

AI 集群 网络基本知识

Content github.com/Infrasys-AI/AIInfra

AI 系统 + 大模型全栈架构图

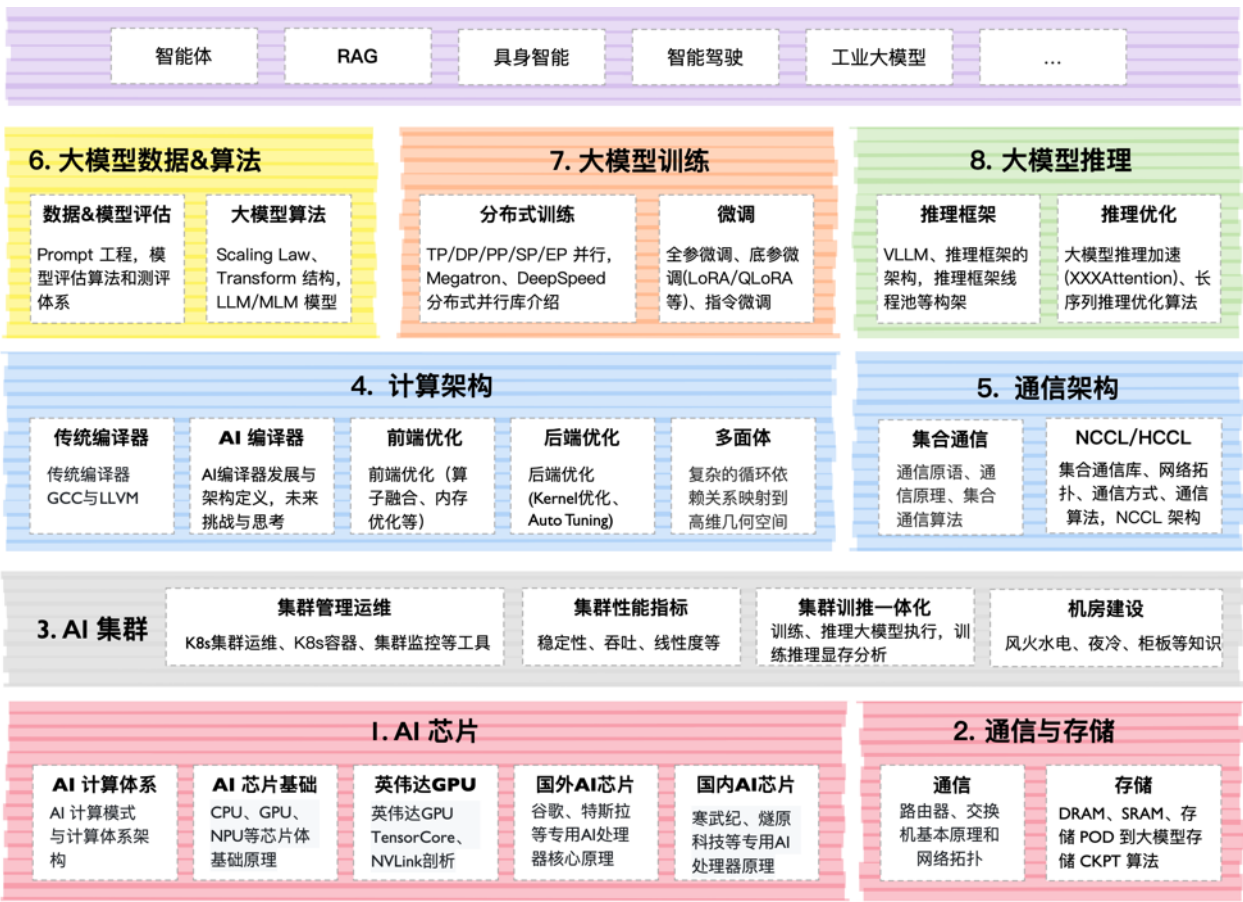


时事
热点

大模型
训推

编译
计算
架构

硬件
体系
结构

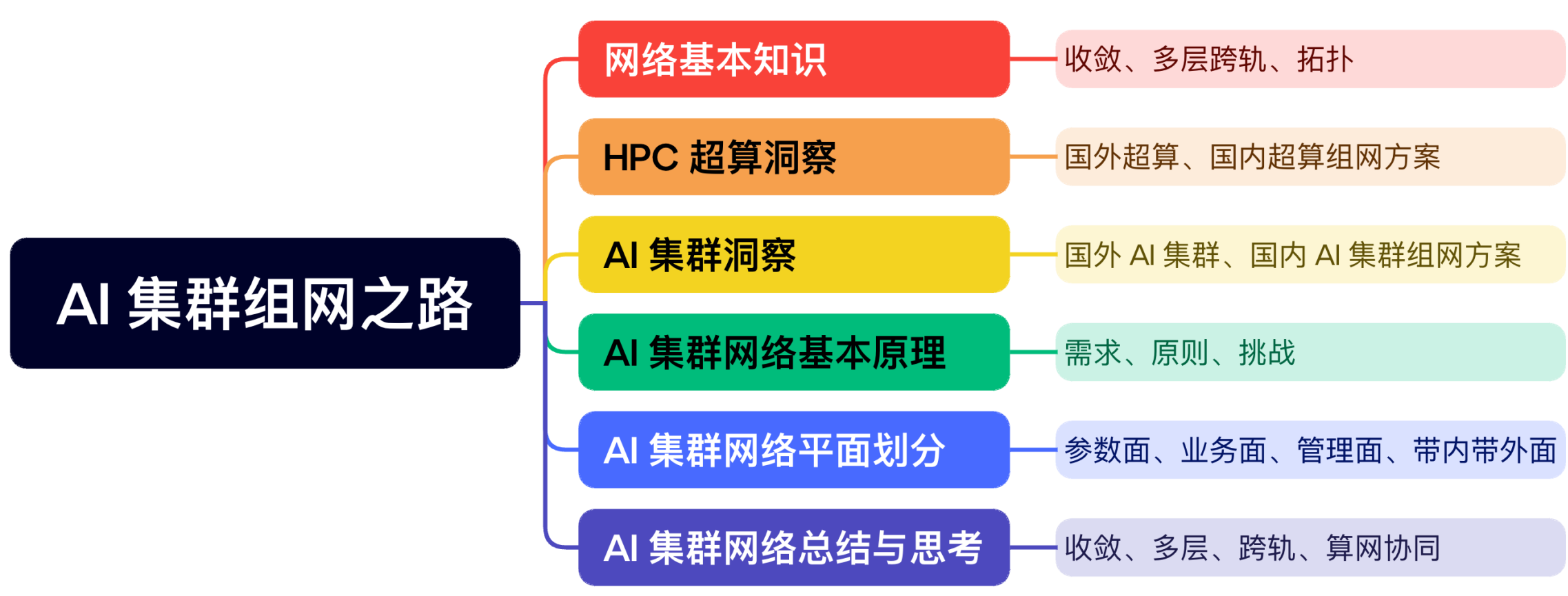


Question?

