



大模型系列 - 集合通信 -

集群互联



ZOMI

大模型业务全流程

数据 & 模型算法

模型训练 & 微调

模型验证 & 推理部署

2. 数据处理

开源数据
数据预处理
向量数据库

3. 模型算法

LLM模型架构
多模态一切皆Tokens

4. 模型训练

混合精度
梯度检查
梯度累积
...

5. 分布式并行

训练集群稳定性

6. 模型微调

全参微调
低参微调
指令微调

7. 模型验证

下游任务
测评标准
bench mark

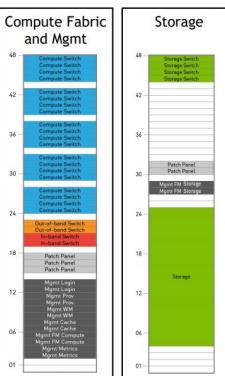
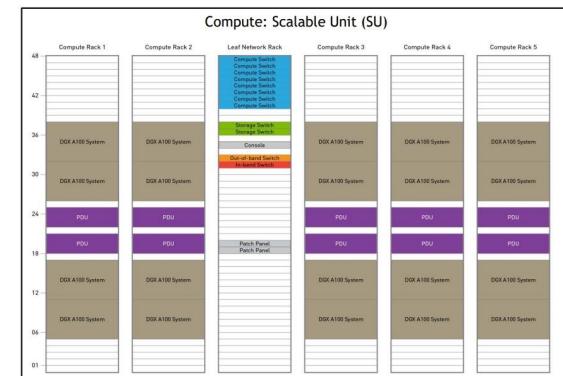
8. 推理与智能体

