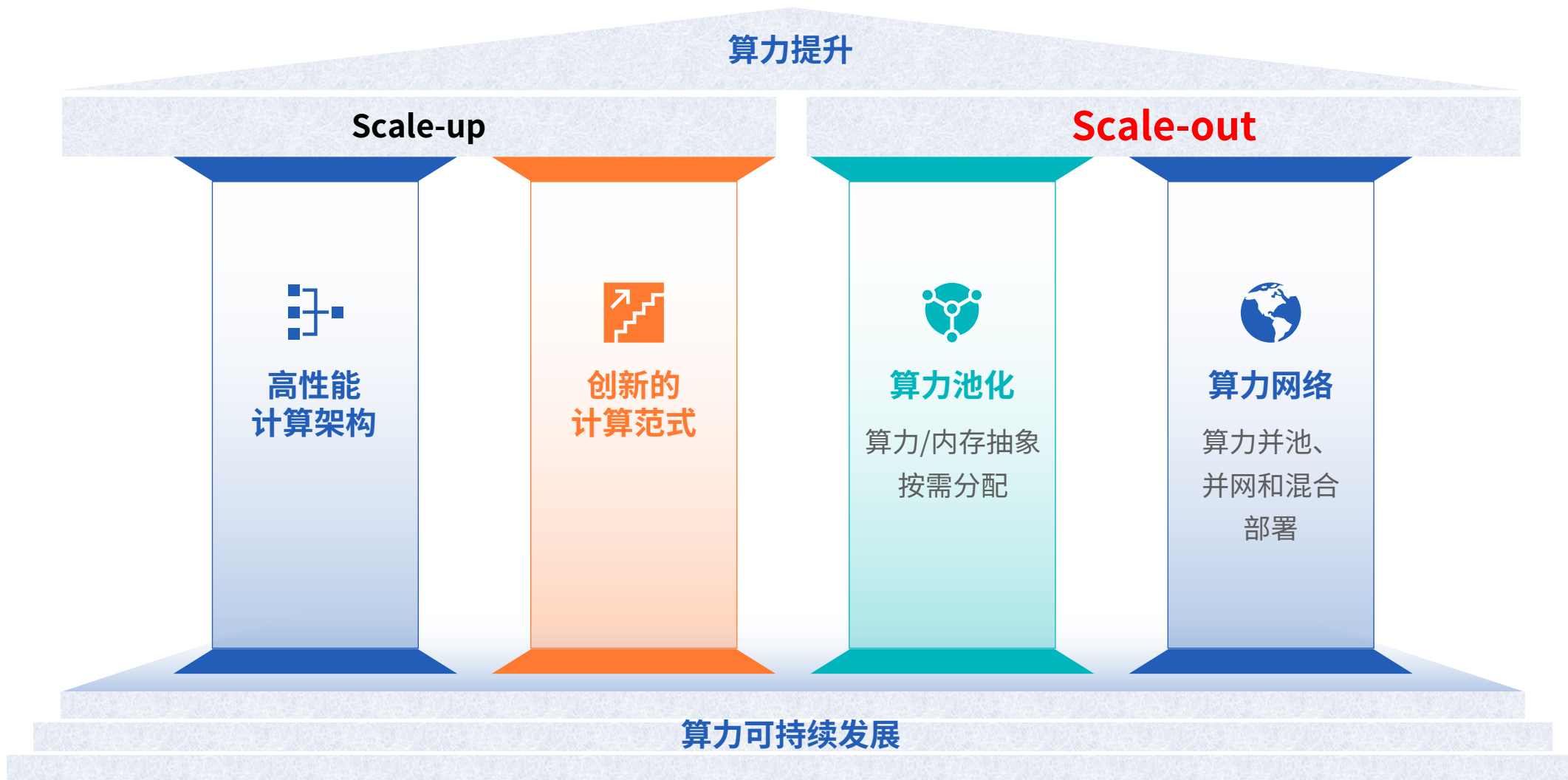
## 芯粒Chiplet 和片上网络



晶圆级GPU



# 异构计算和资源池化是算力提升的必要条件

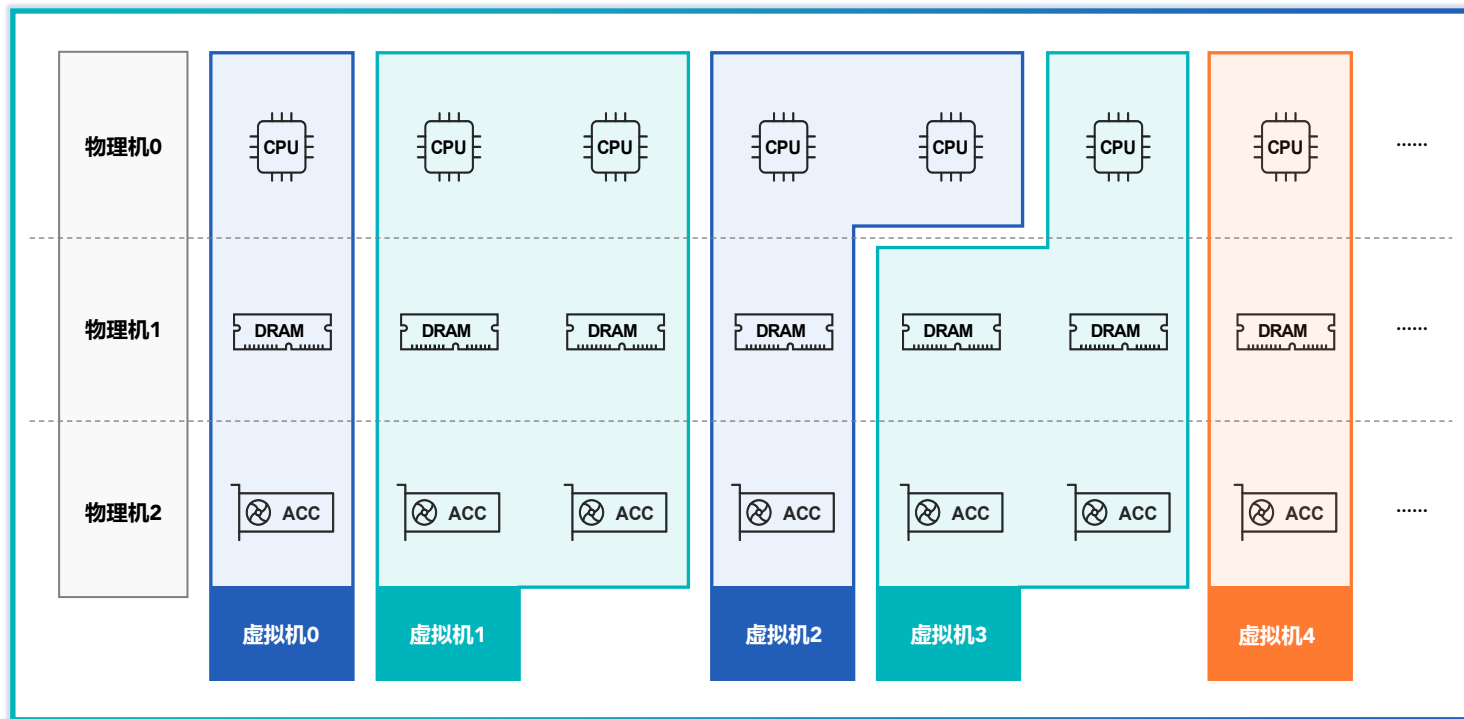






# 异构计算池化 – scale-out 范例

## 可重构解耦数据中心



柔性