- 万卡集群，感觉很难，但不就是把 GPU/NPU 都放在一起堆料吗？
- 但你想过吗？多层和单层组网、带宽收敛比、网络拓扑、多平面网络、交换槽位和端口、铜互联和光互联、单轨和多轨通信、框和盒关系。



Content



找不到，根本找不到呀家人们

Google

GPU AI 集群 单轨通信

×

🔍

🗨️

📷

🔍

All

News

Images

Short videos

Web

Finance

Books

More ▾

Tools ▾

Saved

锐捷

Gpu 服务器

Nvidia

算力

Ai fabric

Nvlink

万卡

Cpu

网络


Ai infra

组网

网络架构

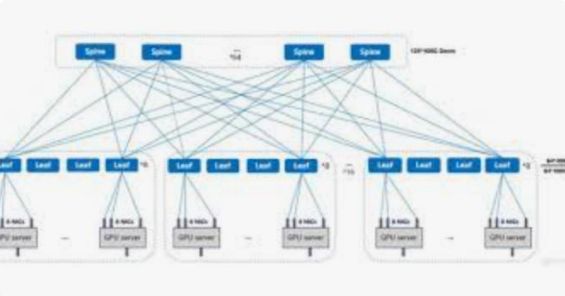
组网方案演进：参数面多轨下行降速，样本面与业务面合一，影响IO读写效率

- 参数面多轨下行：以参面100G RoCE接入，单轨高带宽
- 无独立数据面，数据面和业务面合一后256k NPS协议访问，访问效率存在风险



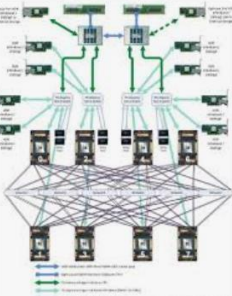
InfoQ

大模型在超大规模集群性能提升实践_AI&...



51CTO博客

AI算力网络单轨和多轨_wx67aeaa48dbd3c的技术...





知乎专栏

AI时代的GPU集群...

小红书案例：训练性能提拉

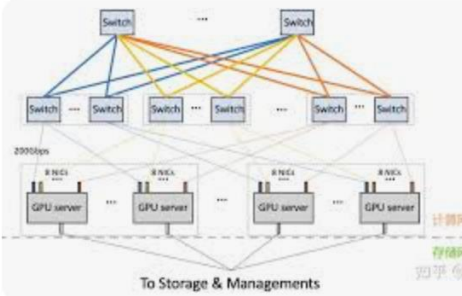
三个方面分析：

- 1. 网络架构：分析网络架构，分析网络带宽，分析网络延迟，分析网络抖动
- 2. 算力资源：分析算力资源，分析算力利用率，分析算力瓶颈
- 3. 数据分布：分析数据分布，分析数据倾斜，分析数据冗余



InfoQ

大模型在超大规模集群性能提升实践_AI...



CSDN博客

AI时代的GPU集群网络算力分析_算力规...



Scale-up vs Scale-out

场景	Scale-up（纵向扩展）	Scale-out（横向扩展）
核心目标	单个“超级GPU”能力超强	千 GPU 实现高效协同
网络重点	NVLink 带宽（喂饱单个猛兽）	低延迟 + 高带宽（万卡 GPU 齐心协作）
致命瓶颈	链路带宽不足 → 算力闲置	网络延迟高 → 集体“等数据”挂机
延迟敏感度	纳秒级（1根头发掉落时长）	微秒级（蜜蜂振翅一次）
数据传输距离	机箱内（铜缆直连）	跨机柜（必须光纤）
拓扑依赖	全连接拓扑（不惜代价直连）	胖树拓扑（成本效率平衡）



