通信设备行业深度报告

光通信跟踪深度：以太网在 AI 算力投资中的 Why、How 与 What

增持（维持）

关键词：#新产品、新技术、新客户

投资要点

- **产业链头号玩家纷纷布局，AI+以太网是确定趋势。**IB 网络在 AI 算力建设前期占据主流，但产业链一直在探索以太网适配 AI 计算的可能性，超以太网联盟（Ultra Ethernet Consortium，UEC）应运而生，博通、英伟达、Meta 等产业链各环节的网络、系统、云计算厂商也持续投入并取得进展，RoCE 有望逐渐取代 IB 的主流地位。
- **为什么用以太网&RoCE？**RDMA 相比传统 TCP/IP 技术更符合 AI 计算高并发、低延迟的要求，是网络技术更优选。IB 和以太网均可支持 RDMA，IB 天然支持 RDMA，是 AI 算力建设初期短时间内快速、保质、保量实现算力落地的局部最优解，以太网产业应用基础深厚、成本低，有望成为后续最优解。
- **以太网如何实现 AI 互联高要求？**AI 互联主要面对两大问题：1）“大象流”显著增加带来的长尾效应——可通过 RoCE 的自适应路由功能解决；2）不同计算进程间数据共接收端导致“多传一”拥塞——可通过 RoCE 的交换机拥塞控制算法+缓存池化解决。
- **以太网带来哪些产业变化？**1）交换机容量提升，并增加自适应路由、拥塞控制等 RoCE 配套功能，同时更加丰富的软硬件也为白盒交换机提供更大发挥空间，有利于其进一步渗透；2）推理需求增长开启叠加 RoCE 到位，云厂加速自建推理算力带来 800G 光模块新增量，同时英伟达客户结构持续优化，基于训练、训推一体上的优势引领 1.6T 等前沿产品迭代应用；3）硅光具备保证光模块供给、承接硅基共封装趋势、降低成本三层产业逻辑，有望加速渗透。
- **投资建议：**我们认为 RoCE 的渗透将有效刺激 AI 算力互联产业链的需求增长、产品技术迭代，行业头部厂商有望以更稳固的份额保持出货增长，国产芯片厂商有望实现技术、客户突破，推荐产业各环节领军者【中际旭创】及【天孚通信】，建议关注【新易盛】、【源杰科技】、【盛科通信】。
- **风险提示：**下游需求不及预期；客户开拓与份额不及预期；产品研发量产不及预期；行业竞争加剧。

表 1：重点公司估值

代码	公司	总市值（亿元）	收盘价（元）	EPS（元）			PE（倍）			投资评级
				2023A	2024E	2025E	2023A	2024E	2025E	
300308	中际旭创	1,636.90	146.00	1.94	6.14	7.41	75.31	23.78	19.70	买入
300394	天孚通信	563.94	101.81	1.32	3.52	5.58	77.26	28.92	18.25	买入

数据来源：Wind，东吴证券研究所，截至 2024 年 6 月 19 日收盘

2024 年 06 月 20 日

证券分析师 张良卫
执业证书：S0600516070001
021-60199793
zhanglw@dwzq.com.cn
研究助理 李博韦
执业证书：S0600123070070
libw@dwzq.com.cn