可管理



经济



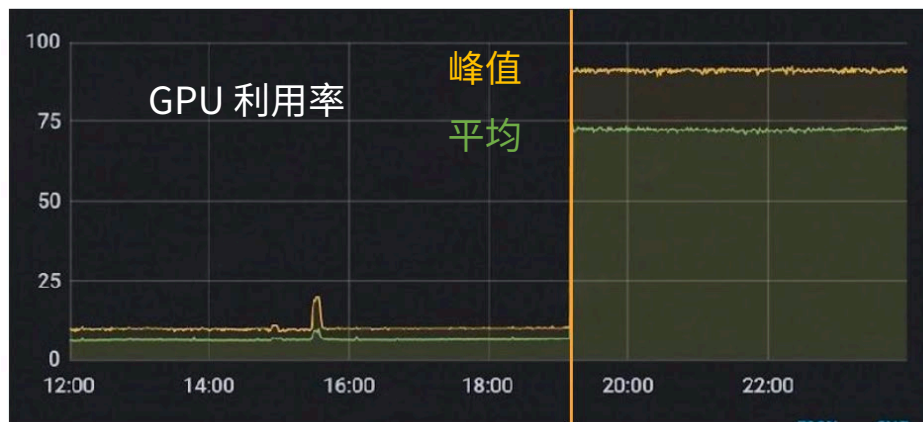
开放



# 异构计算池化 – scale-out 范例

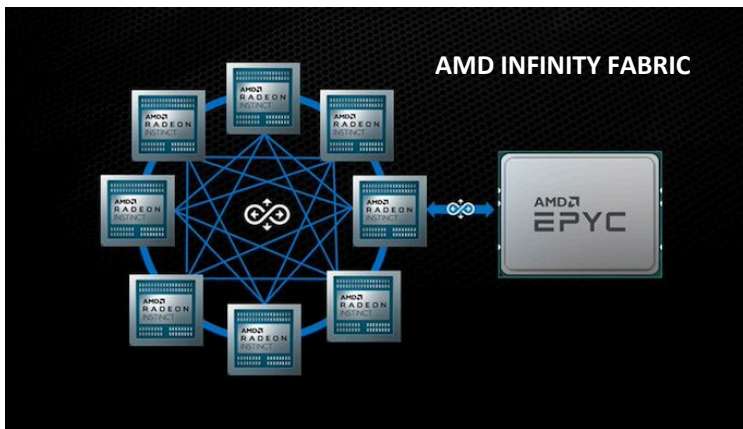
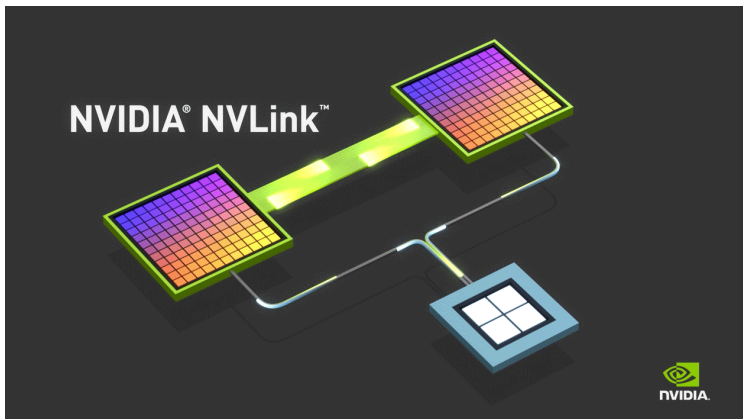
## GPU资源池化对利用率的影响