量化压缩
推理加速
9. Agent智能体

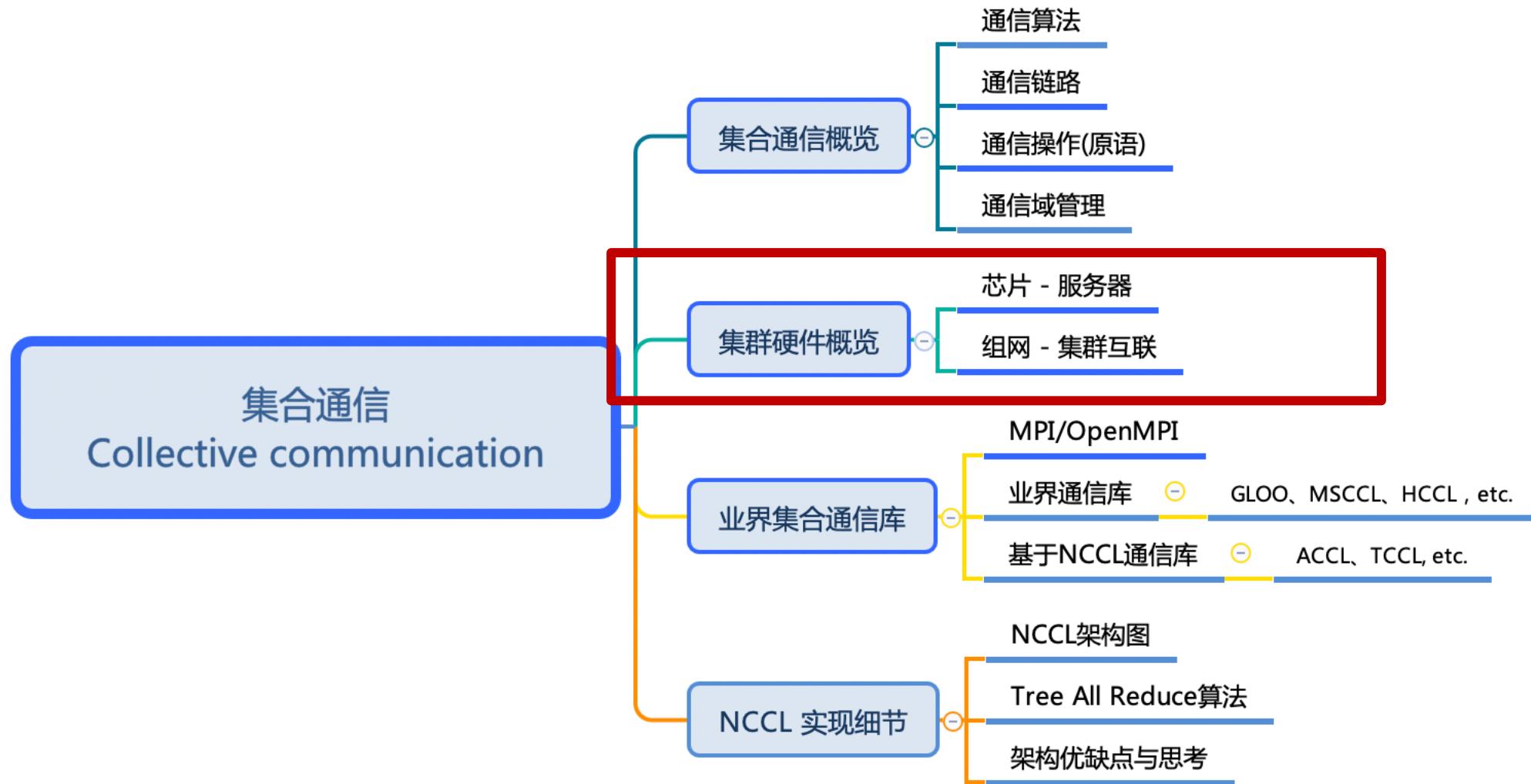
1. AI 集群建设

计算、存储、
网络
AI 集群机房建
设
AI 集群上线与
运维

集群算力准备

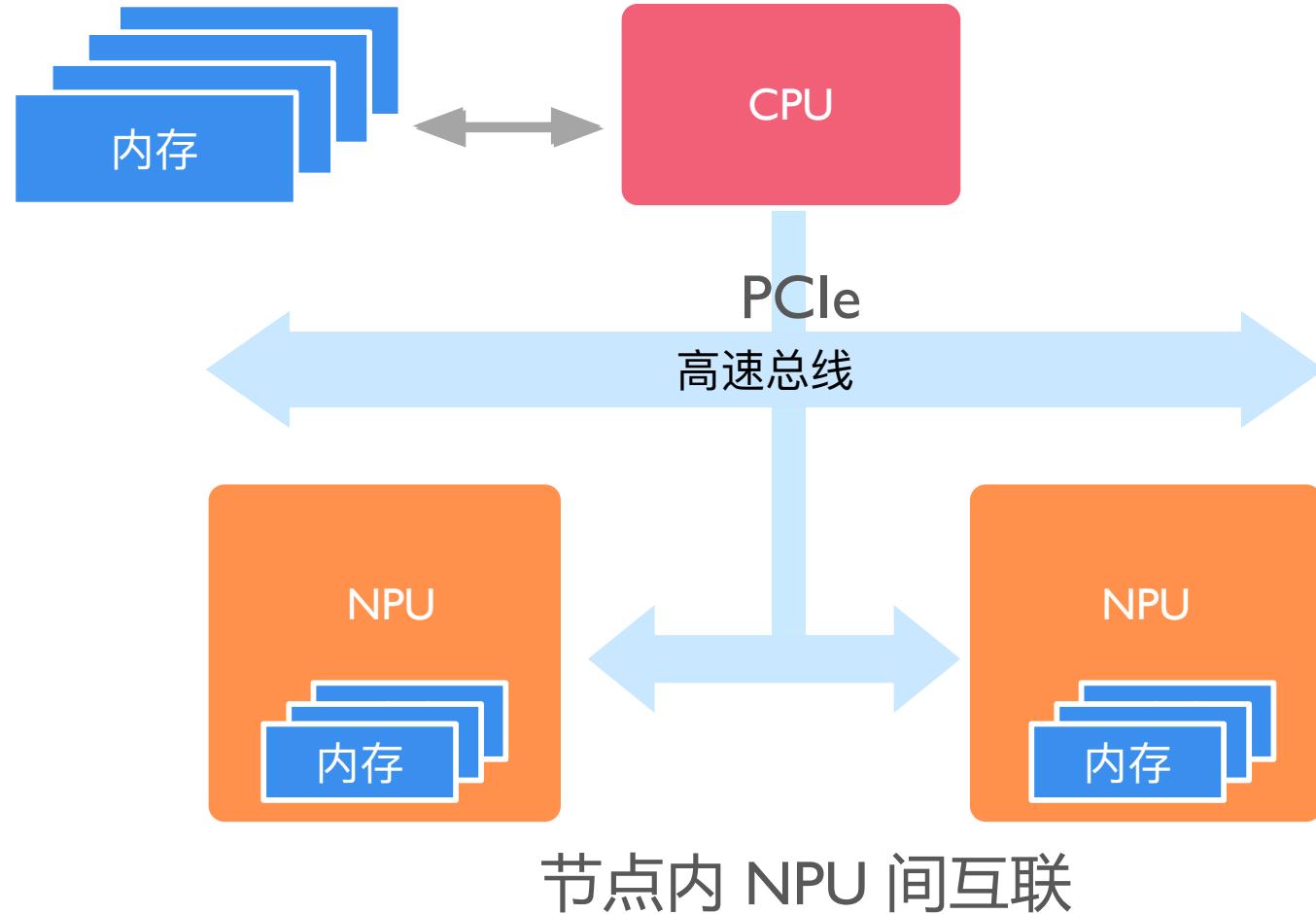


思维导图 XMind



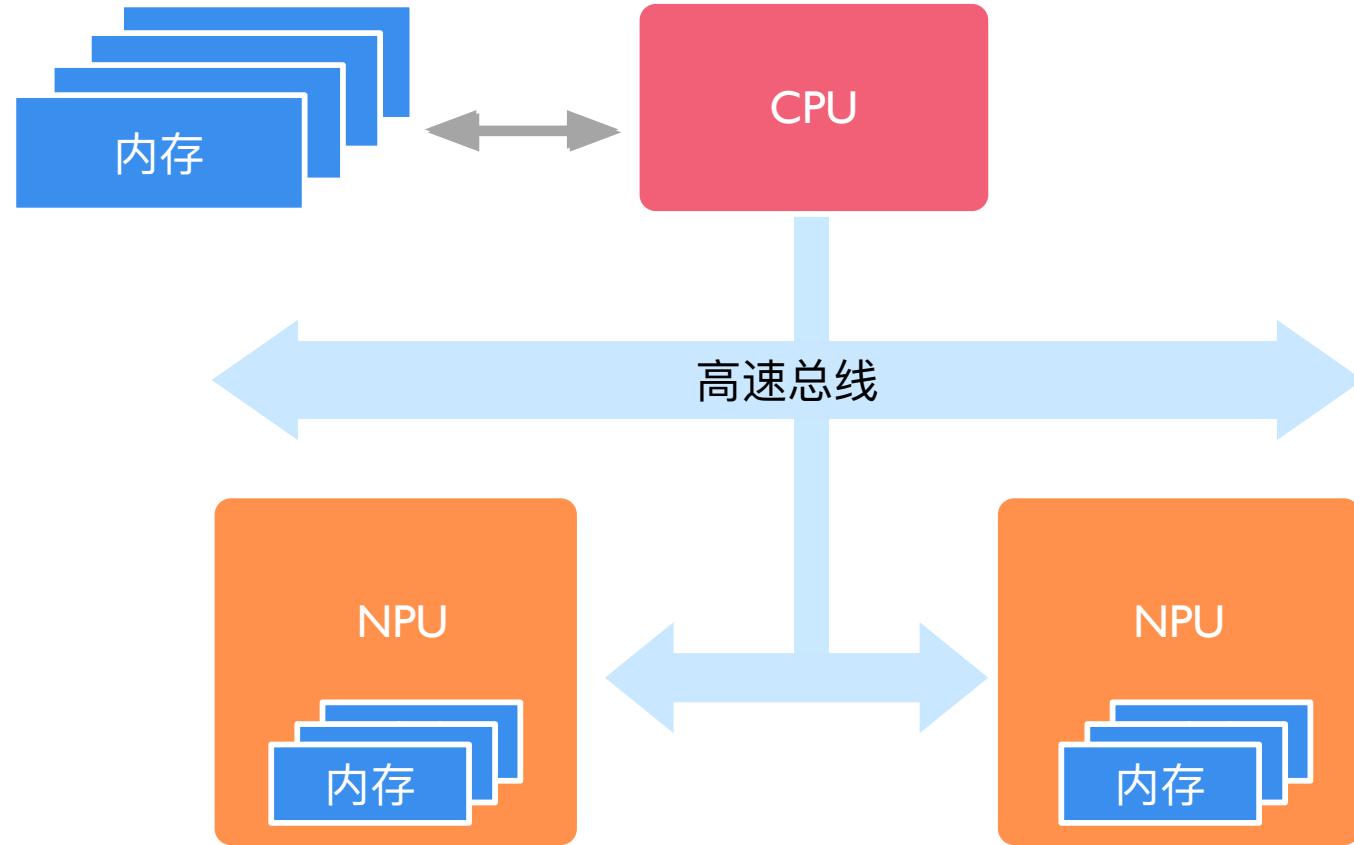
01. NPU 间互联

AI 服务器：一般架构



- 节点内 NPU 间互联
 - PCIe
 - 共享内存
 - 私有协议直连

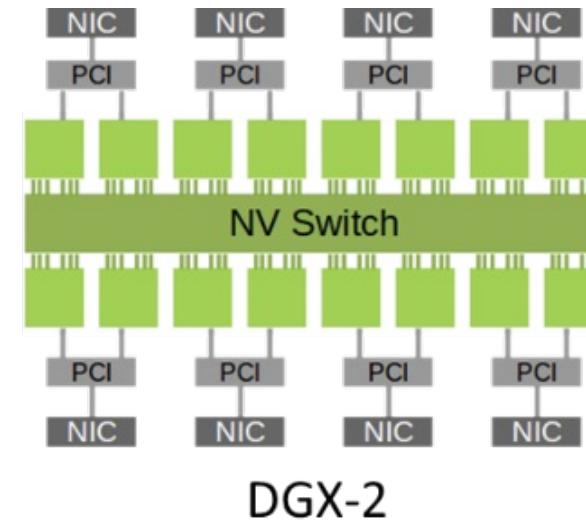
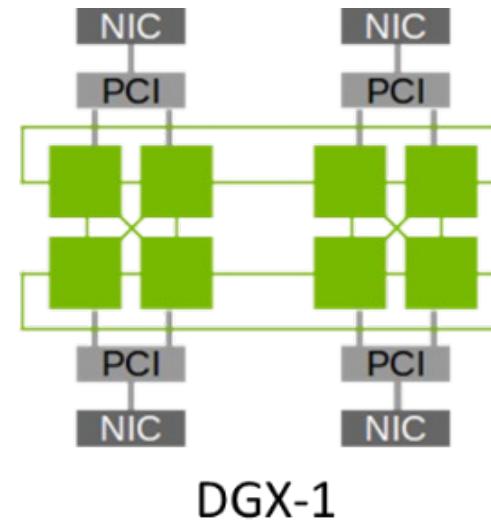
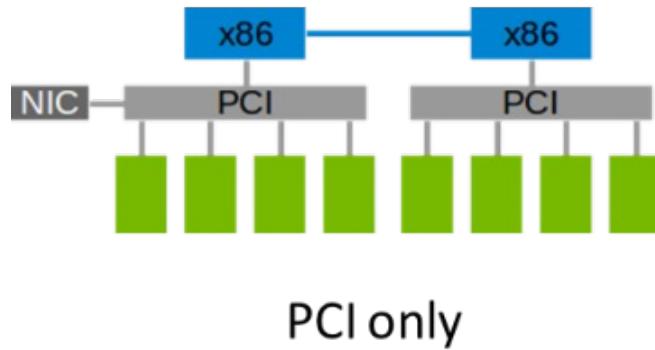
AI 服务器：一般架构



- CPU host, NPU Devices, 其各自有本地内存；
- 在 AI 集群下，NPU 之间需要高速交换数据，为了保证通信性能，NPU 之间设计专用高速互联通道；
- AI 计算的发展逐渐让 H-D 间由传统 PCIe 向多节点无损网络演进；

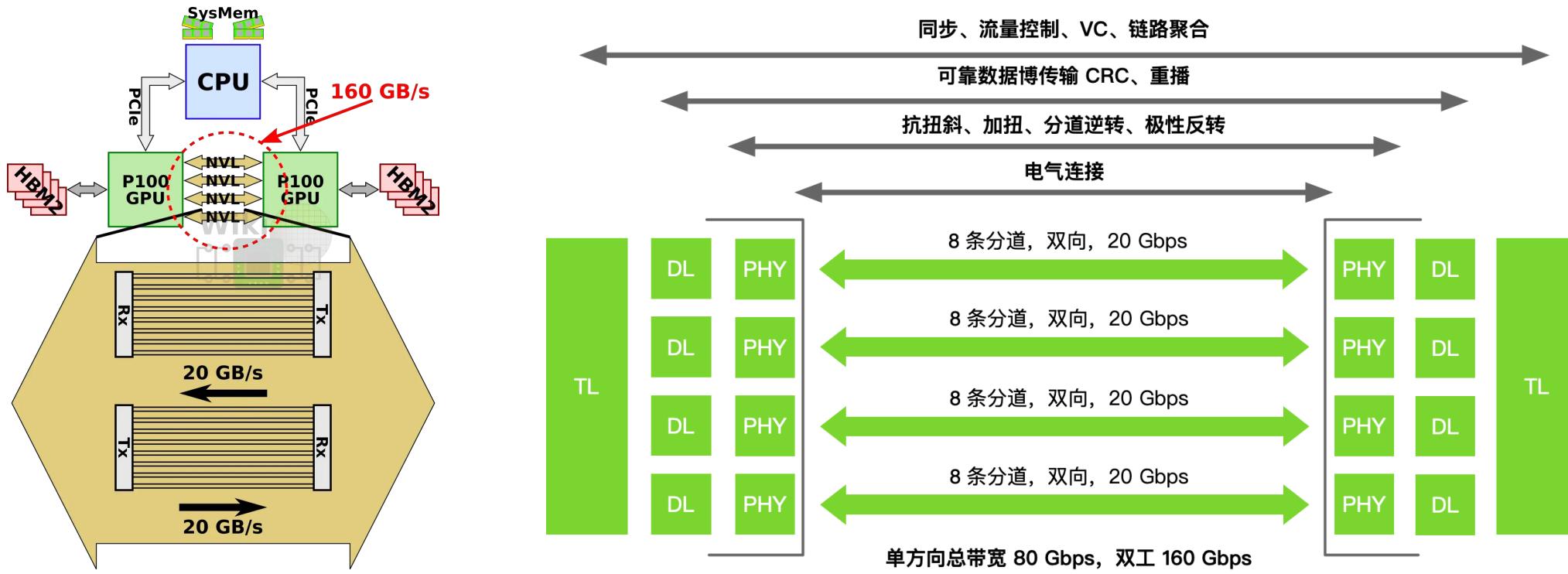
PCIe 互联

- PCIe5.0 应用最广泛计算机中外围设备互连标准，大多场景下高效和可靠，随 AI 计算规模越来越大，并行处理需求急剧增加，AI 计算集群需要万卡 NPU 互联。PCIe受限于带宽、延迟、数据传输效率，已成为 AI 集群互联瓶颈。因此，多节点无损网络协议诞生， NVLink。