行业走势

相关研究

《光模块跟踪点评之（四）：英伟达发布新产品—催生算力带宽需求，为 AI 及传统推理提供更优选择》

2024-03-19

《光通信：美股 AI 科技公司密集新高后，如何看光模块板块投资逻辑？》

2024-01-22

内容目录

图表目录

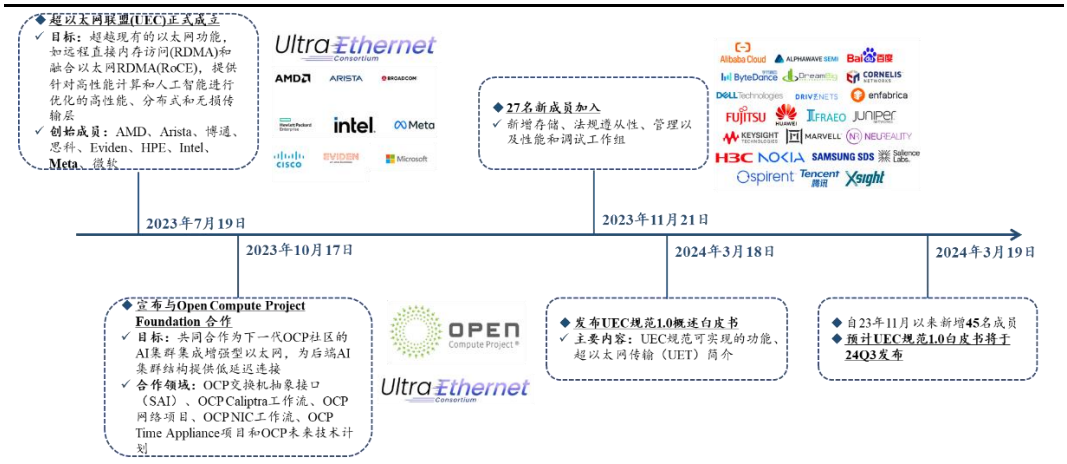
1. 为什么是 RoCE 取代 IB?

1.1. 产业链巨头相继入局，RoCE 有望取代 IB

IB 网络在 AI 算力建设前期占据主流，但产业链一直在探索以太网适配 AI 计算的可能性，超以太网联盟（Ultra Ethernet Consortium，UEC）应运而生，博通、英伟达、Meta 等产业链各环节的网络、系统、云计算厂商也持续投入并取得进展，RoCE 有望逐渐取代 IB 的主流地位。

在传统以太网上延展，超以太网联盟聚集头部玩家。超以太网联盟（UEC）于 2023 年 7 月 19 日成立，由 Linux 基金会及其联合开发基金会倡议主办，目标是超越现有的以太网功能，以 RDMA 和 RoCE 等提供面向 HPC 和 AI 计算的高性能、分布式和无损传输层，其初创成员包括 AMD、Arista、博通、思科、Eviden、HPE、Intel、Meta 和微软。截至 2024 年 3 月 19 日，UEC 目前已新增 45 名新成员，并已发布 UEC 规范 1.0 概述白皮书，简述了 UEC 规范可实现八大功能和超以太网传输（UET）的性能优势，预计 UEC 正式规范 1.0 白皮书将在 24Q3 发布。