Source: Bytedance @ GTC 2022



Before sharing

After sharing

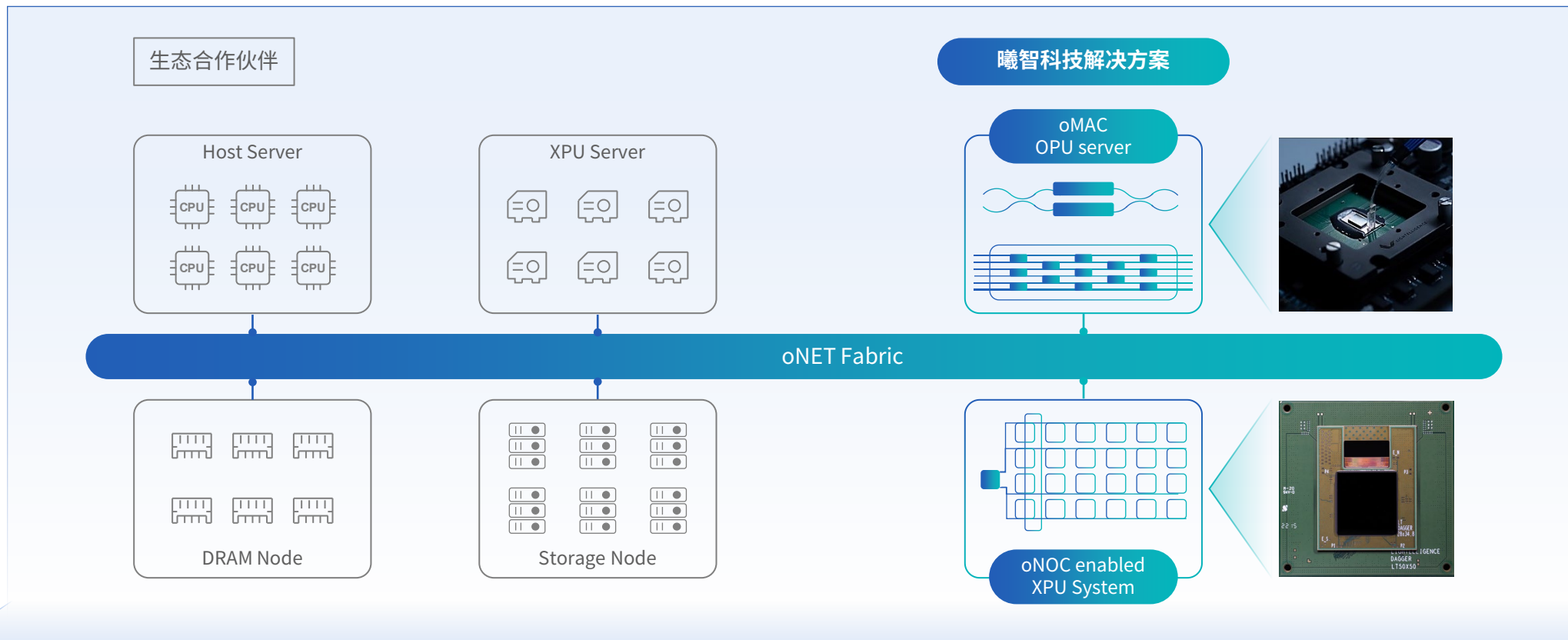


03

数据中心新范式与产品



# 曦智科技支撑未来计算趋势的关键技术



光子矩阵计算  
oMAC

片上光子网络  
oNOC

片间光子网络  
oNET

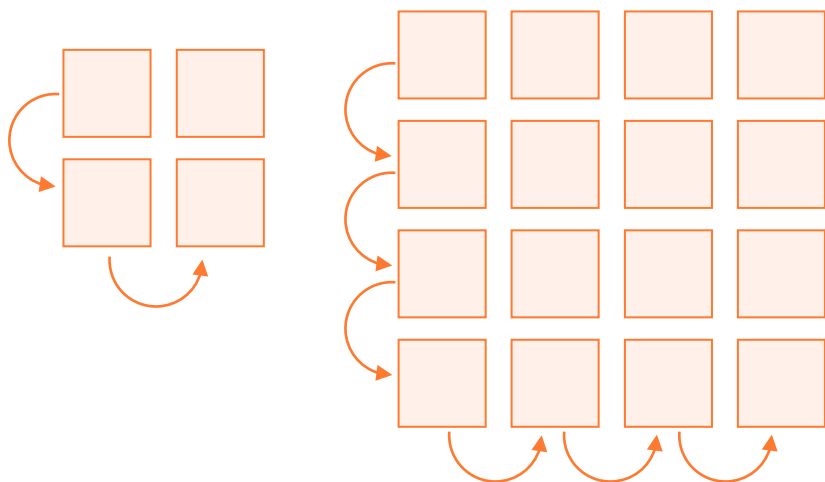
助力数据中心可持续发展的高能效光计算和光网络技术





# 变革性的光子矩阵计算 oMAC

## eMAC(电乘加器)延时

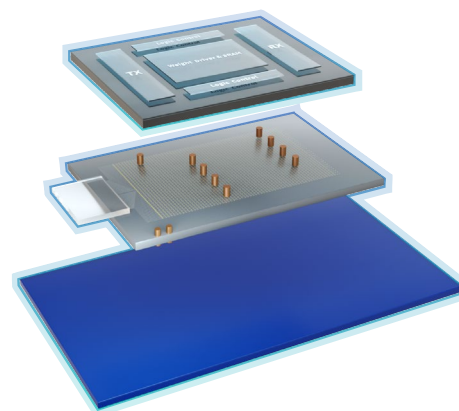
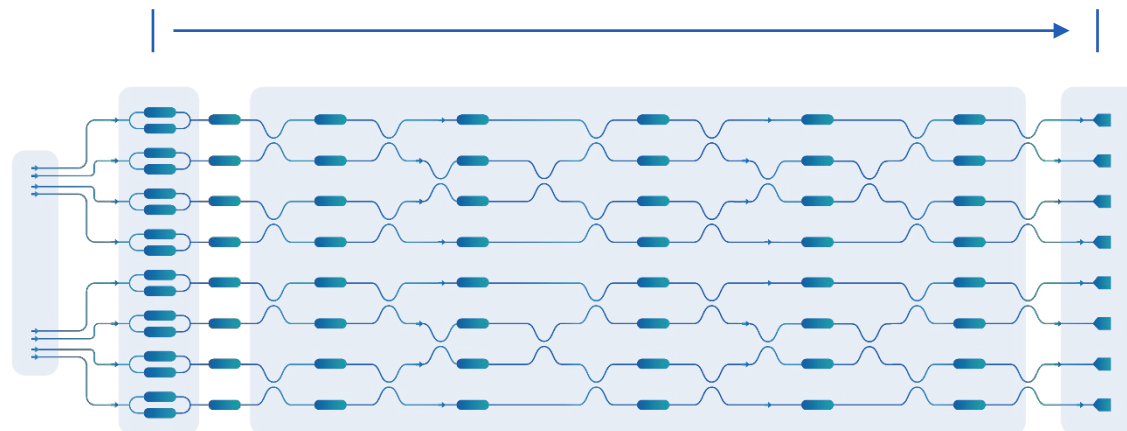