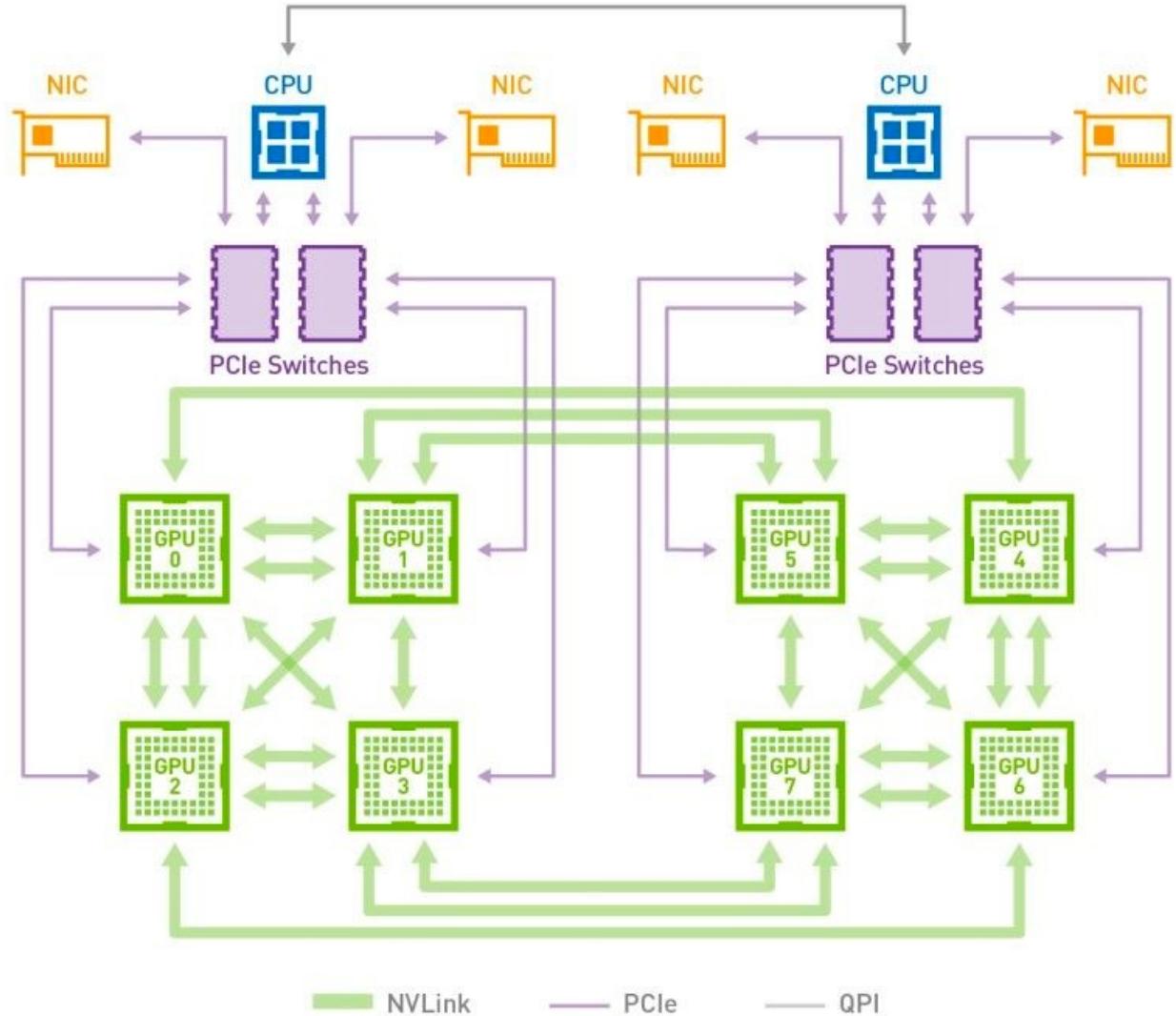
NVLink 出现

1. NVLink 设计目的，突破 PCIe 屏障，实现 GPU2GPU 及 CPU2GPU 片间高效数据交互。
2. NVLink 由软件协议组成，通过 PC 板上多对导线实现，让 GPU2GPU 间高速度收发显存数据。

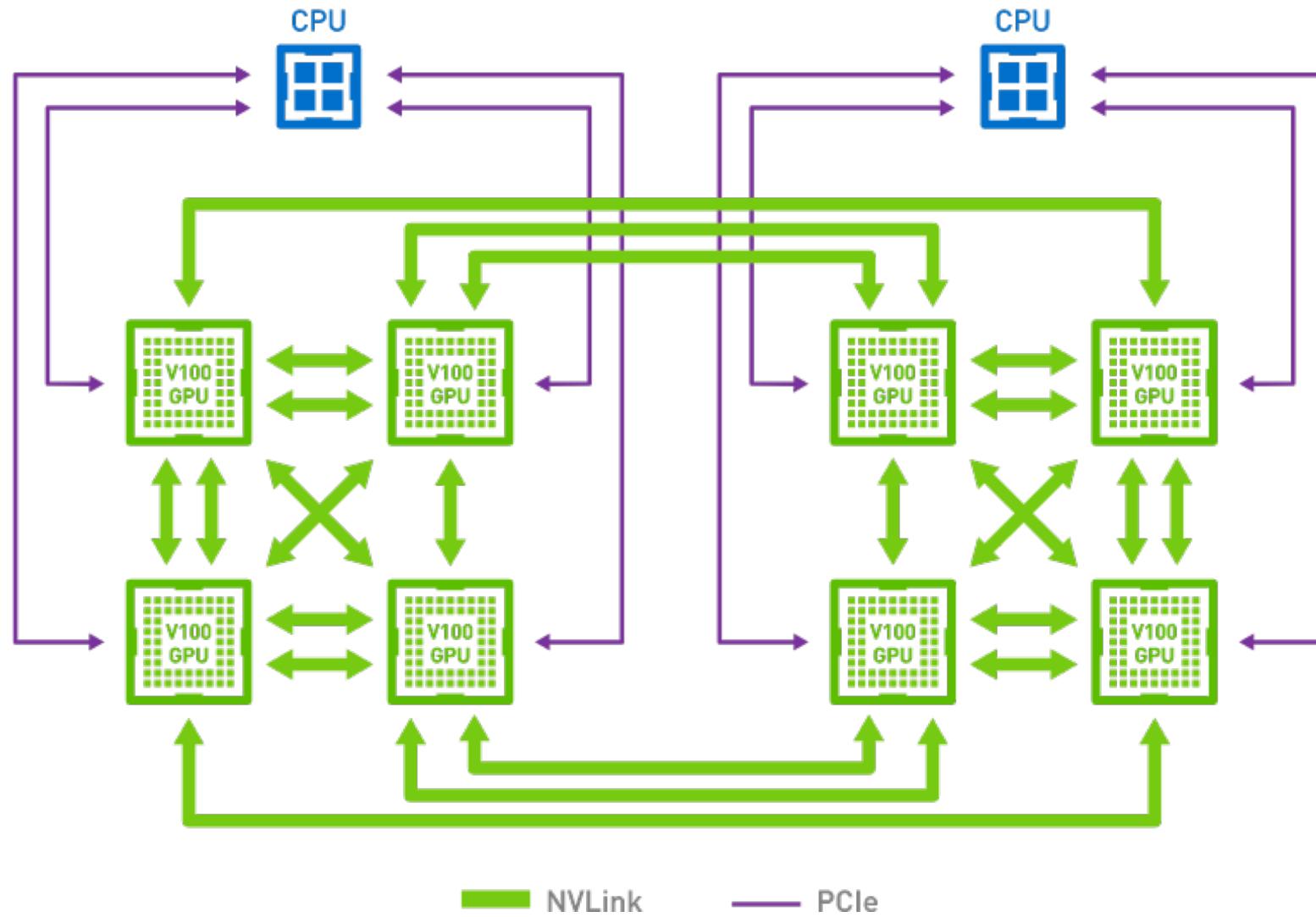


NVLink 出现

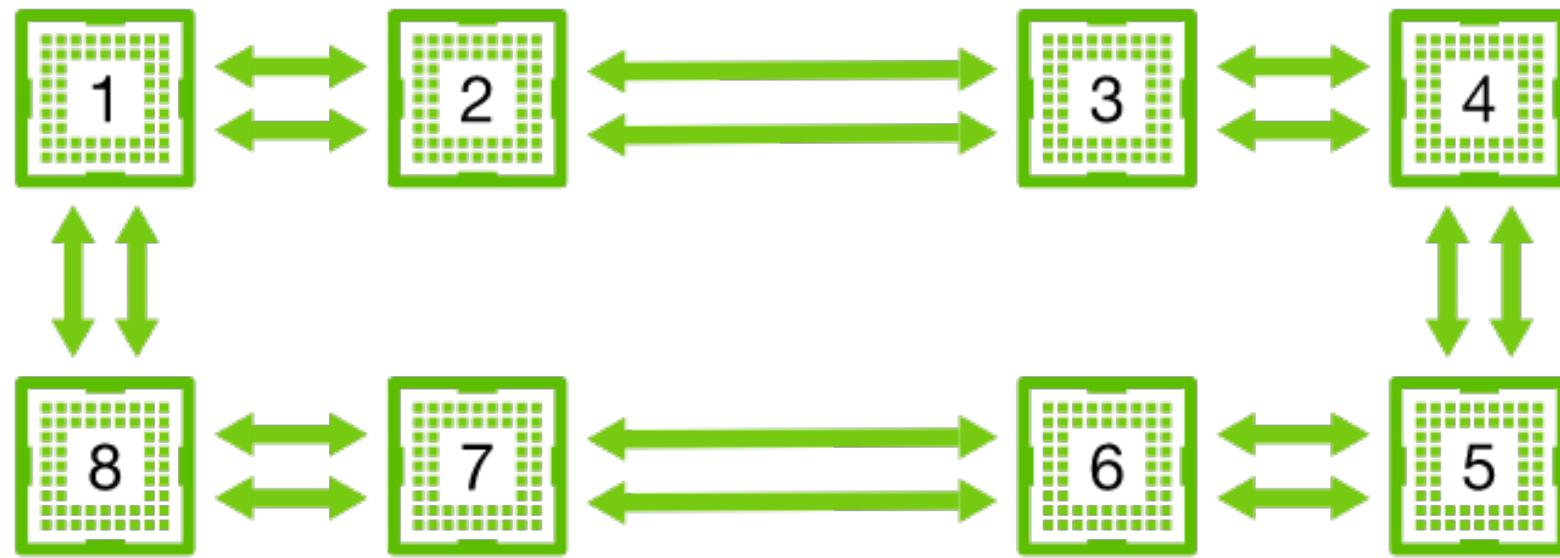
1. 通过 NVLink GPU 间互联链路采用点对点私有互联协议，绕开 PCIe 总线，直连多个 GPU 并组建成 GPU 计算阵列。
2. 第一代 NVLink 组成 GPU Mesh 结构，第四代 NVLink 集成在 NVSwitch 中。第四代 NVLink 传输速率实现 900GB/s，同期 PCIe 5.0 最大支持 128GB/s 双向带宽。



环同步 (Ring Synchronization)