目录

1. 网络架构基础：城市交通系统
2. 网络效率关键：带宽与流量管理
3. 网络硬件基础：交通基础设施
4. 通信机制：数据传输方式，交规与道路标准



01

网络架构基础

城市交通系统



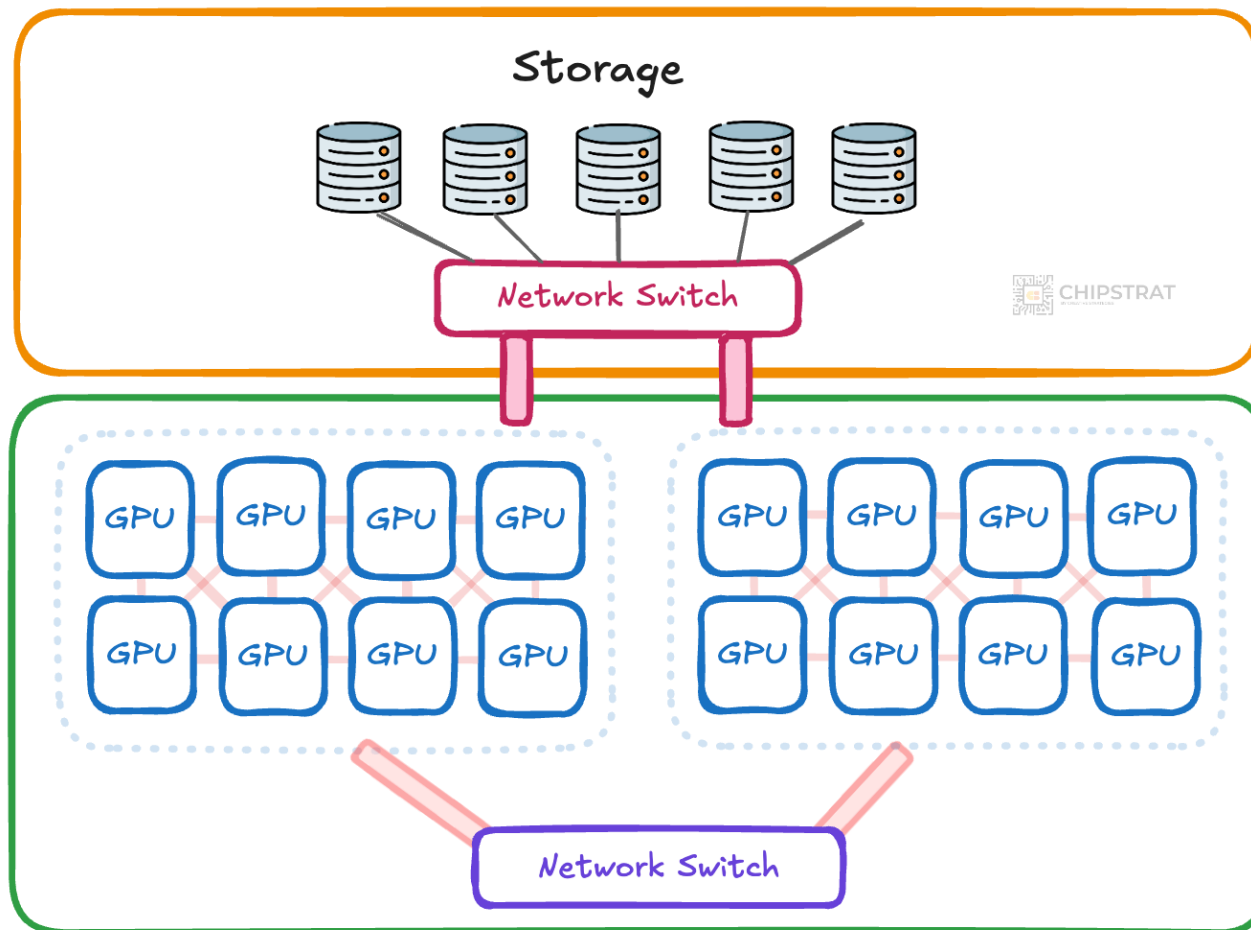
网络层级结构：单层 vs 多层组网结构

- **单层组网**：像地铁系统，所有站点直接相连（如 NVLink 全互联），适合 Scale Up（单节点内多 GPU 通信）
- **多层组网**：像地铁+公交+高架的立体交通，分核心层/汇聚层/接入层，适合 Scale Out（多服务器集群）
- Leaf-Spine 三层胖树架构：
 - Leaf 层交换机（柜顶）：直连 GPU 服务器
 - Spine 层交换机（核心）：高速互联 Leaf 交换机
- 无阻塞规则：核心层带宽 \geq 所有 Leaf 层带宽总和

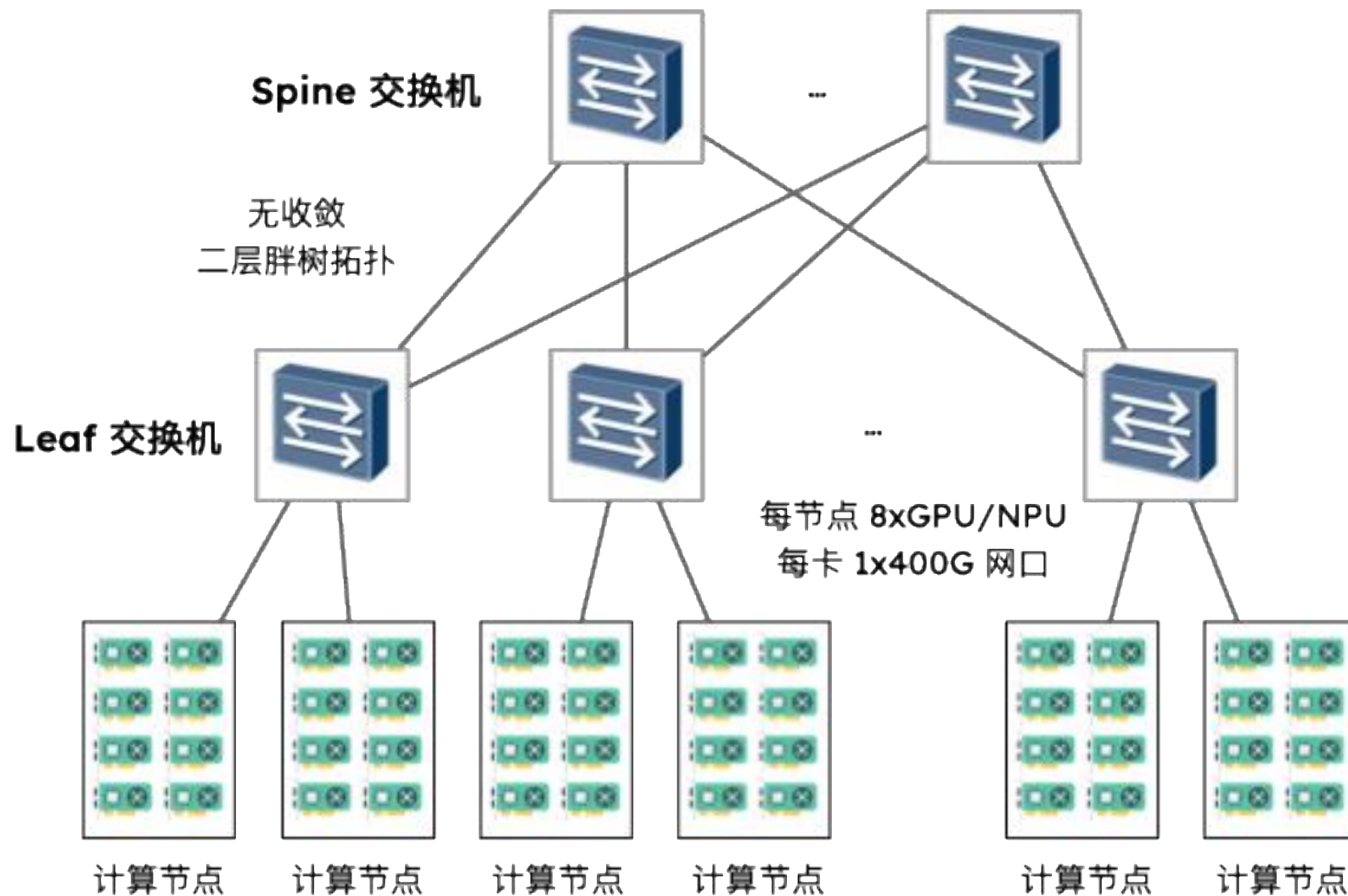


网络层级结构：单层 vs 多层组网结构

FRONT-END NETWORK



组网：优先选400G的场景 vs 优先选200G的场景



02

网络效率关键 带宽与流量管理



带宽收敛比：无收敛 vs N:1 收敛

- 网络“入口宽度”与“出口宽度”的比例，收敛比 = Σ 下行端口带宽 / Σ 上行端口带宽
- **1:1无收敛**：10 条入口车道对应 10 条出口车道（理想状态）
- **3:1收敛**：3 条入口车道挤进 1 条出口车道（必然堵车）
- 90% AI 训练集群卡顿源于网络收敛比过高，就像 100 辆车挤进 30 个收费口
- AI 训练中，GPU 需要频繁交换海量参数
- 高收敛比导致“数据堵车”，GPU 大部分时间在等数据，算力浪费
- 现代 AI 集群正从传统的 3:1 收敛向 1:1 无收敛演进，400G 高性能网卡成为主流



带宽收敛比：无收敛 vs N:1 收敛

- 网络“入口宽度”与“出口宽度”的比例，收敛比 = Σ 下行端口带宽 / Σ 上行端口带宽

256台GPU \rightleftharpoons 柜顶交换机 (40Gx256)



核心交换机集群 (10.24Tbps总带宽)

256台GPU \rightarrow 40Gx256 需求带宽



核心交换机仅2.56Tbps \leftarrow 瓶颈淤血!