### 延时增长模型:

- 冯·诺伊曼:  $k*N^2$
- 脉动矩阵:  $k*N$
- ASIC:  $N+\log(N)$

随矩阵规模急剧增长

## oMAC(光乘加器)延时

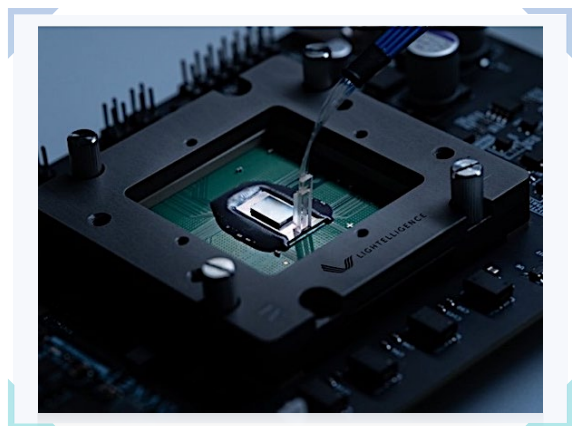
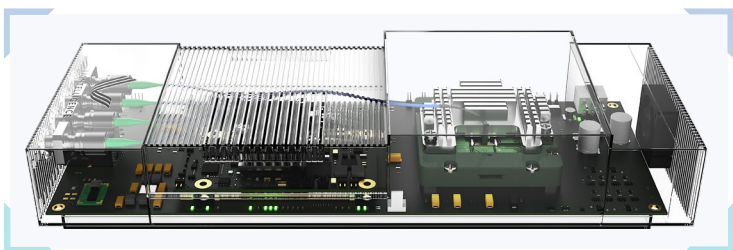


- 低延迟: 单个矩阵乘法延迟 $<5\text{ns}$
- 低能耗:  $>10\text{Tops/W}$

延迟对矩阵大小不敏感



## 世界上第一个64x64光子张量协处理器



(发布于2021年)

### PACE--Photonic Arithmetic Computing Engine 光子计算引擎

- 全球首款64\*64 oMAC计算芯片
- 单个光子芯片中集成超过10,000个集成光子元器件
- 运行单一计算问题速度可达目前高端GPU数百倍
- 利用光计算超低延迟特性

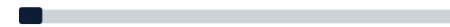
伊辛问题单次循环计算时间

**3**<sub>ns</sub>    **2,560**<sub>ns</sub>  
PACE    A10GPU



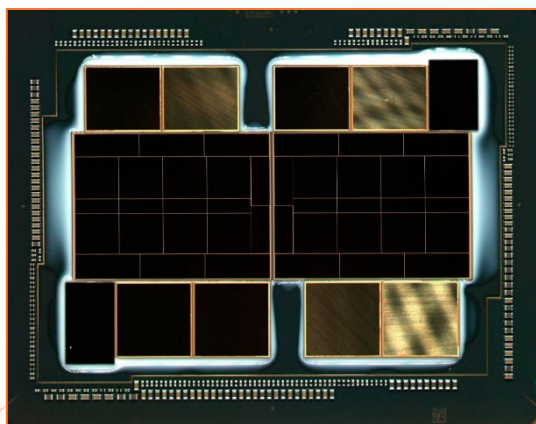
伊辛问题求解时间

**0.015**<sub>ms</sub>    **12.8**<sub>ms</sub>  
PACE    A10GPU

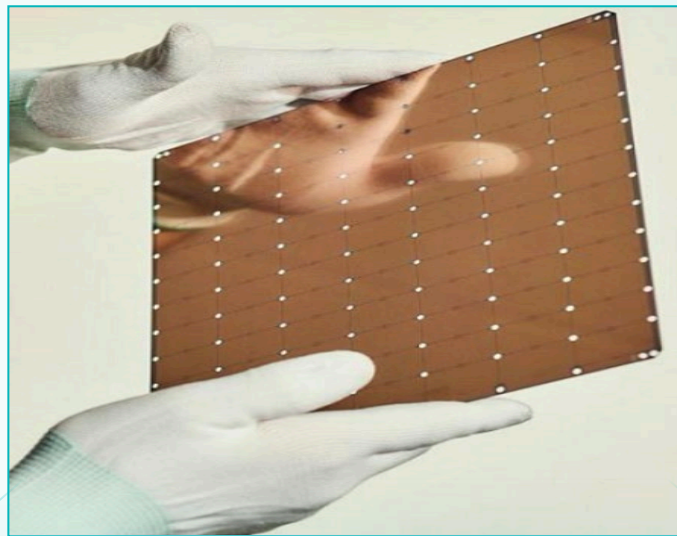




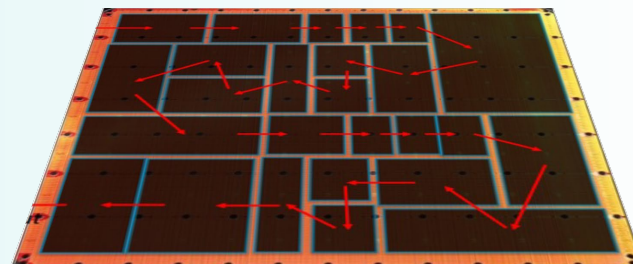
# 大芯片的片上电互连边际成本急剧升高



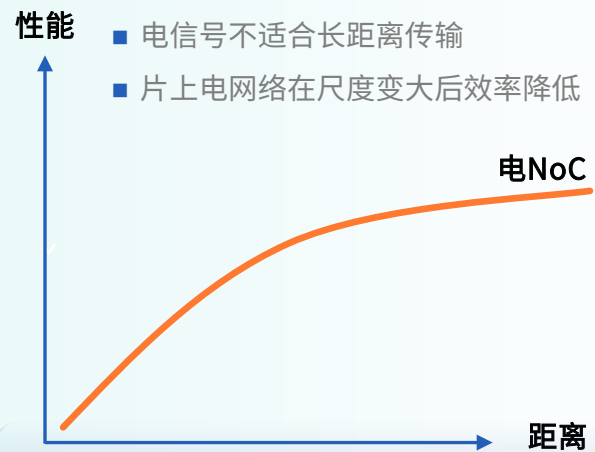
基于Chiplet的MCM GPU (Intel)  
>1,200 mm<sup>2</sup>