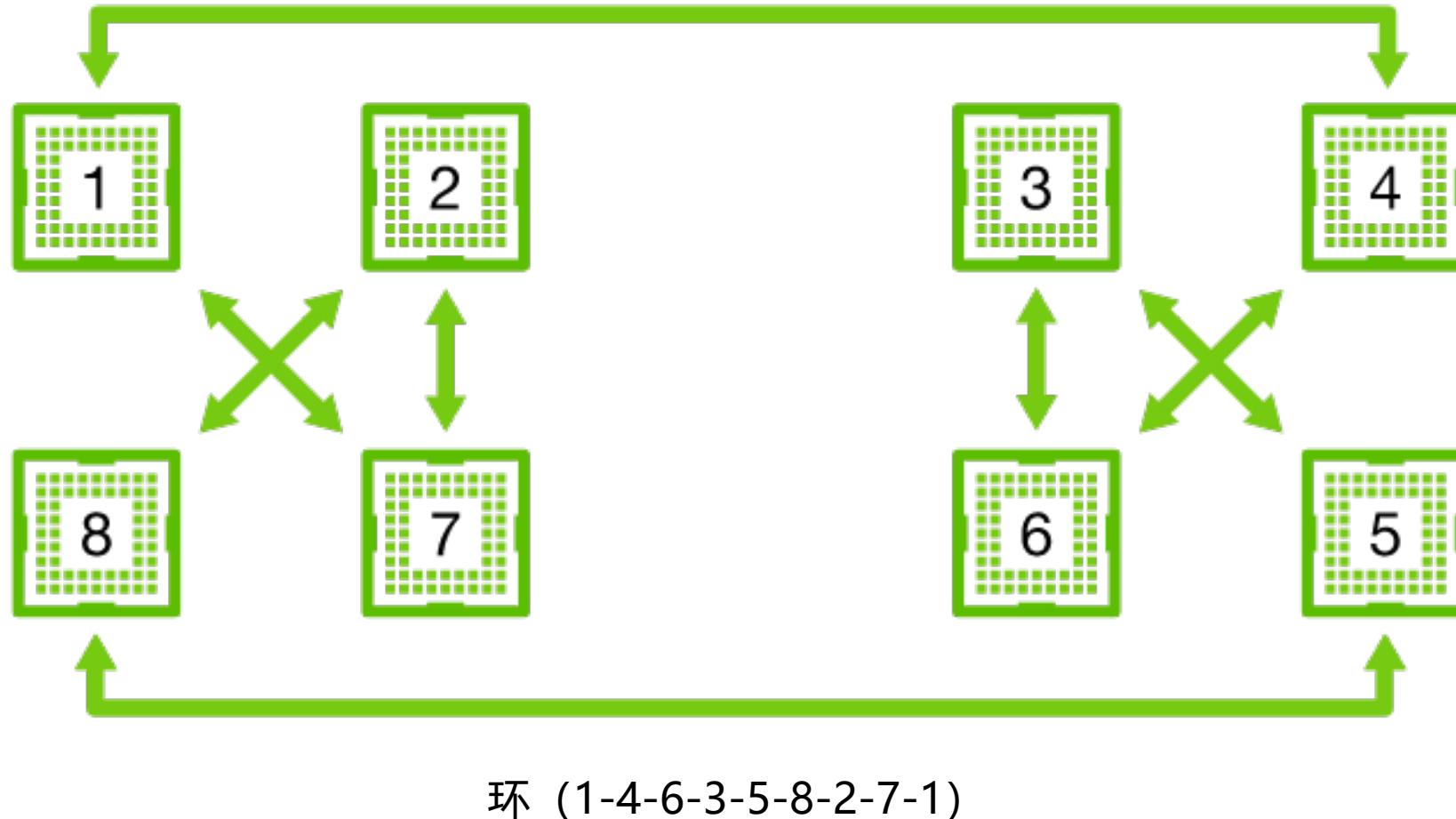


环同步 (Ring Synchronization)



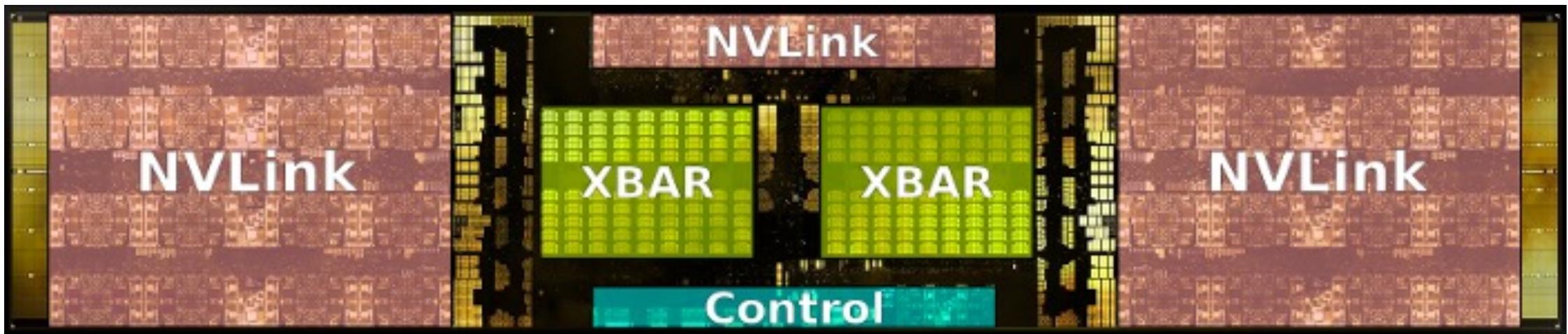
环 (1-2-3-4-5-6-7-8-1)

环同步 (Ring Synchronization)



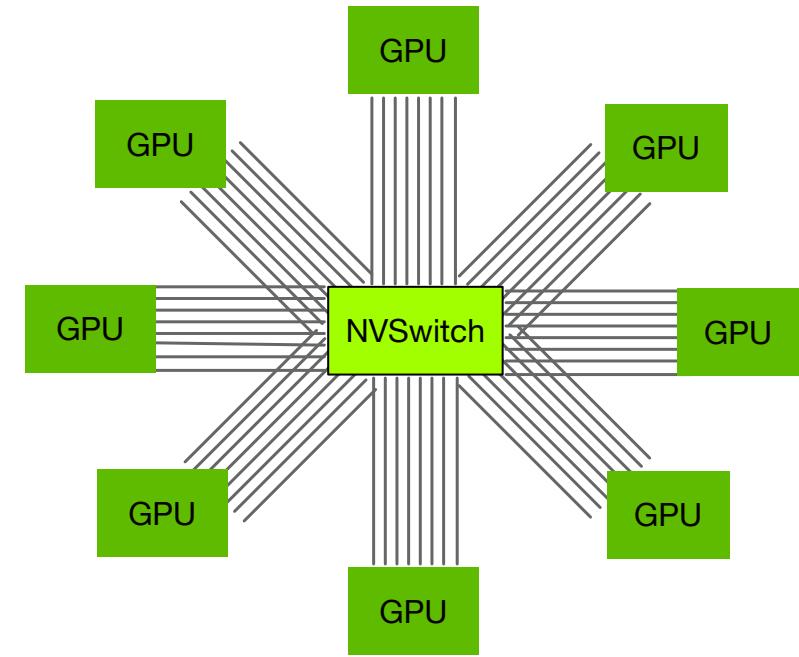
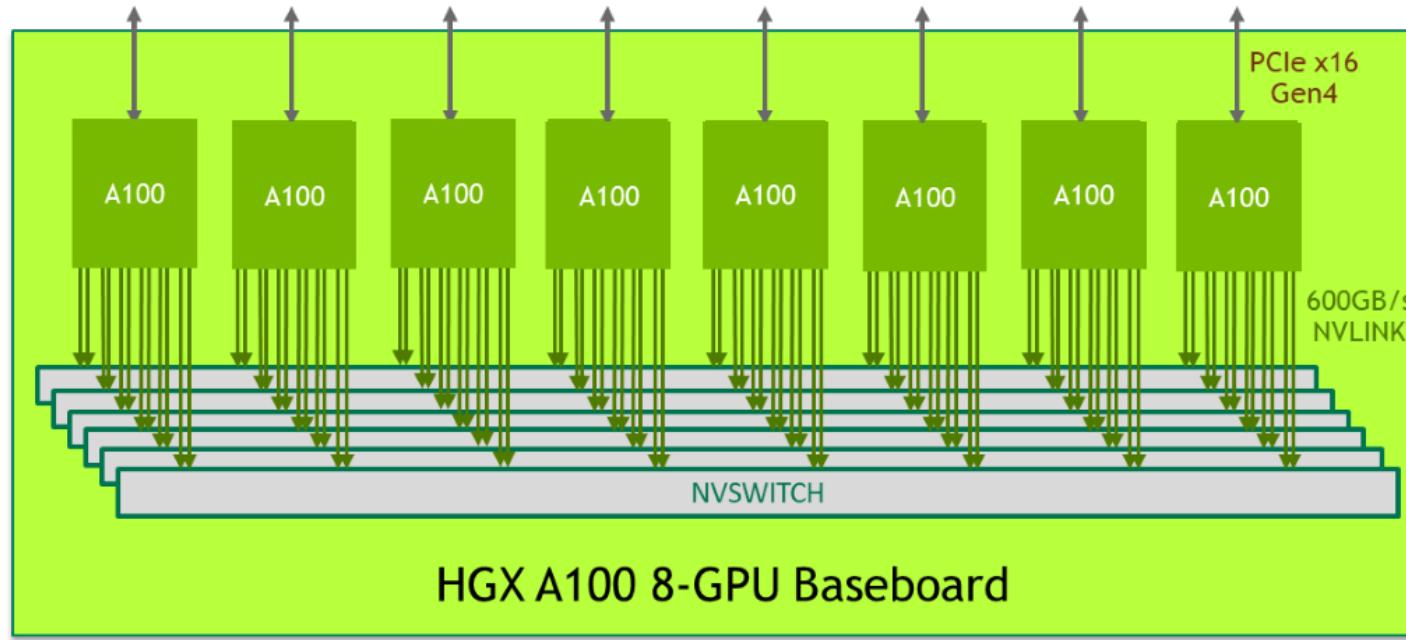
NVSwitch 出现

- NVSwitch 基于 NVLink 组成，可实现多节点 GPU 直联。不但绕开服务器内 PCIe 互联，还绕开服务器间以太网通讯，降低 GPU 通讯延迟，增强 AI 计算中数据同步效率。
- NVSwitch 基于无损数据交互互联网络，并非一颗简单芯片，而是一套复杂系统，需要系统的片间互联协议算法和匹配产品。