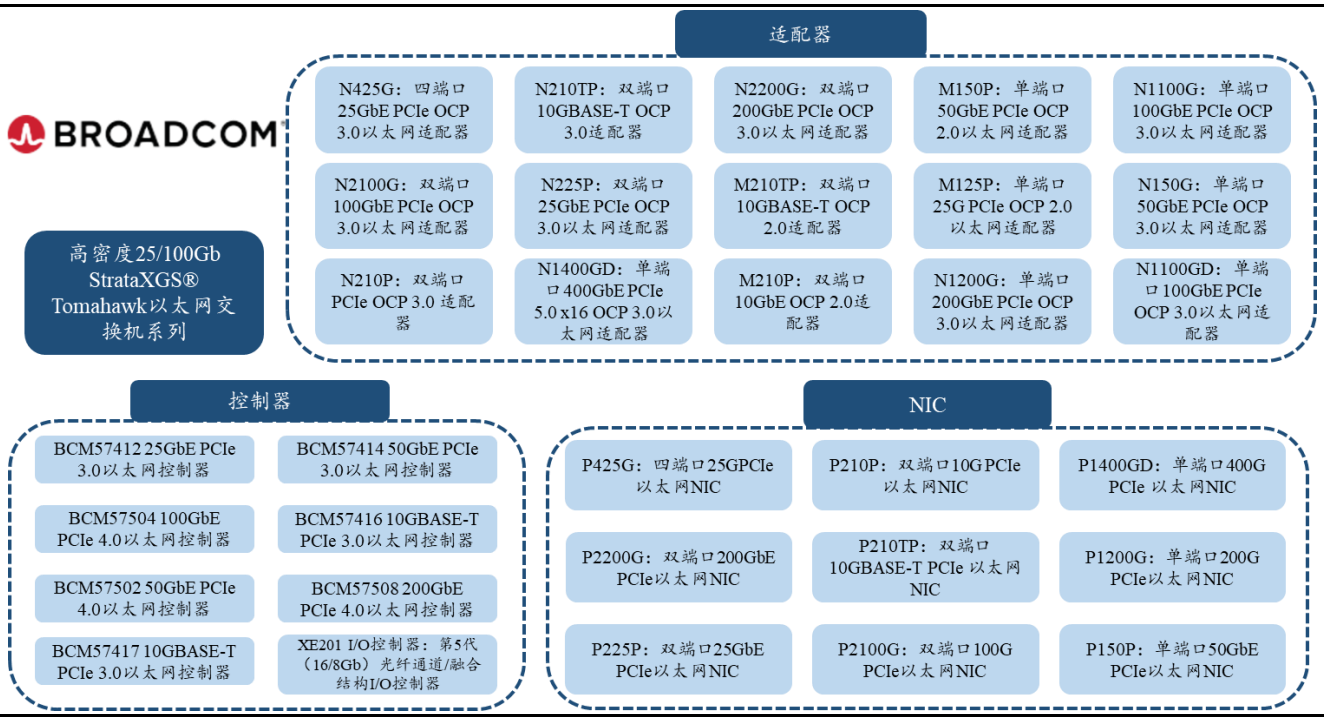
图1：超以太网联盟历史沿革



数据来源：公司官网，东吴证券研究所

博通持续丰富产品线，积极布局 RoCE 领域。博通是全球领先的有线和无线通信半导体公司，目前已在行业深耕 60 余年，拥有深厚的技术积累与丰富的产品组合。在 RoCE 领域，公司从控制器、适配器、NIC、交换机四方面入手，目前已有超 30 种相关产品，近期博通基于第四代 RoCE 推出单端口 400GbE 以太网适配器 N1400GD 和单端口 400G PCIe 以太网 NIC P1400GD，主要应用于 AI、云计算、高性能计算和存储的网络构建。

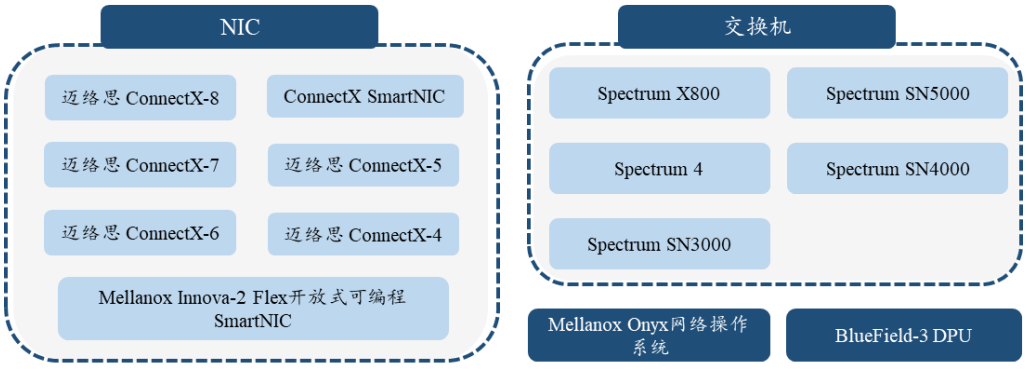
图2：博通 RoCE 领域产品布局



数据来源：公司官网，东吴证券研究所

英伟达积极拥抱 RoCE，引领 AI 计算新风向。英伟达在 NIC 和交换机方向进行布局，尽管此前英伟达是 InfiniBand 的主要推动者及供应商，但也持续在 RoCE 方向布局，陆续推出 Spectrum SN4000 和 Spectrum SN5000 交换机则，并于今年推出与 IB 新产品同规格的 Spectrum X800 交换机，同时计划于 2025 年推出 512 端口的 Spectrum Ultra X800 交换机，于 2026 年推出带宽相比 X800 翻倍的 X1600.。