晶圆级计算引擎(Cerebras)  
>46,000 mm<sup>2</sup>



- 电信号不适合长距离传输
- 片上电网络在尺度变大后效率降低

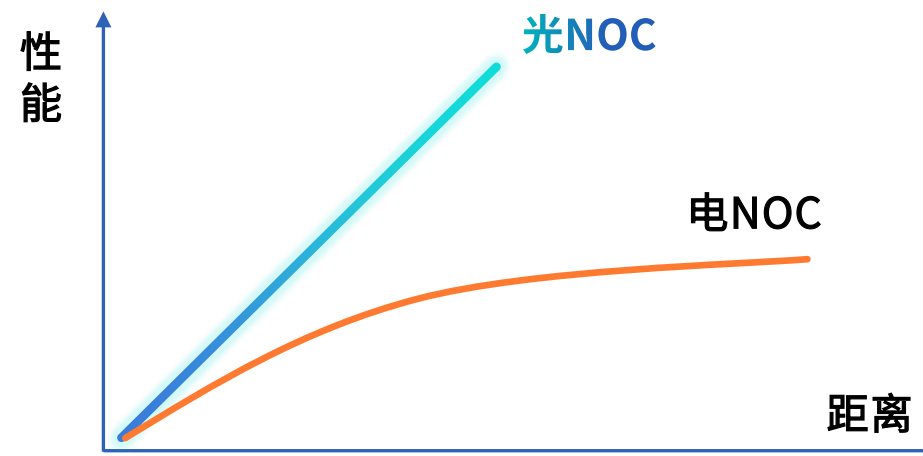
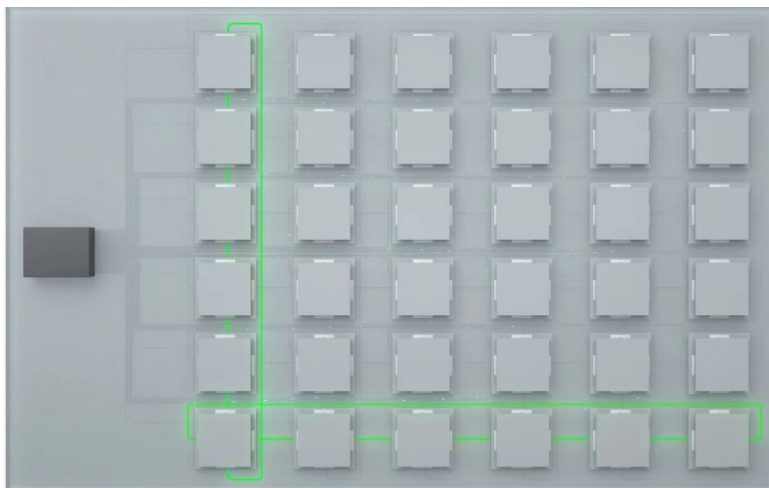
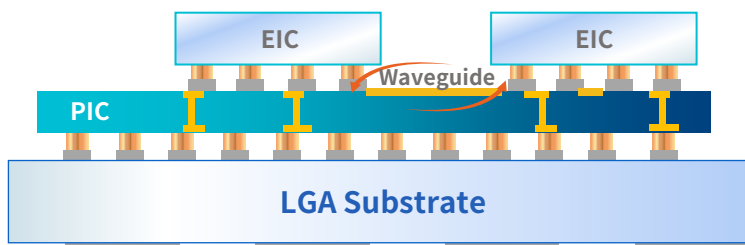