网络拓扑：决定数据传输路径

- 快递分拣中心布局，Scale Out 集群必须用高级拓扑架构，否则万卡 AI 集群服务器间通信会像没有导航的快递系统
- Scale Up：使用简单的全互联拓扑 Full mesh
- Scale Out：必须用高级拓扑架构

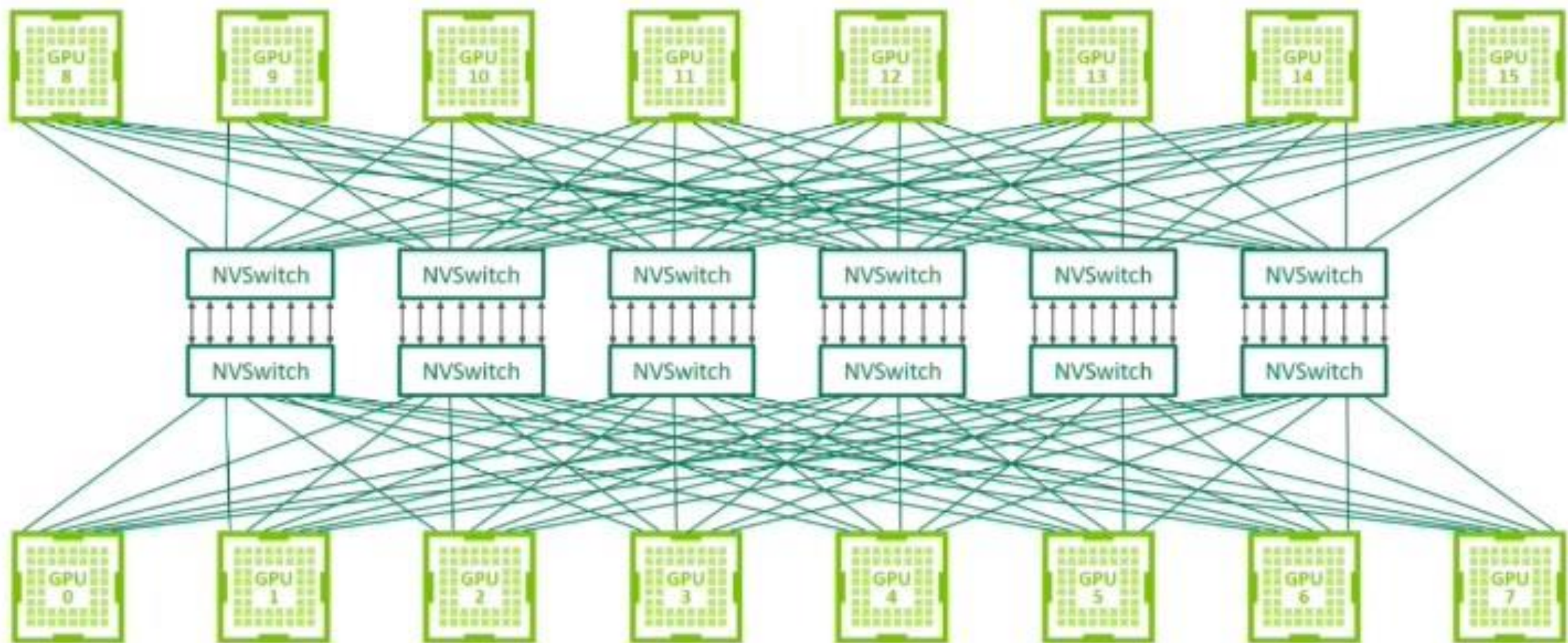


网络拓扑：决定数据传输路径

- Scale Up：使用简单的全互联拓扑 Full mesh
 - 优势：任意两点直连（延迟最低）
 - 致命伤：GPU 数量 (N)，连线数 = $N(N-1)/2$ ，如 8 GPU 需 28 条线路