图3：英伟达 RoCE 领域产品布局

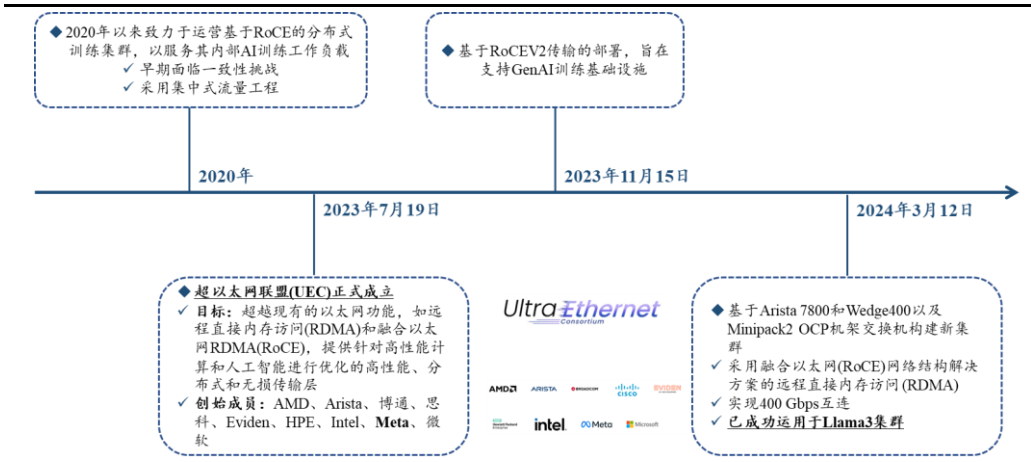


数据来源：公司官网，东吴证券研究所

Meta 布局多年，RoCE 成功应用于 Llama3 集群。2020 年以来，Meta 始终致力于运营基于 RoCE 的分布式训练集群，但早期面临一致性挑战。为实现 RoCE 的 AI 计算应用落地，Meta 作为创始成员成立超以太网联盟，并积极推进 RoCE 的部署。公司使用

Arista 7800 和 Wedge 400 等组成的 RoCE 网络能够实现 400G 互连，现已成功运用于 Llama3 集群。

图4：Meta RoCE 领域布局

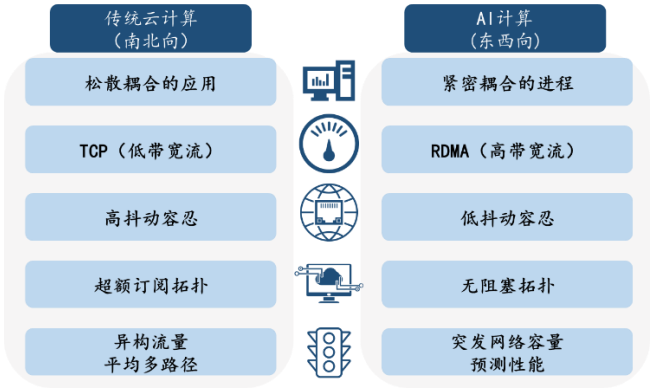


数据来源：SDNLAB，公司官网，东吴证券研究所

1.2. 为什么之前是 IB，现在是 RoCE？

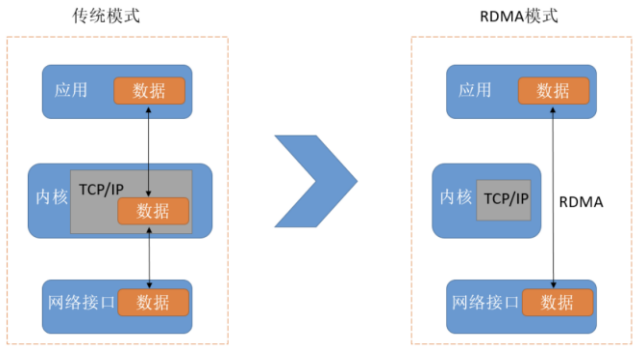
RDMA 相比传统 TCP/IP 技术更符合 AI 计算高并发、低延迟的要求，是更优选。和之前的 TCP/IP 软硬件架构相比，RDMA 使得通信系统直接通过网卡访问 GPU 显存数据，流程无需经过操作系统或 CPU，这种高吞吐、低延迟的网络通信非常适合在大规模并行 AI 计算集群中使用。