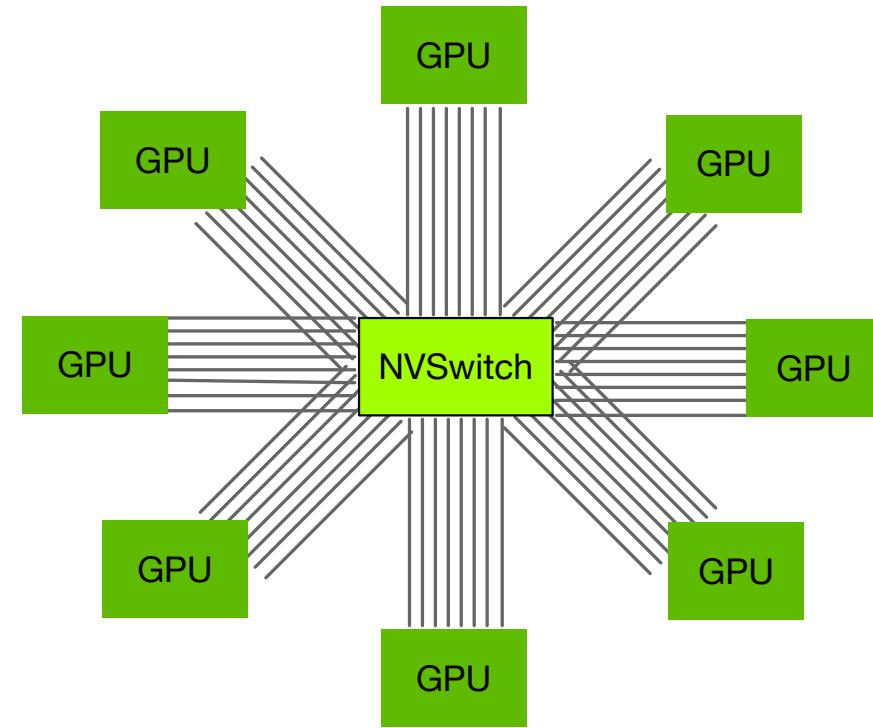
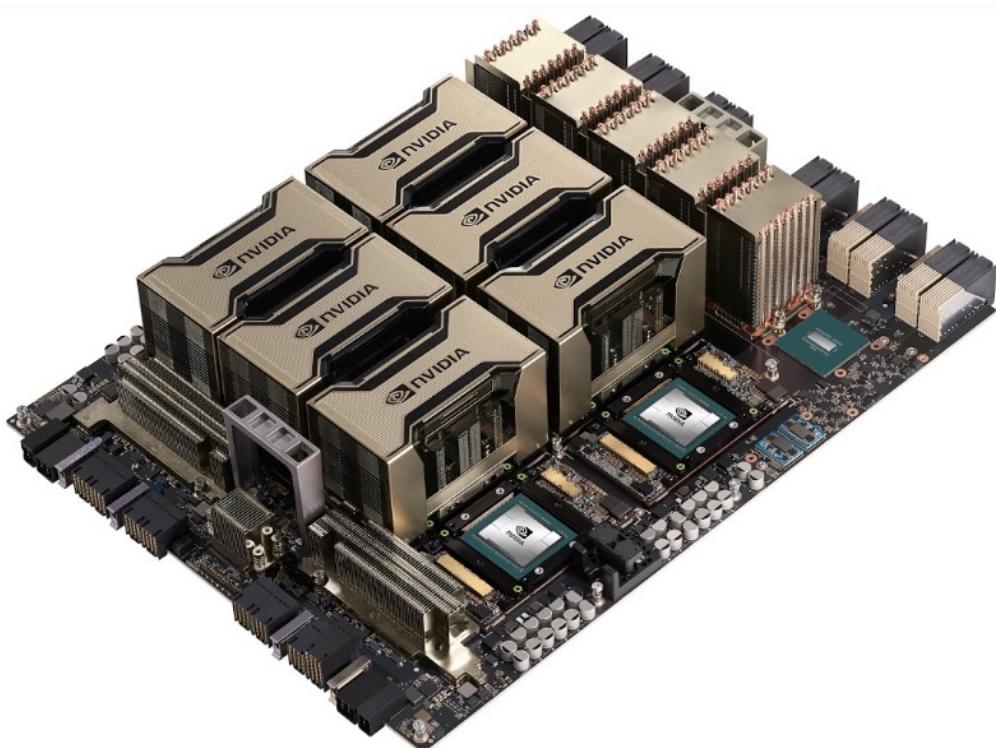
NVSwitch 出现

- NVIDIA DGX 其系统中 8 个 GPU 上 NVLink 通过 NVSwitch 芯片共享快速、直接连接，共同组成了一个 NVLink 网络，使服务器中的每一个 GPU 都成为 AI 集群中一部分。



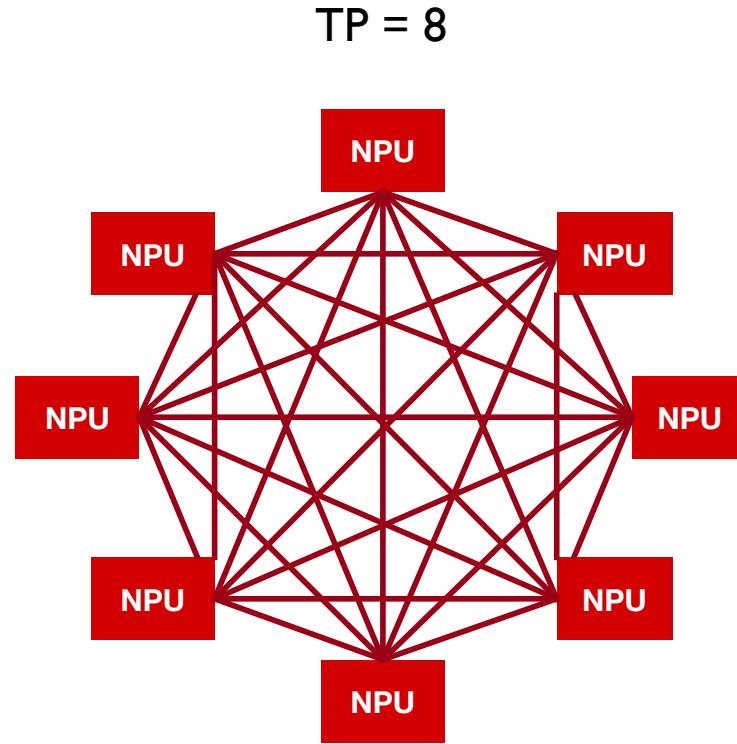
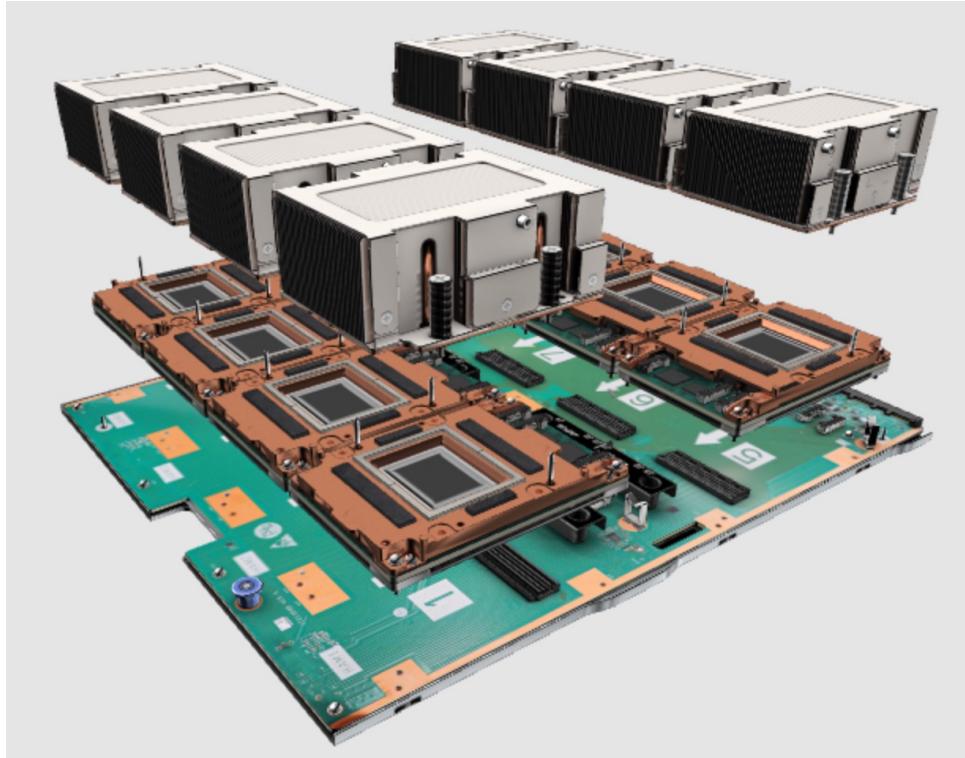
NVSwitch 出现

- 星型拓扑：TP=2/4/8 都可以将 AI 节点内通信链路用满。



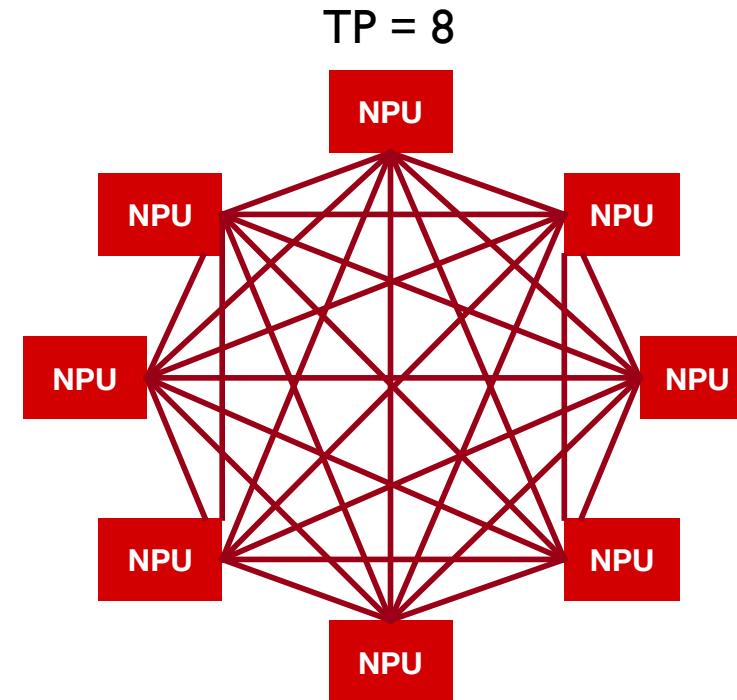
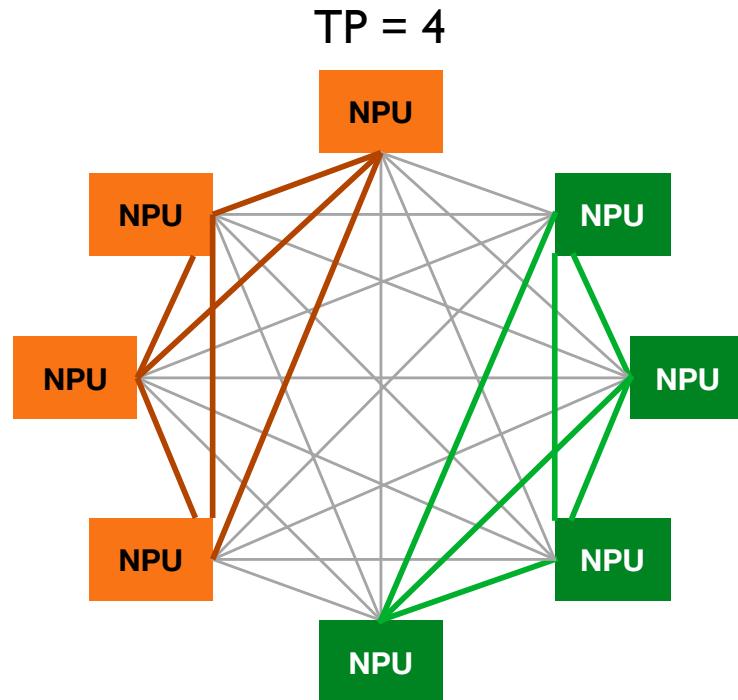
Ascend 服务器形态