网络拓扑：决定数据传输路径

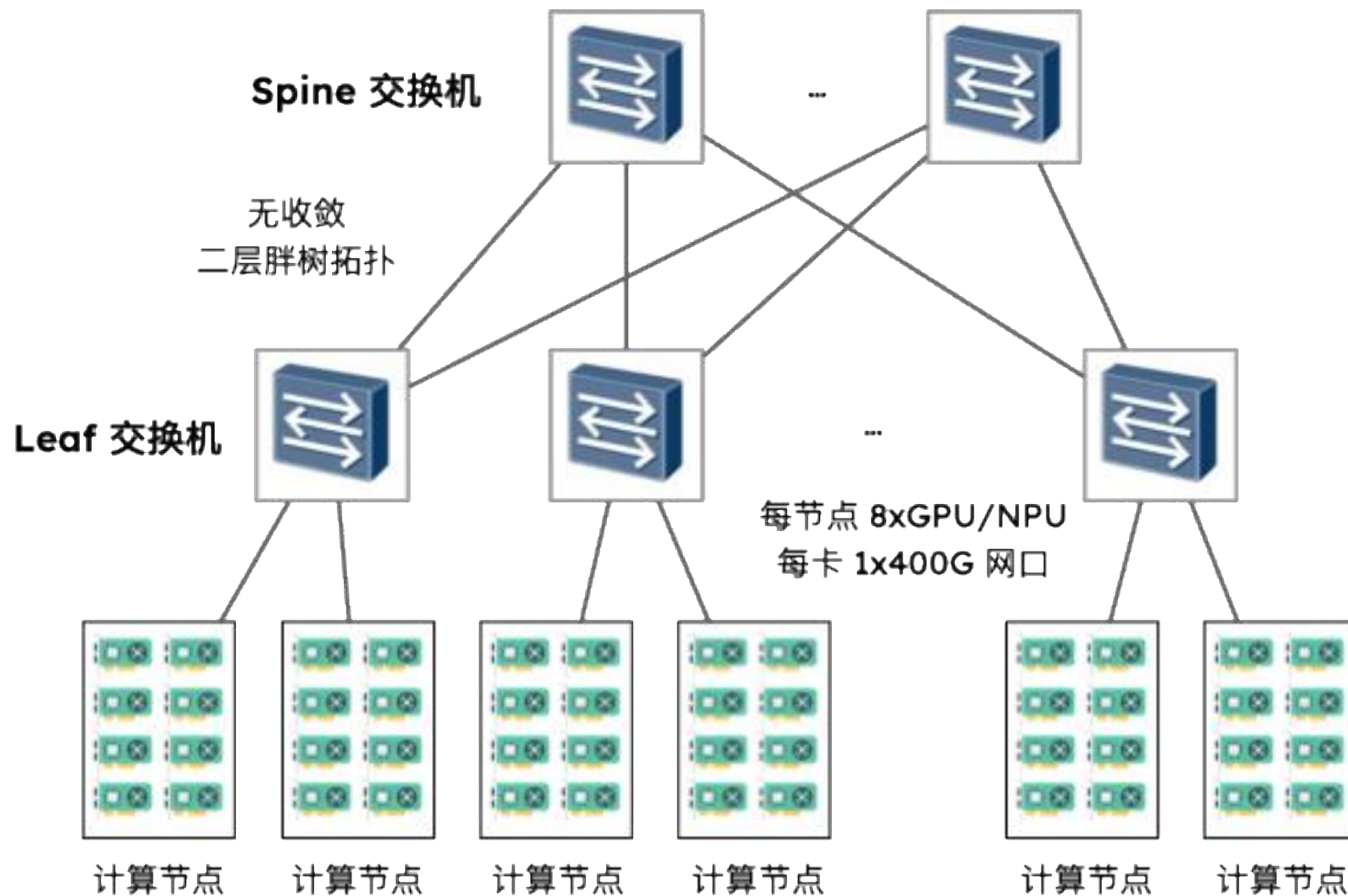


网络拓扑：决定数据传输路径

- Scale Out：高级拓扑架构
 - Fat-Tree：像立体停车场，任意两点间有多个路径
 - Dragonfly：像机场中转系统，先同城快运再跨省转运
 - Clos：现代AI集群的主流选择，可无限扩展
- 路径冗余：任意两点存在多条通路



组网：优先选400G的场景 vs 优先选200G的场景



03

网络硬件基础 交通基础设施

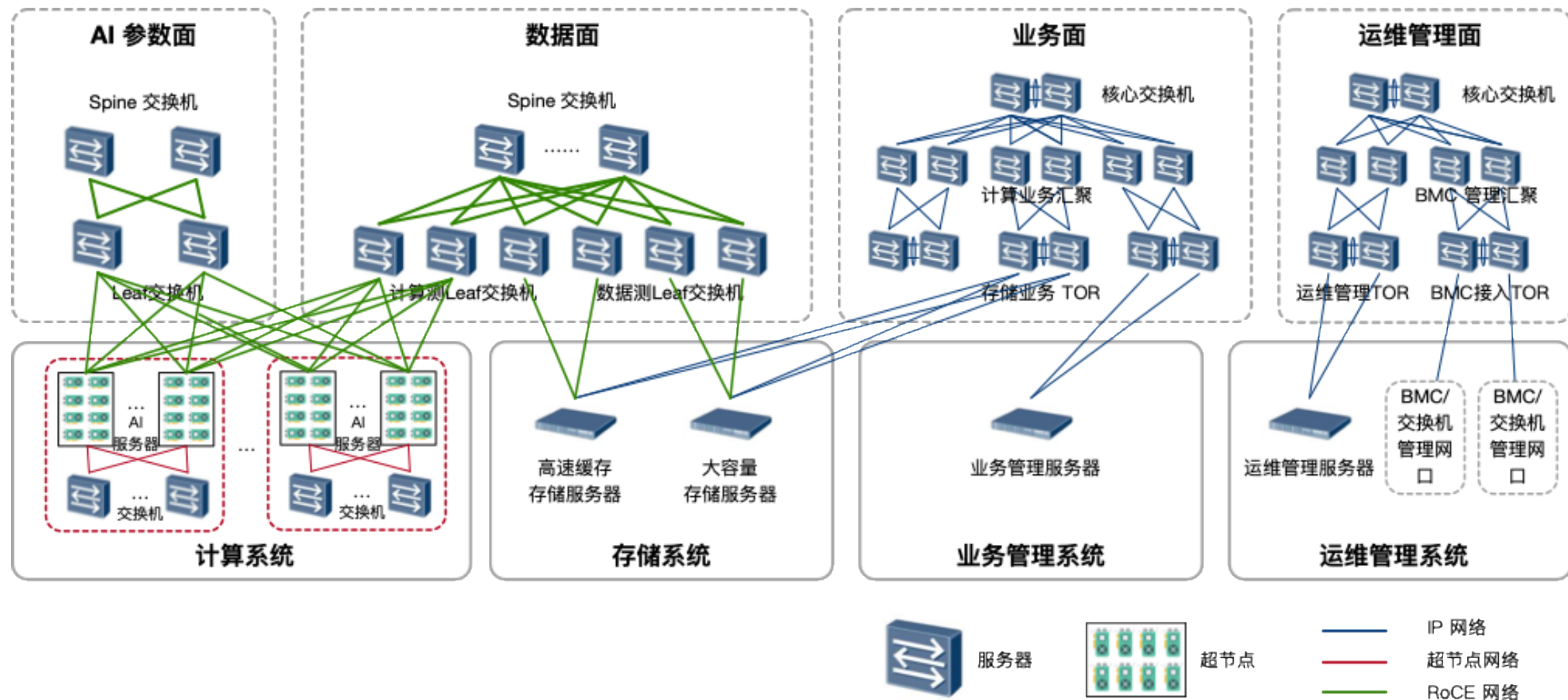


多平面网络，业务隔离术，数据使用专用网络通道

- 数据面 or 参数面：运输货物的卡车（传输训练数据）
- 存储面 or 数据面：专用冷链车（处理模型参数）
- 控制面 or 带内外面：调度中心的电话线（管理通信指令）
- 混合业务会导致：e.g. 运蔬菜的卡车占用了运药品的通道
- AI 集群通过多平面优化，显著提升训练效率