图5：传统云计算和 AI 计算部分特性对比



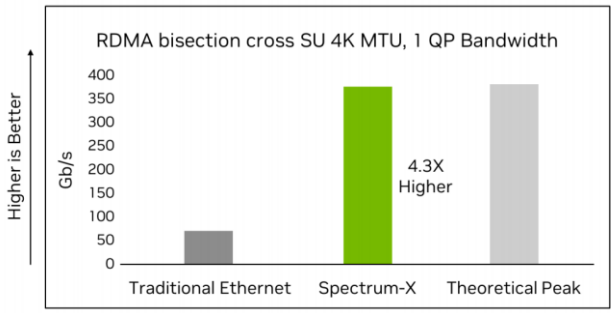
数据来源：英伟达，东吴证券研究所

图6：RDMA 和传统 TCP/Ip 实现方式比较



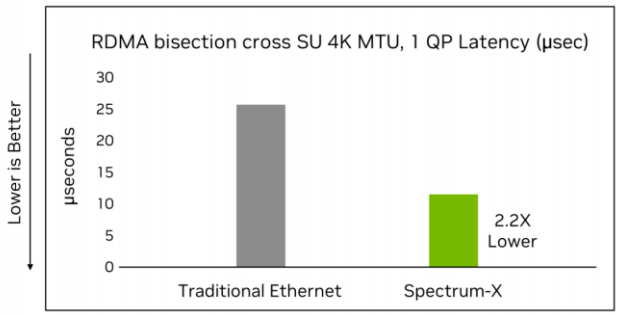
数据来源：华为，东吴证券研究所

图7：RDMA 相比传统以太网有更高的实际带宽



数据来源：英伟达，东吴证券研究所

图8：RDMA 相比传统以太网有更低的实际延迟



数据来源：英伟达，东吴证券研究所

目前支持RDMA的网络有Infiniband、RoCE(RDMA over Converged Ethernet)、iWARP，各类网络特性各异：

- **Infiniband：**专为 RDMA 设计，从硬件级别保证可靠传输，应用效果好，无需做针对性的设计研发但是需要 IB 网卡和交换机支持，成本高昂
- **RoCE：**基于以太网和传输层 UDP 协议设计，消耗的资源更少，可以使用普通的以太网交换机，但需要专门支持 RoCE 的网卡。
- **iWARP：**基于以太网传输层 TCP 协议，利用 TCP 达到可靠传输。相比 RoCE，在大型组网的情况下，iWARP 的大量 TCP 连接会占用大量的内存资源（RoCE 的 UDP 连接不需要），对系统规格要求更高。可以使用普通的以太网交换机，但需要专门支持 iWARP 的网卡。

图9：三类 RDMA 网络对比

维度	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
云计算应用基础	中	好	好
RDMA适配性	天然适配	基于以太网额外开发	基于以太网额外开发

数据来源：华为，东吴证券研究所

在 AI 算力建设浪潮中，IB 是早期局部最优解，RoCE 是更广泛最优解。在 AI 算力建设加速之初，高吞吐、低延迟的网络要求需要支持 RDMA 的网络通信，从英伟达 H 系列 GPU 持续性地供不应求也可以看出，短时间内快速、保质、保量实现算力落地是各算力投资方的核心诉求，因此英伟达的 GPU 加上天然适配 RDMA 的 IB 网络架构是当时的最优解。

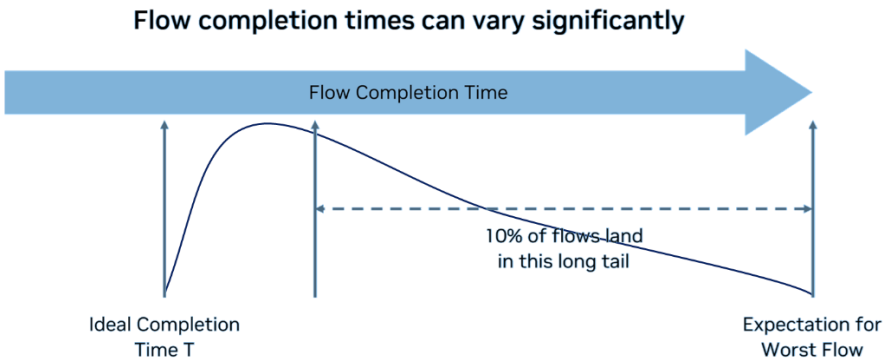
长期看，以太网/RoCE 相比 IB 在云计算领域有更深厚的产业应用基础，实现成本也更低，随着技术越来越成熟，且推理需求逐渐起势，以太网将逐步将来到 AI 算力舞台中心。