单个封装单元或板卡内的  
算力提升需要更好的互联解决方案



# 变革性的片上光子网络 (oNOC) 实现高效的计算/存储互联

一个封装单元内



高带宽、低能耗、低延迟、  
距离不敏感



通用性强，可将不同类型的  
电子芯片与之结合，为芯片间  
提供高速、低能耗的互连

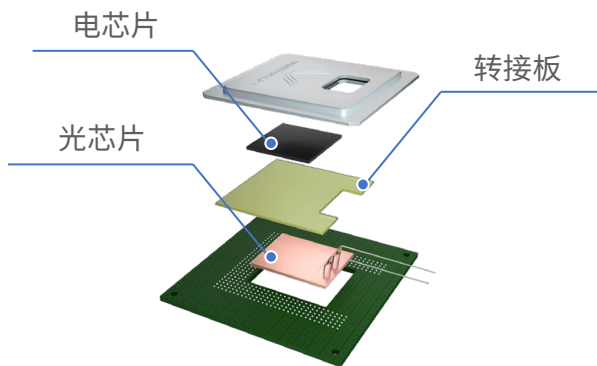


有利于任何有高带宽  
需求的应用场景



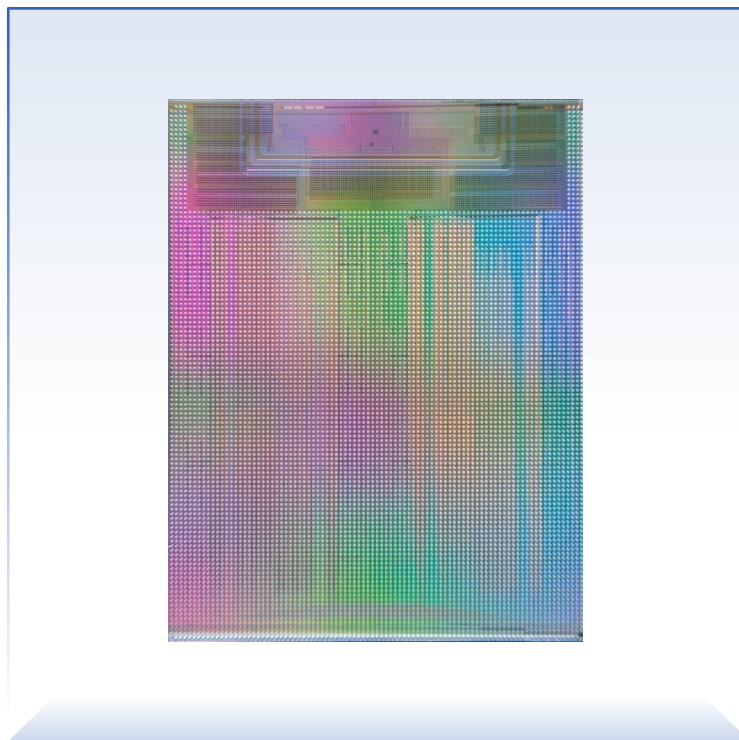


# 首个基于片上光网络(oNOC)的计算芯片

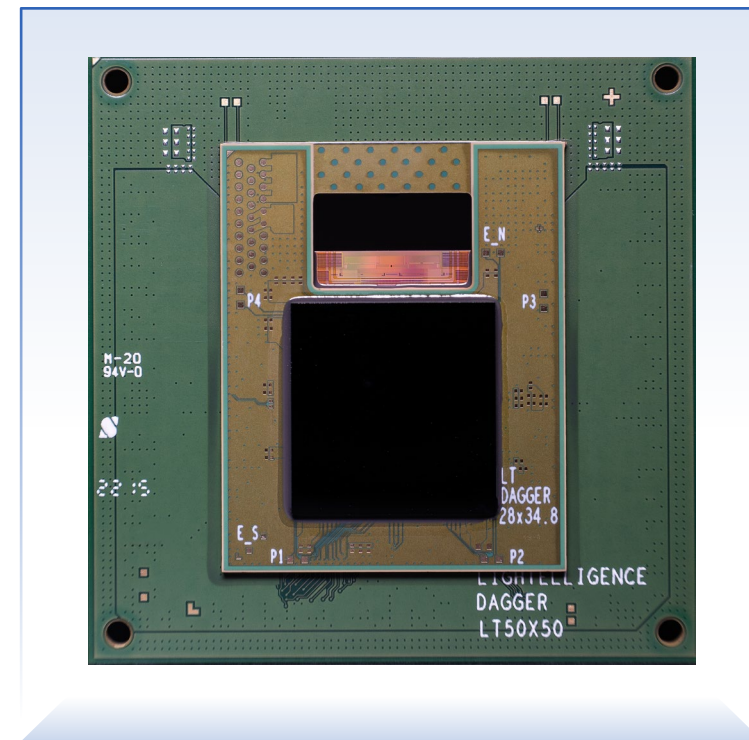


广播延时	1ns
最长广播距离	50mm
通道数	512ch
单通道频率	4GHz
片上总带宽	2Tbps

\*首个芯片实测数据，未来提升空间巨大



光芯片实际图片

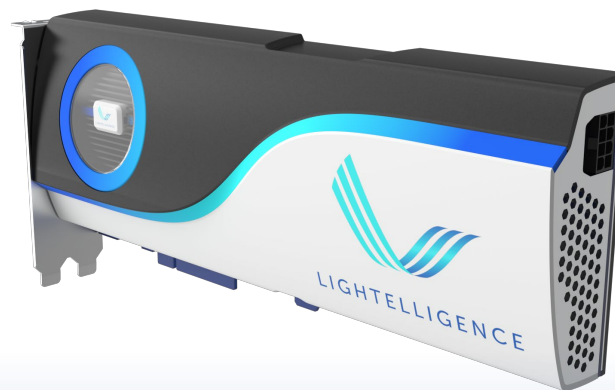
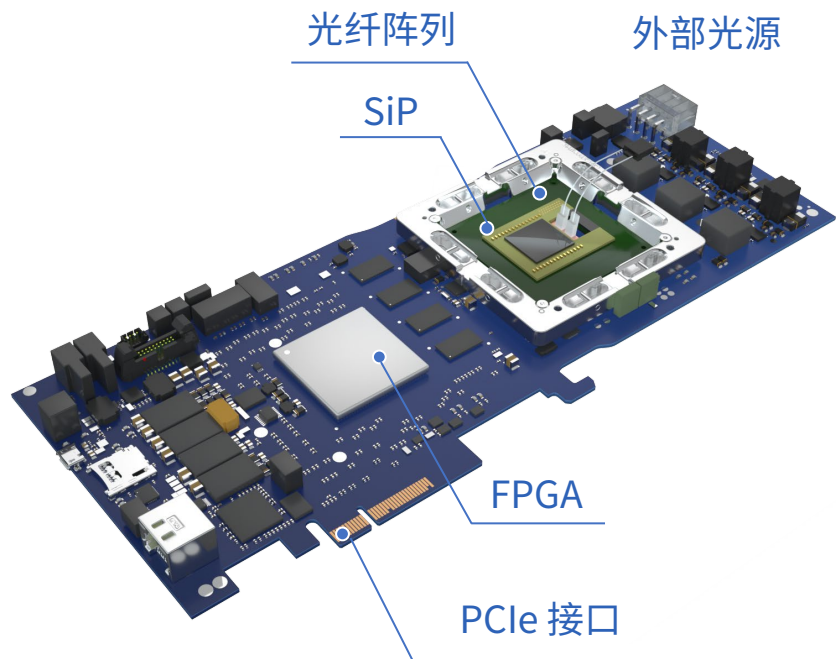