万卡AI 集群网络平面划分



框式 vs 盒式设备

- 万卡AI集群需要的交换机端口数远超普通数据中心，传统网络架构无法支撑。

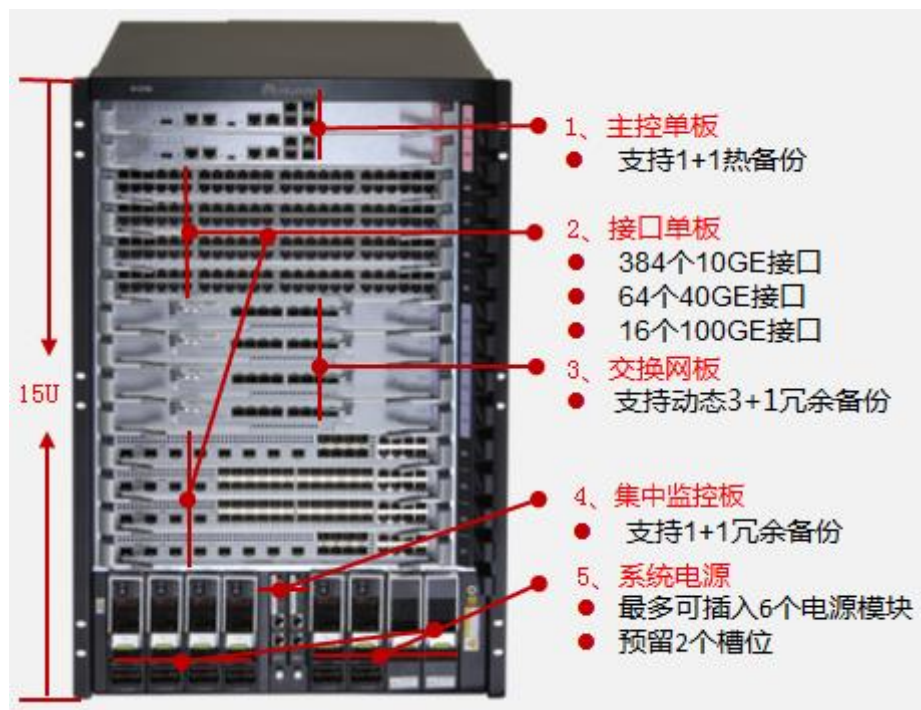
特性	框式交换机（Chassis）	盒式交换机（ToR）
物理形态	机架式铁框（多槽位）	独立铁盒（固定端口）
槽位价值	插业务板卡→灵活扩展	整机更换→扩容量跳跃
典型成本	空框50万+单板10万	单台3万 8万
适用场景	核心层（高密度大流量）	接入层（柜顶连接GPU）
特点	模块化设计，可热插拔，高可靠性	一体化设计，成本低，部署简单



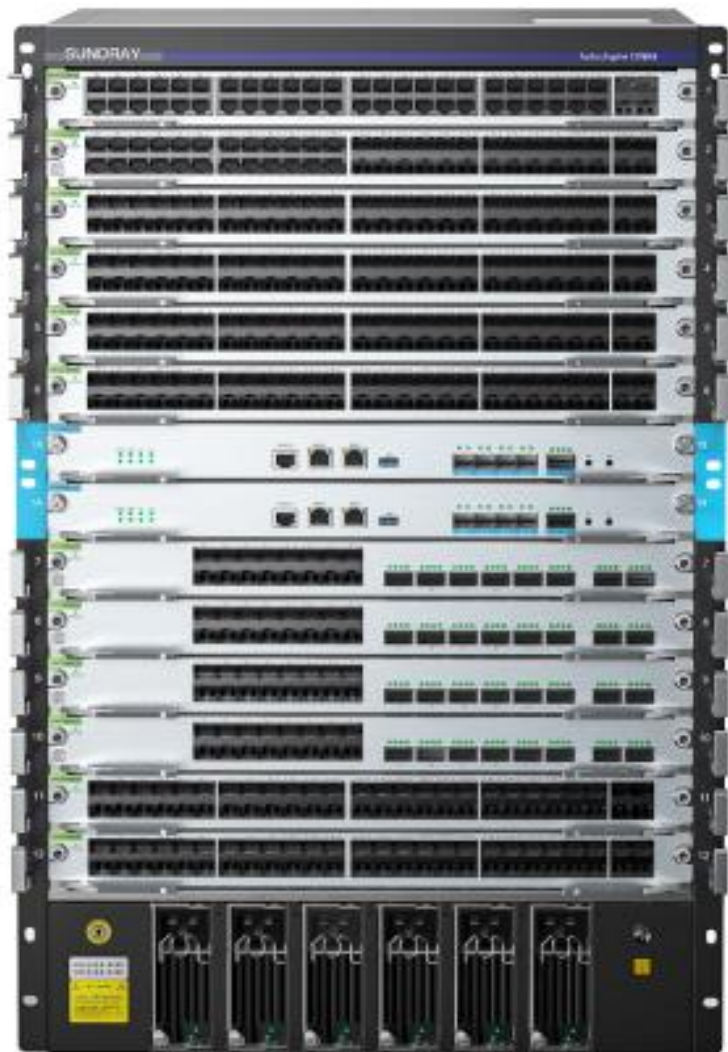
框式 vs 盒式设备

- 为什么框式更重要？

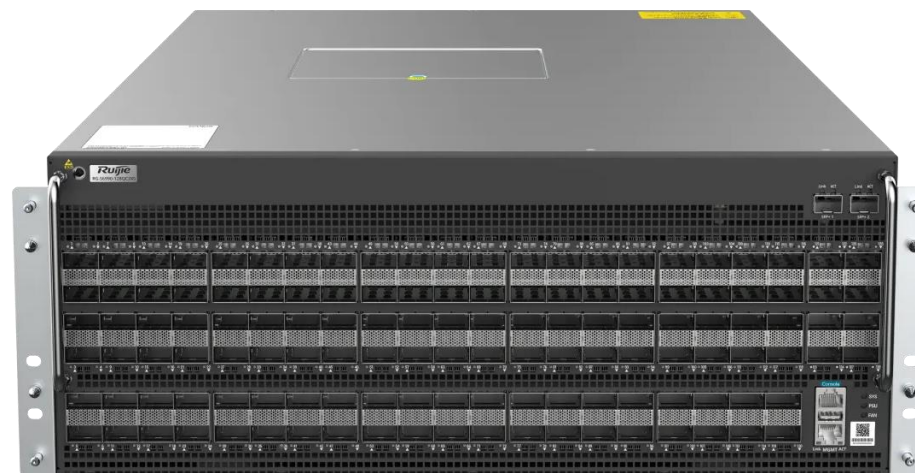
1. 万卡集群需要 7×24 小时不间断运行
2. 框式设备可在线更换故障模块，无需停机