2. RDMA 如何在技术上满足 AI 计算的互联要求？

2.1. AI 计算面临的潜在互联问题

“大象流”显著增多，拥塞与长尾效应更明显。传统云计算及相应算法产生的数据流基本为占用内存小、波动范围小的流量，因此虽然网络为非全局路由，按照既定策略为流量分配路径也不会过多出现拥塞；AI 计算产生的数据流中大象流（Elephant Flow）显著增加，对于少数被分配较多大象流的路径，其传输时间将显著高于大部分路径，这就会产生“长尾效应”，大部分路径传输完成后闲置等待少数路径完成传输，系统利用率因此打折扣。

图10：AI 计算的数据流传输容易出现长尾相应

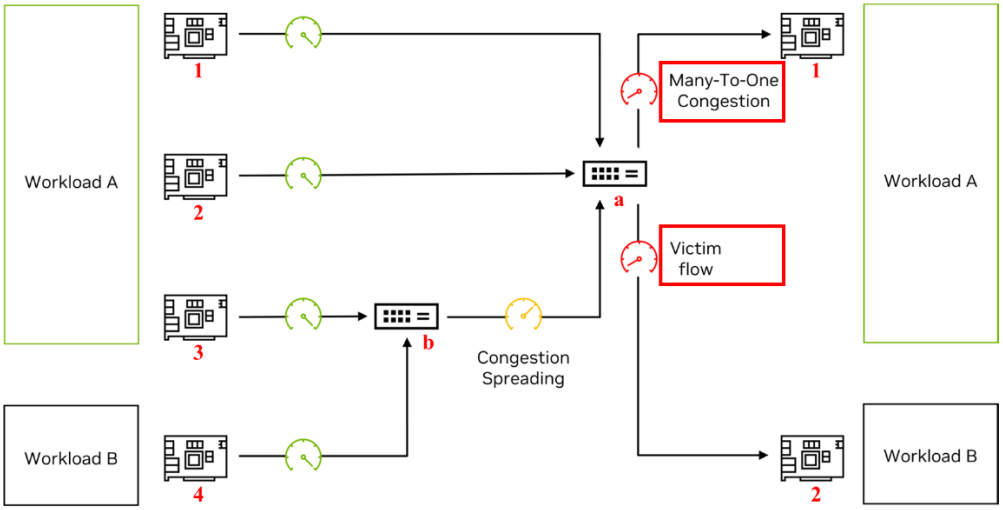


数据来源：英伟达，东吴证券研究所

不同计算进程间数据共接收端，容易出现“受害者流量”。AI 推理集群必然会出现多个负载处理多个用户需求或多条并发请求的情况，不同负载由不同端口输出数据，传输路径上有共用的叶、脊交换机，则共接收端的“多传一”（Many-To-One）现象容易出现网络背压、拥塞传播甚至丢包。

例如下图中，负载 A 由网卡 1、2、3 输出的路径与负载 B 由网卡 4 输出的路径共用交换机 a，且路径 3 与路径 4 共用交换机 b，在常规网络架构下，路径 1、2、3 均按最大带宽连接交换机 a，交换机 a 处出现拥塞，网路背压导致连接交换机 b 的路径也出现拥塞，路径 4 数据流的稳态带宽受到影响，成为“受害者流量”（Victim Flow）。