- Ascend 服务器 NPU 卡间通过灵渠总线互联（ Full-Mesh 拓扑）；
- 每张卡出 Eth 网口，连接到 AI 参数面，通过 ROCE 互联。



Ascend 服务器形态

- Full-Mesh 拓扑:
 - TP=4, 仅仅能利用 3/7 的 AI 节点内通信链路；存在仅能利用一条 AI 节点内通信链路的 PP 流量。
 - TP=8, 可以将 AI 节点内通信链路用满；PP 通过 AI 节点 RDMA 链路通信。



片内互联挑战与技术壁垒

- **片间接口复杂性：**
 - 如何利用高速片间接口，使其满足芯片直连需求，又可完成设备交换功能；
- **交换算法扩展性：**
 - 如何利用已有计算体系实现 AI 集群内高速数据传输需求；
- **网络协议适配性：**
 - 如何构建覆盖 CPU、NPU、片内存储领域的网络协议，无缝适配已有硬件设备或上层软件系统。

片内互联挑战与技术壁垒

- 目前仅有英伟达、华为昇腾、AMD、Google 等少数公司实现高速多节点片间互联网络和私有协议，因此在 AI 训练服务器集群构成竞争壁垒。同时，这也是国内 N + 高性能 AI 芯片公司，却少高速互联芯片企业重要原因。