框式 vs 盒式设备



64个800GE端口或者128个400GE端口



提供128个400GE端口，2个10GE端口



铜缆 vs 光纤：传输介质之战

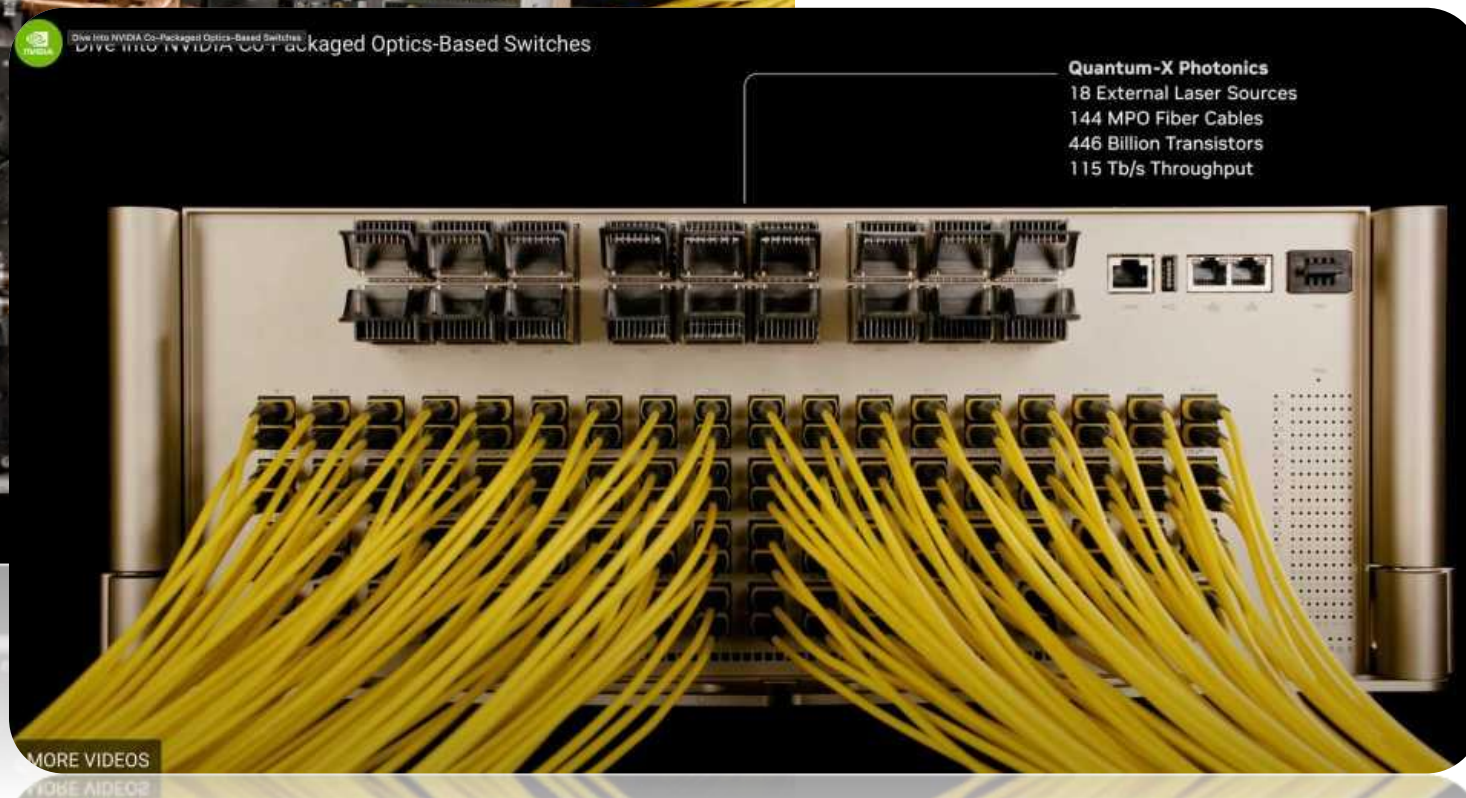
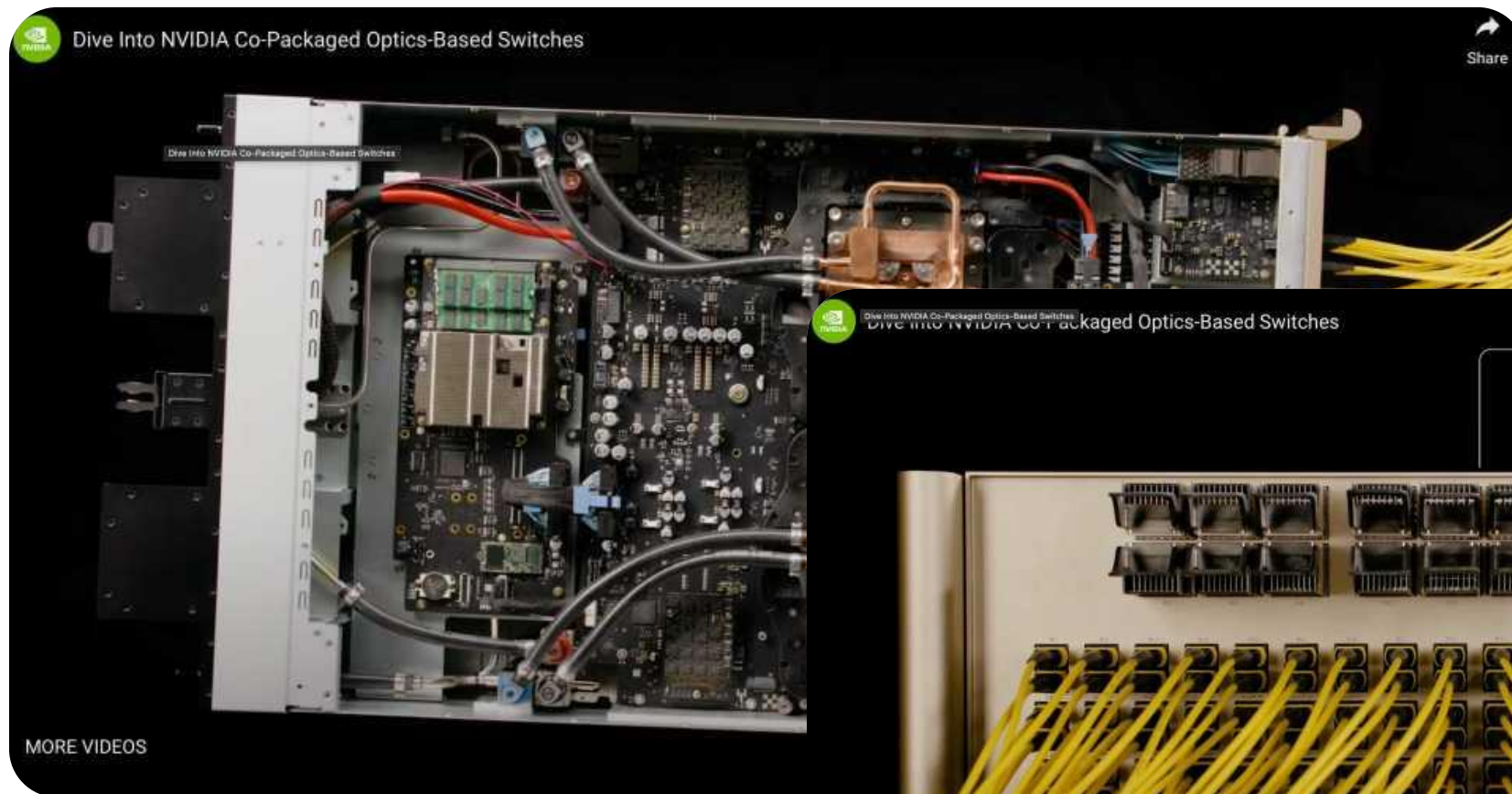
指标	高速铜缆（DAC）	光纤（AOC/光模块）
传输原理	电信号（电压波动）	光信号（激光脉冲）
信号衰减	7米后剧烈衰减（> 3dB）	100米衰减≈0.1dB
电磁干扰	怕强电机房（丢包率↑）	完全免疫
成本真相	\$200/3米，低	\$2000/100米（含光模块），高
抉择关键	仅限机柜内	跨机柜/跨楼层必选
传输距离	<5米	>100米
传输速度	25G-100G	100G-400G
适用场景	服务器内部、短距离连接	机柜间、数据中心间连接