封装实际图片



# 光子计算产品线

世界上第一个基于片上光网络的计算产品，据此证明算力扩展的下一代架构



**可执行AI模型的推理任务及其他计算加速**

光NoC All-to-All 广播

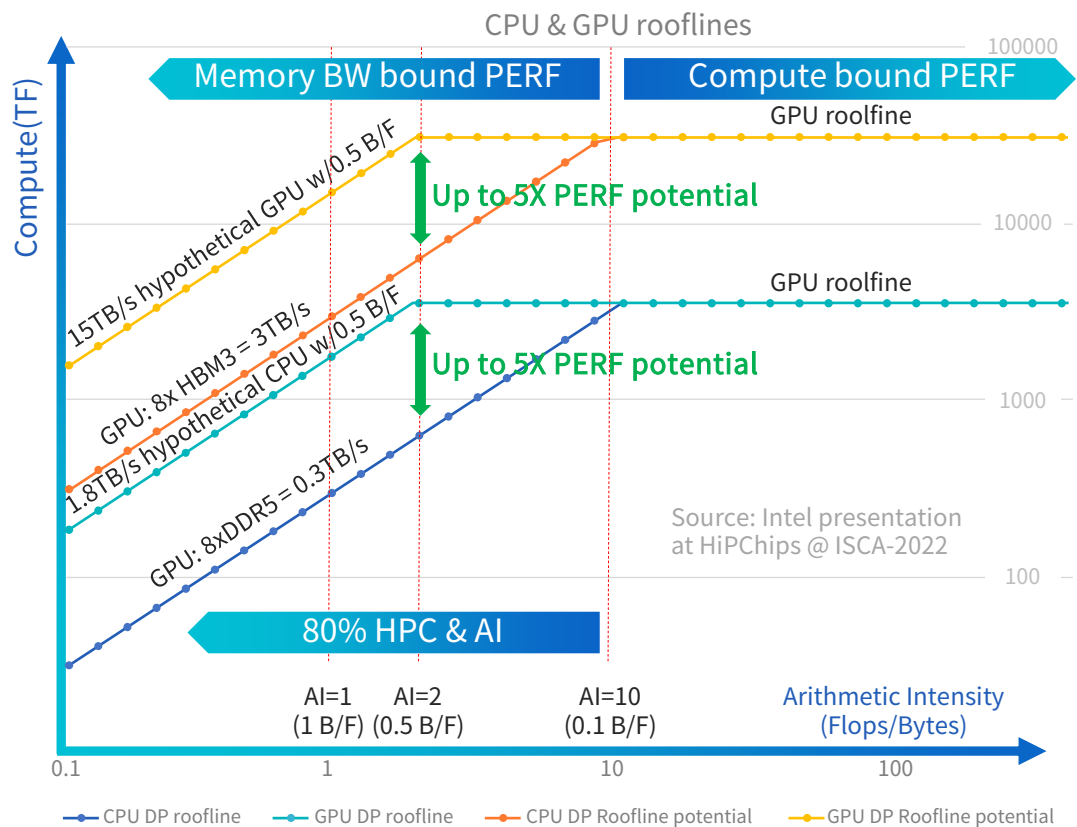
64 个计算核心

Lightelligence SDK

PCIe 标准尺寸



# 分布式计算需要高效互连基础设施



## 资源池化和算力网络需要高效的互连基础设施:



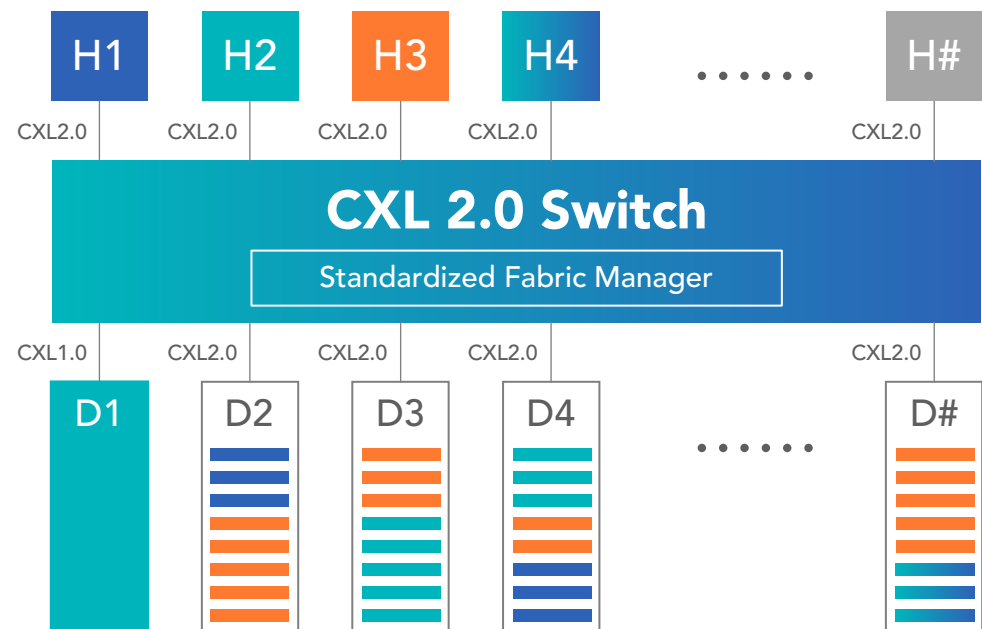
≥ 80% HPC和AI应用都有较低的计算密度，而且CPU/GPU计算架构展示了类似趋势



同一个计算架构下，内存带宽增加(0.5B/F, AI=2)对性能提升效果明显



# CXL成为硬件资源解耦架构的主导协议



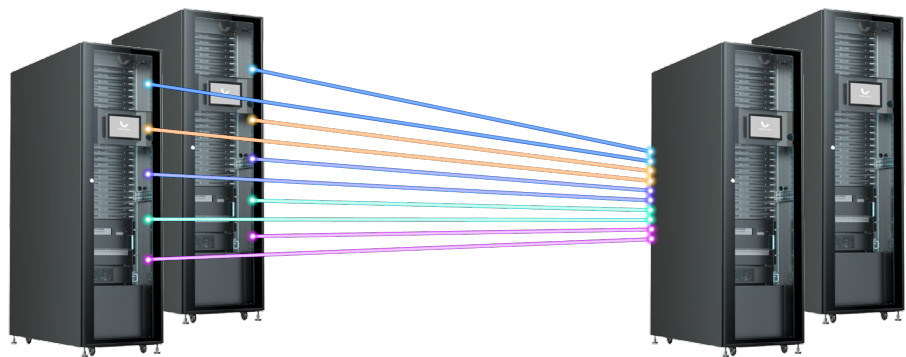
	缓存一致性	延时	内存解耦
CXL	是	~100ns	支持
RDMA (ethernet)	否	~3μs	不支持