intelliifusion

云天励飞

Cambricon

寒武纪

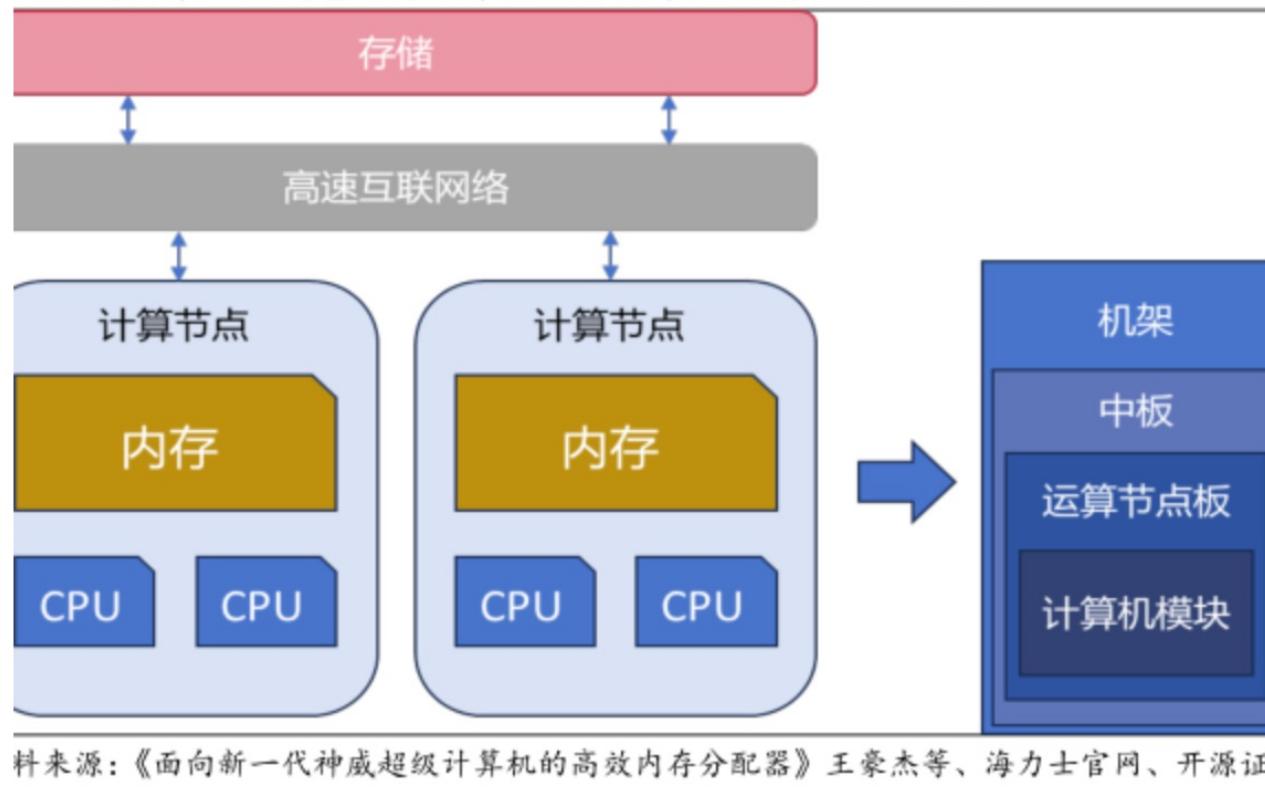
Axera
爱芯元智

芯驰 SemiDrive

02. 节点间互联

SuperPOD

48：超算系统主要包括存储节点、计算节点与高速互联网络



超算系统主要包括存储节点、计算节点与高速互联网络

• 节点内 NPU 间互联

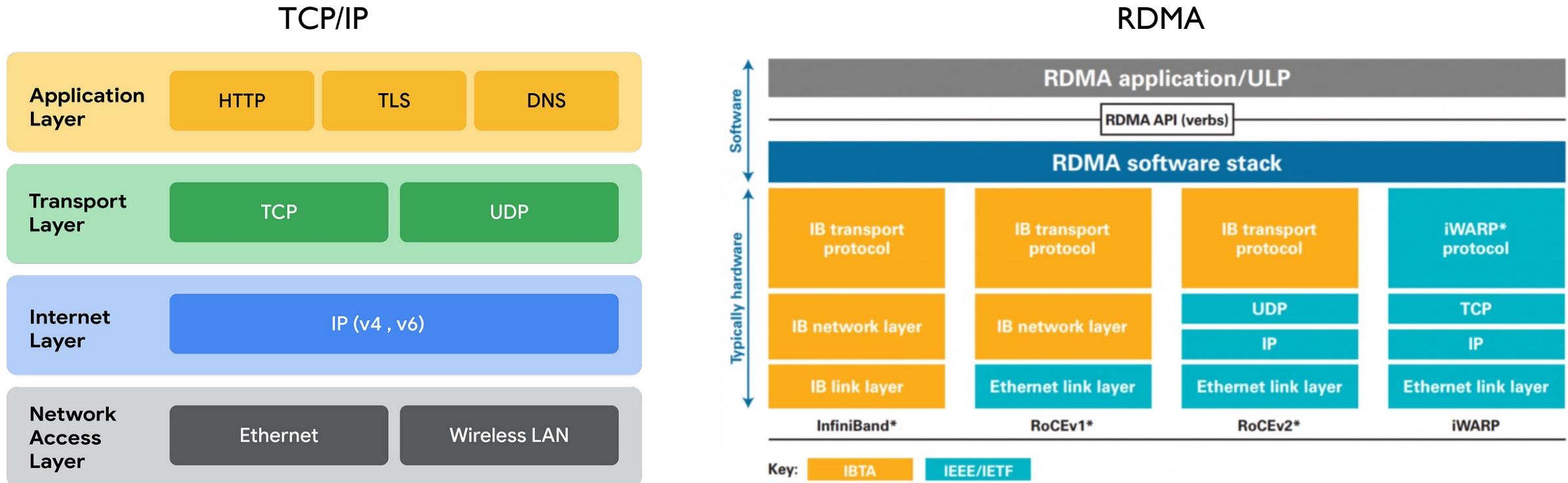
- PCIe
- 共享内存
- 私有协议直连

• 计算节点间通信

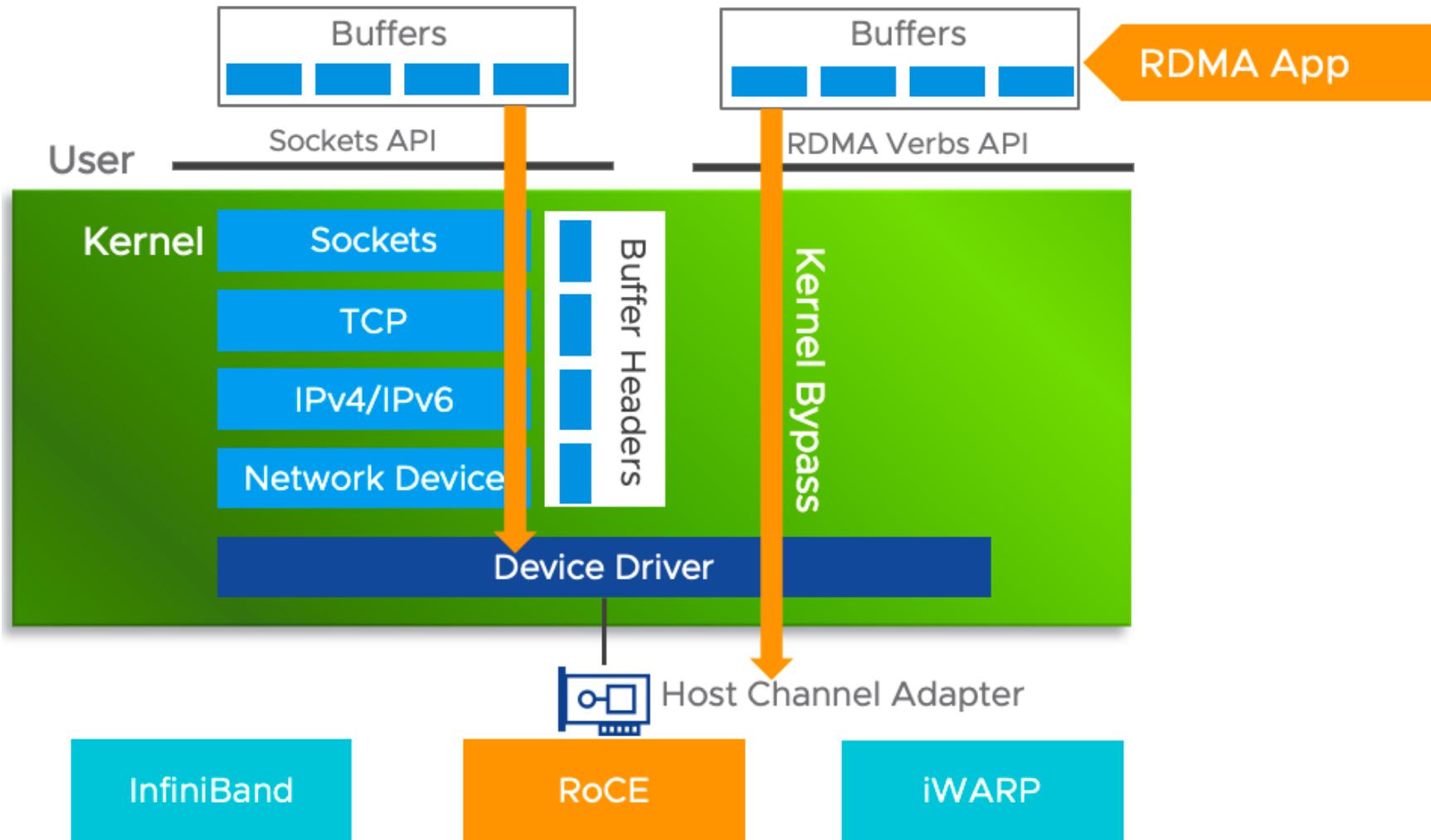
- TCP/IP 网络
- RDMA 网络

从 TCP/IP 到 RDMA

- 万卡集群已经成为大模型训练标配，面对大模型规模庞大的数据交互，传统 TCP/IP 协议逐渐被 RDMA（远程直接内存访问）技术取代。



RDMA 主要实现方式

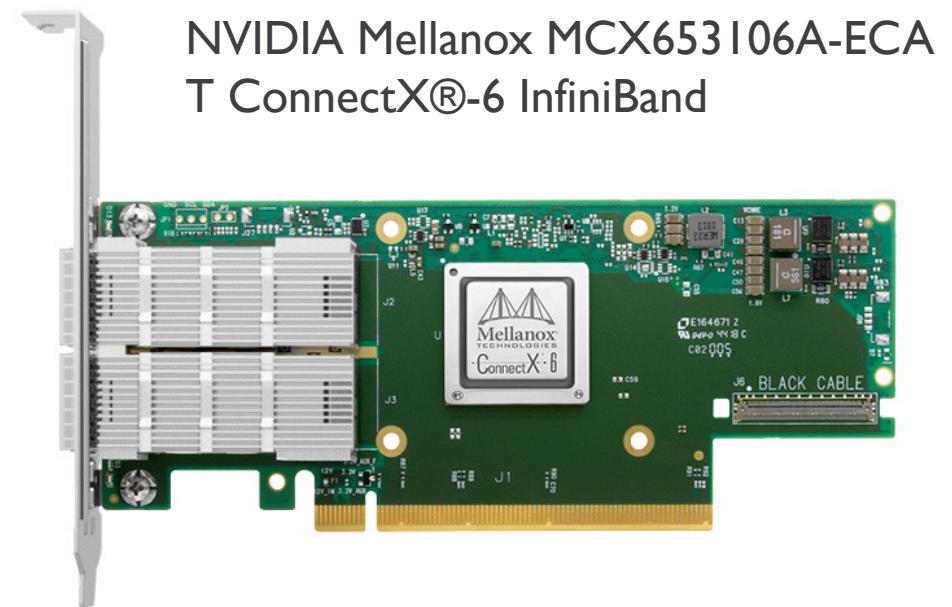
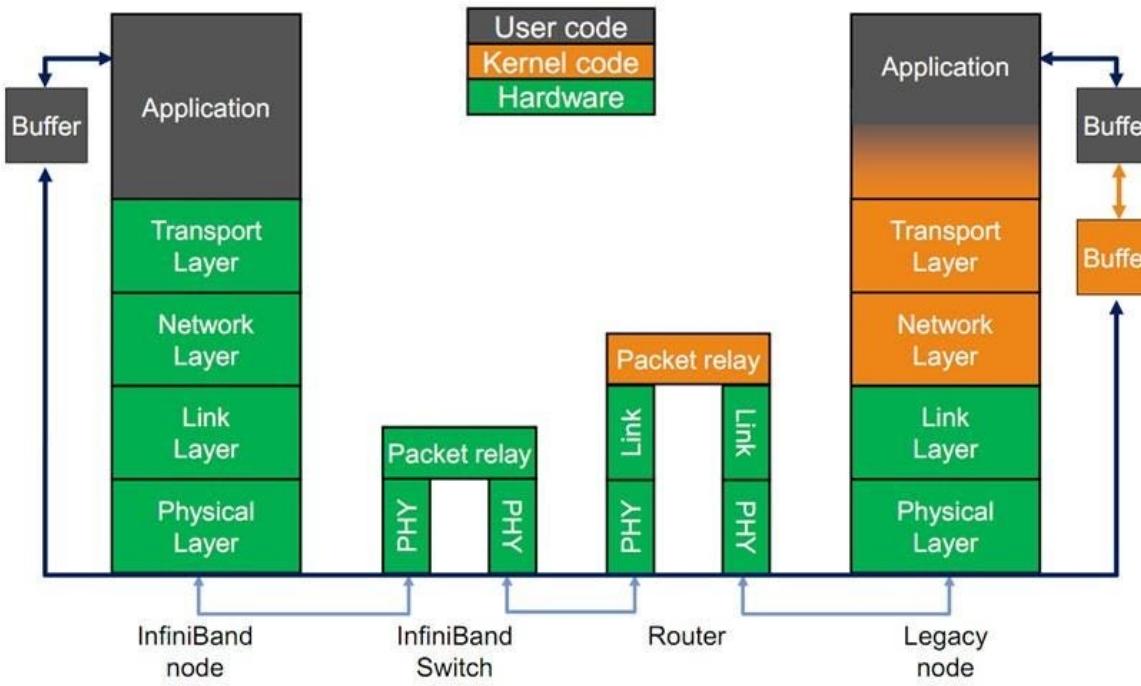


RDMA 主要实现方式

	InfiniBand	iWARP	RoCE
性能	最好	稍差 (受TCP影响)	与 InfiniBand 相当
成本	高	中	低
稳定性	好	差	较好
交换机	IB交换机	以太网交换机	以太网交换机

InfiniBand：从超算到AI

- InfiniBand (IB) 网络起源于超算 HPC，为超算提供低延迟、高吞吐量的数据传输解决方案。2010 年底 Mellanox 完成对 InfiniBand 交换机厂商 Voltaire 公司收购，2021 年英伟达以 69 亿美元现金完成以色列芯片商 Mellanox Technologies 收购。



InfiniBand：从超算到AI

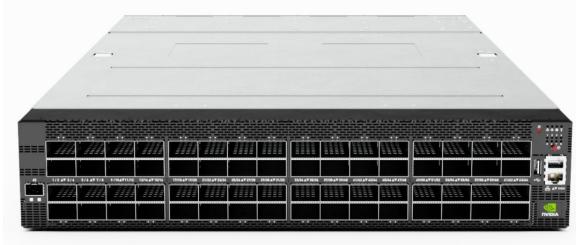
- 使用 InfiniBand 部署 AI 训练集群远比基于以太网昂贵，网络成本占 AI 集群成本 ~20%。
- 随着 AI 集群规模化增长，很多互联网企业重新寻求性价比高的网络形式。

6857-48S6CQ-EI-6800系列-华为交换机



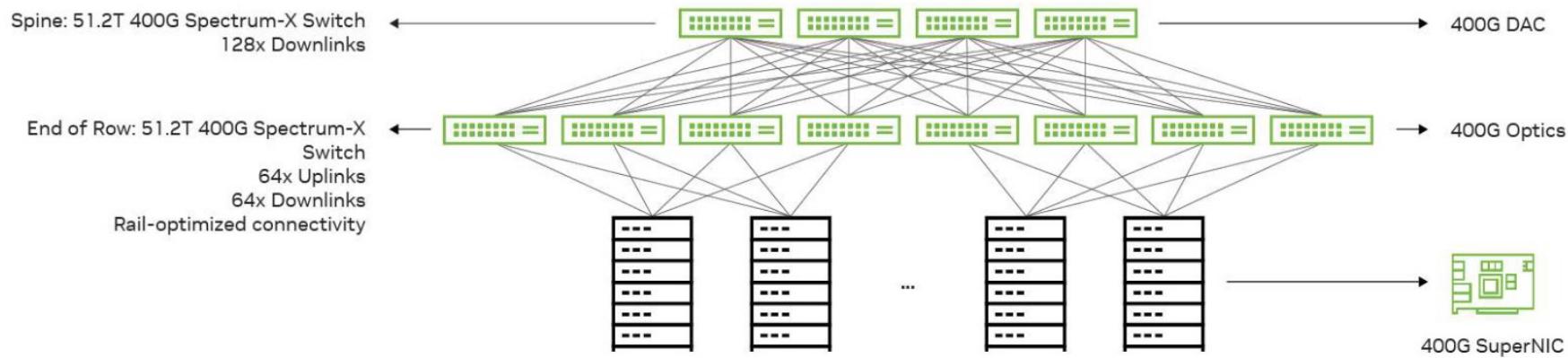
RoCE

- 下一代高速以太网，基于以太网 RoCE (RDMA over Converged Ethernet) 网络协议，可利用现有以太网基础设施，构建网络性能接近或等同的 RDMA 网络。