铜缆 vs 光纤：传输介质之战



铜缆 vs 光纤：传输介质之战



04

通信机制 数据传输方式



单轨 vs 多轨通信

- 单轨通信:

1. 像单车道公路，数据排队传输
2. 常见于传统以太网
3. 容易拥堵，不适合AI训练

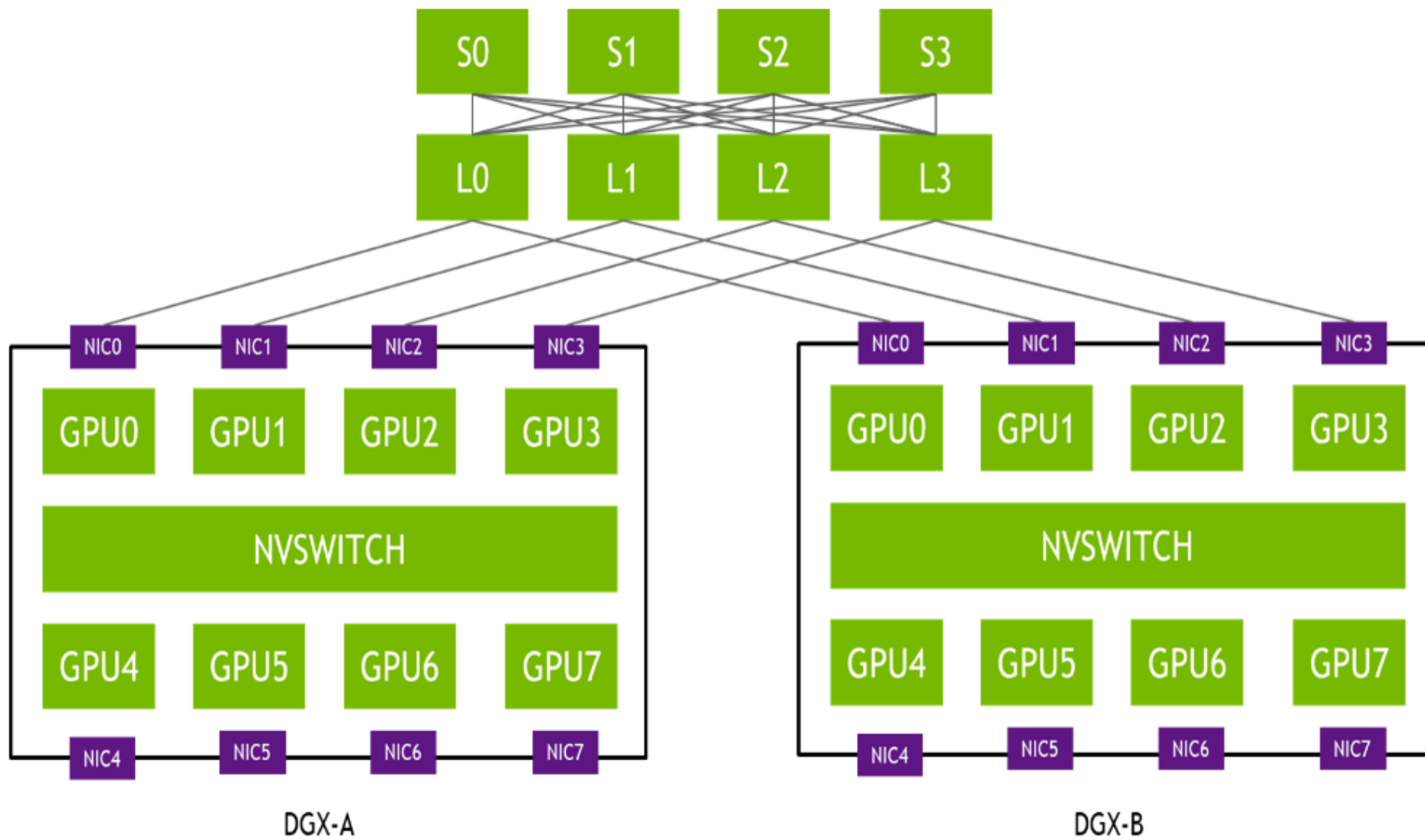
- 多轨通信:

1. 像多车道高速公路，数据并行传输
2. NVLink 4.0支持18条并行"数据车道"
3. 是Scale Up性能的关键

- Scale Up 依赖多轨通信实现单机内 GPU 闪电交互
- Scale Out 则需要智能路由确保跨服务器数据流高效



单轨 vs 多轨通信





总结与思考



Scale-up vs Scale-out

场景	Scale-up（纵向扩展）	Scale-out（横向扩展）
核心目标	单个“超级GPU”能力超强	千 GPU 实现高效协同
网络重点	NVLink 带宽（喂饱单个猛兽）	低延迟 + 高带宽（万卡 GPU 齐心协作）
致命瓶颈	链路带宽不足 → 算力闲置	网络延迟高 → 集体“等数据”挂机
延迟敏感度	纳秒级（1根头发掉落时长）	微秒级（蜜蜂振翅一次）
数据传输距离	机箱内（铜缆直连）	跨机柜（必须光纤）
拓扑依赖	全连接拓扑（不惜代价直连）	胖树拓扑（成本效率平衡）





Thank you

把 AllInfra 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 Infra to every person, home and
organization for a fully connected,
intelligent world.

Copyright © 2025 [Infrasys-AI](#) org. 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. [Infrasys-AI](#) org. may change the information at any time without notice.



ZOMI

GitHub github.com/Infrasys-AI/AllInfra

Book infrasys-ai.github.io



引用与参考

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