获得产业头部企业广泛支持

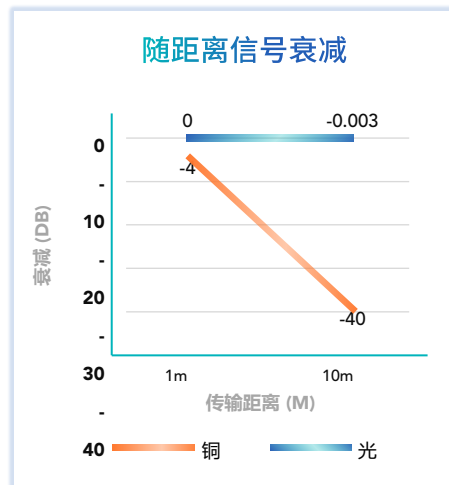


# 光的物理优越性提供了更大范围的高效scale-out

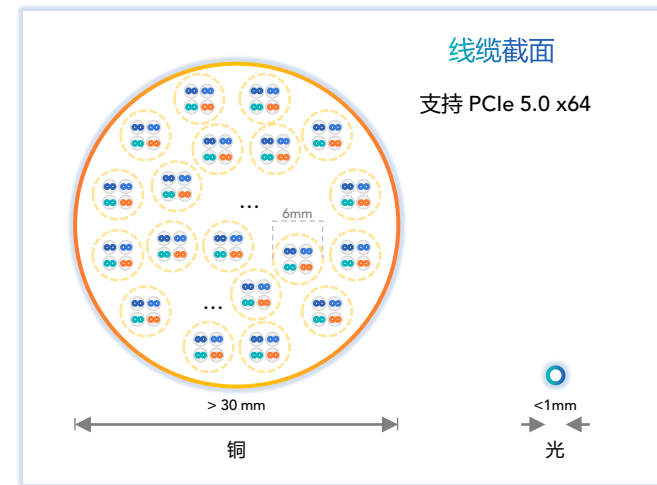


## 机架之间互连

低延时、长距离、低功耗



假设 AWG26 线缆, PCIe 5.0 信号

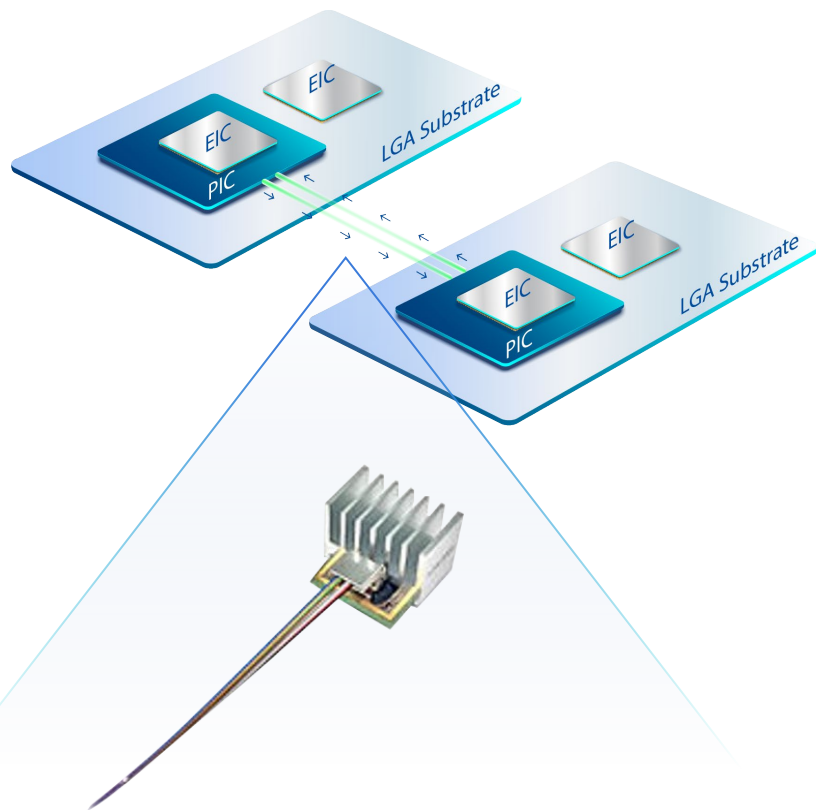


32 根直径 > 6mm (CAT8) 线缆 16 根直径 = 0.125mm 光纤

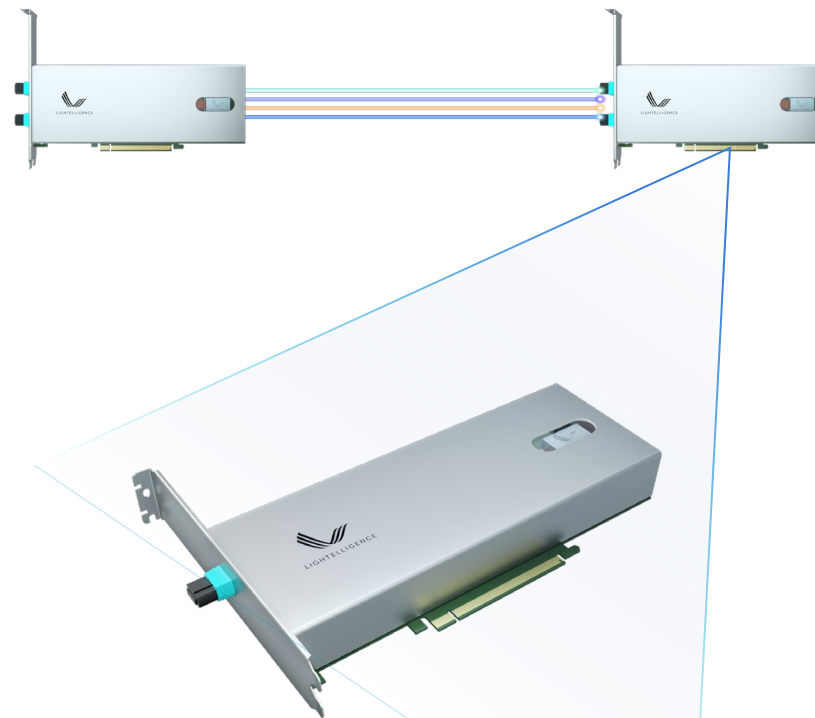
铜线缆勉强支持相邻服务器间CXL扩展



# 硅光CXL可能的产品形态



OBO (板载光学)



PCIe板卡



One More Thing



# 系列白皮书重磅发布

《大规模光电集成赋能全局弹性计算系统白皮书》

01

《光子矩阵计算 (oMAC) 技术白皮书》

02

《片上光网络 (oNOC) 技术白皮书》

03

《片间光网络 (oNET) 技术白皮书》

04



谢谢观看

Thank you for your time.