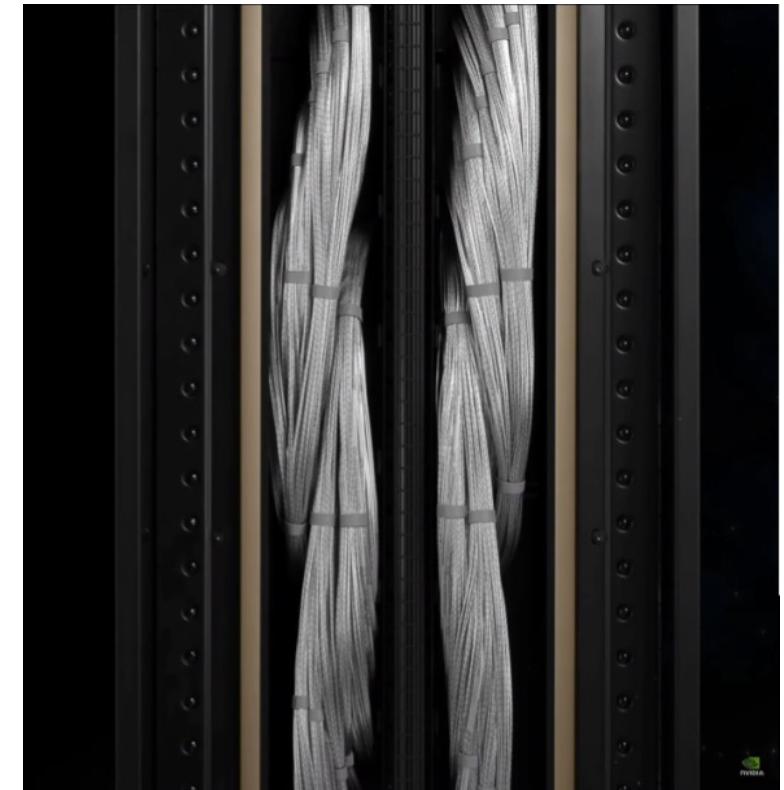
NVIDIA Spectrum-4

NVIDIA Spectrum SN5000 系列交换机是第五代 Spectrum 以太网交换机，专为加速超大规模生成式 AI 网络而打造。SN5000 交换机的端口速度高达 800Gb/s，可为每个数据中心提供加速以太网，且不会影响性能和功能。



RoCE

- AI 计算集群的高性能远距离传输，主要以 InfiniBand 和 RoCE 为主。目前，超远距离线缆的解决方案也在从传统铜线向光学传输不断进化。

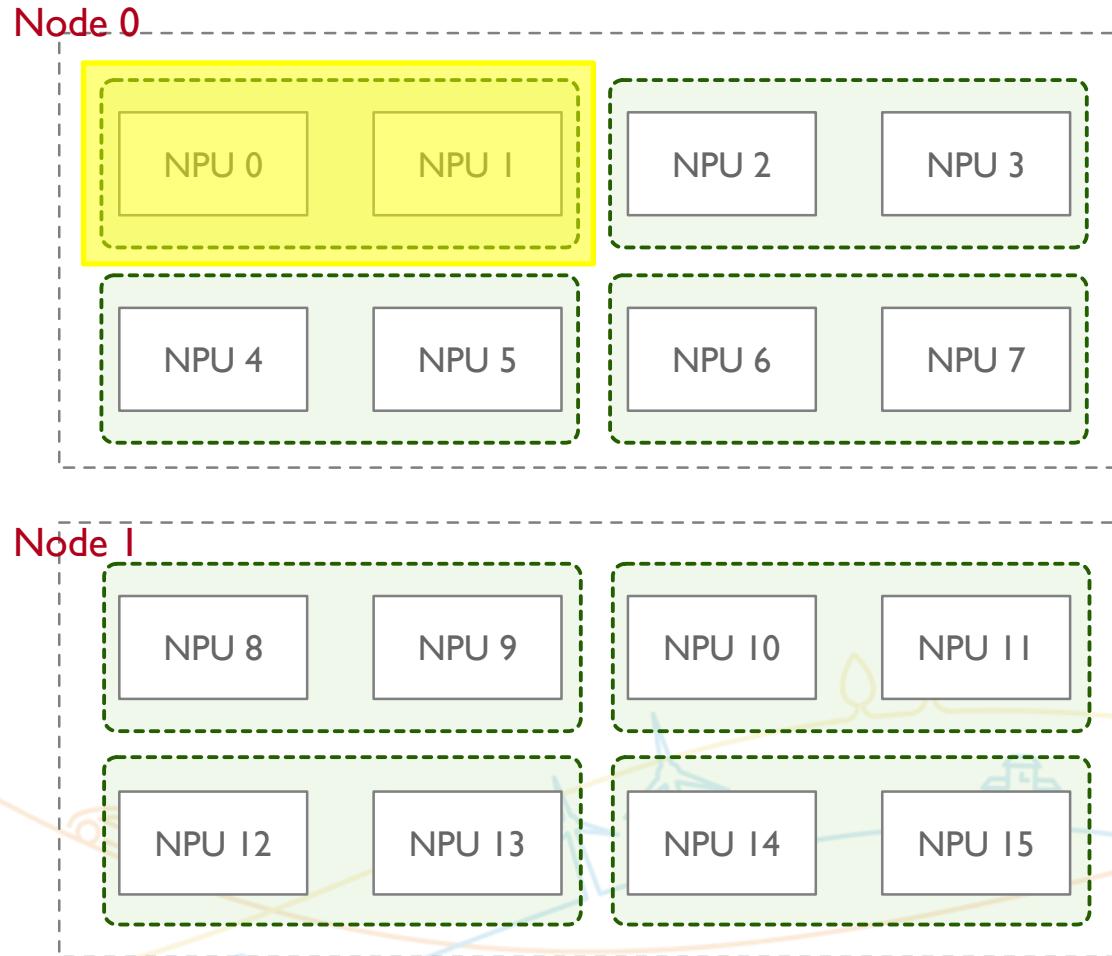


- <https://info.support.huawei.com/computing/server3D/res/server/atlas900poda2/index.html?lang=cn>
- <https://info.support.huawei.com/computing/server3D/res/server/atlas800ta2/index.html?lang=cn>

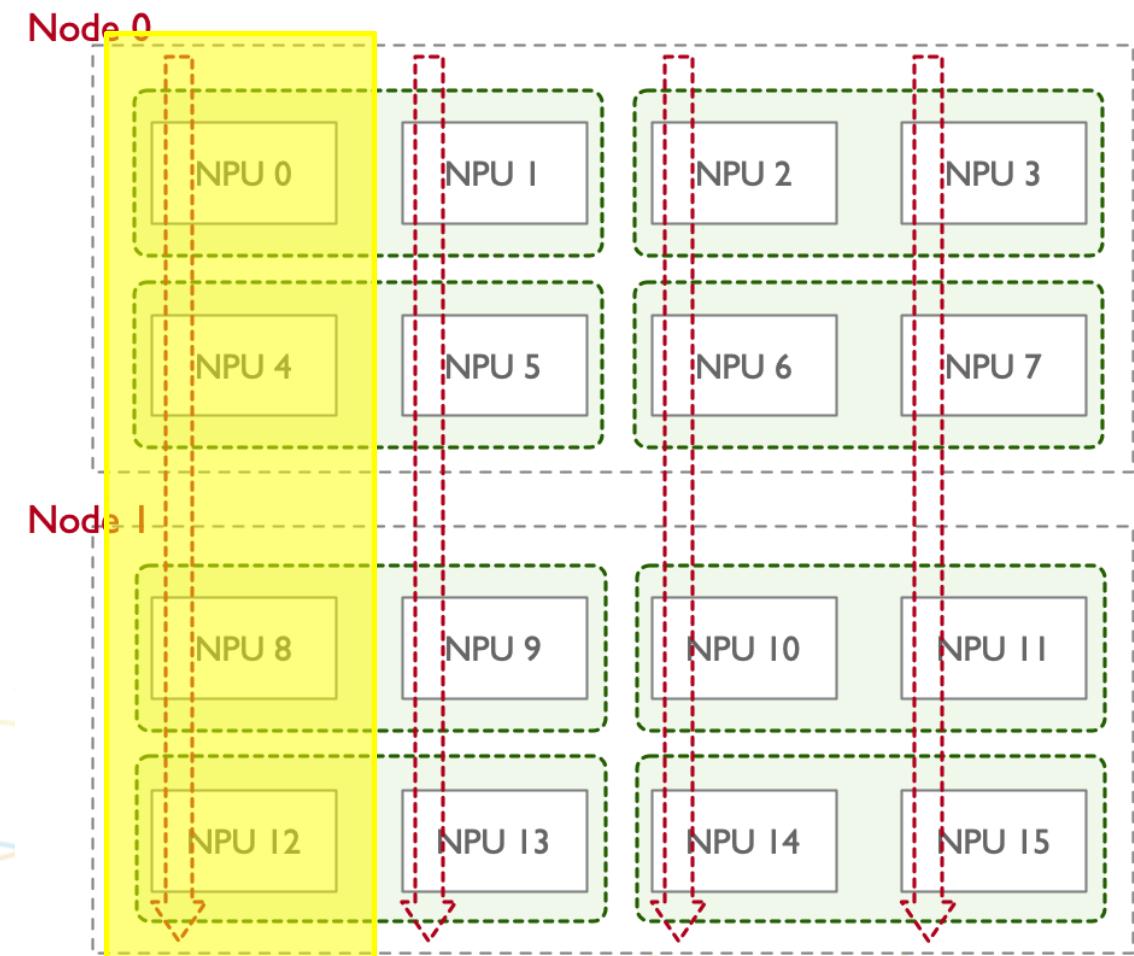
小结与思考

TP && PP 通信域

- Tensor Parallel



- Pipeline Parallel



思考

1. 进程 (process) 与进程组 (group) 关系:

- 每个进程客观上唯一（一个进程对应一个 Process ID）；
- 同一个进程可以属于多个进程组（每个进程在不同进程组中有个各自 Rank ID）；
- 一个通信域对于一个 SoC AI 芯片的计算进程会维护一个 Rank；

2. 多 Die 合封后，Rank 和 Group 到底是一个芯片一个 Rank 还是一个 Die 一个 Rank？



总结

- **AI 计算互联演变中，主要有三大趋势：**
 1. Die 间，多芯粒互联技术和合封技术正加速崛起。
 2. 片间，由 PCIe 向多节点无损网络演进；
 3. 集群间，互联方式从 TCP/IP 向 RDMA 架构转变；





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 © 2023 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.



Course chenzomi12.github.io

GitHub github.com/chenzomi12/DeepLearningSystem