图11：共接收端“多传一”带来拥塞

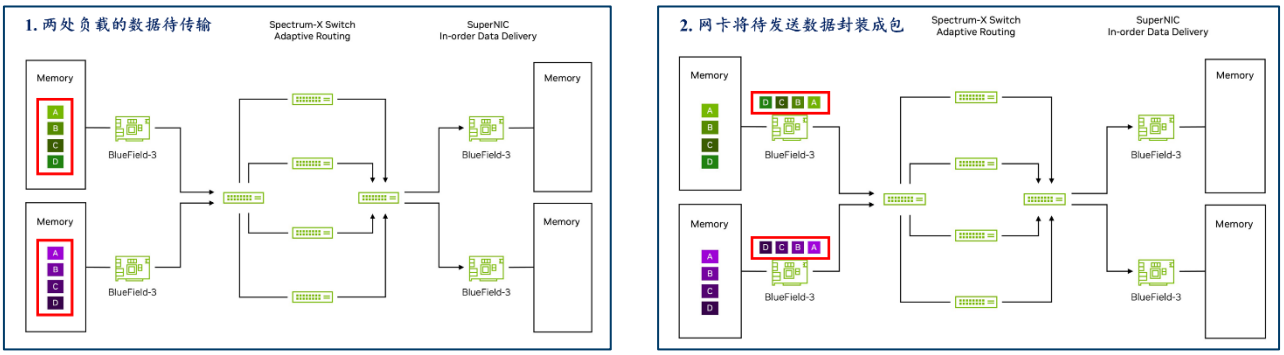


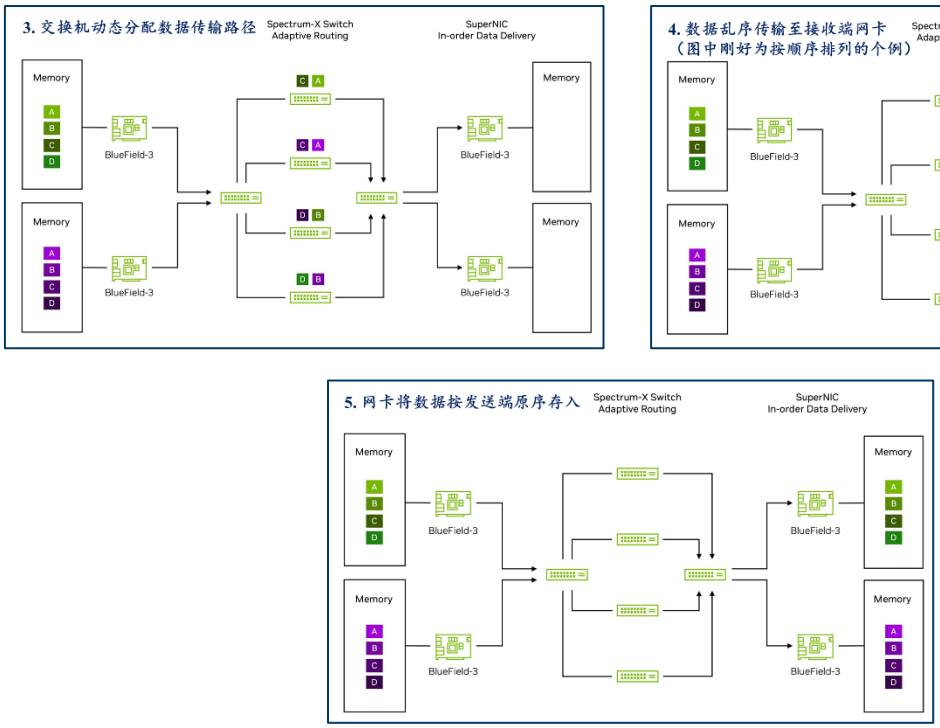
数据来源：英伟达，东吴证券研究所

2.2. RDMA 网络如何解决潜在问题？

“自适应路由”基于网卡及交换机，可解决“大象流”带来的长尾效应。1）交换机根据各端口数据输出队列状态判断该端口的负荷情况，并将新数据路由至当前负荷最小的端口/路径，这样可有效实现各端口负载均衡；**2）**重新路由后的数据一般会按照与原序列不同的顺序到达网卡，网卡利用 DDP 协议（数据报文中的 DDP 前缀包含识别数据原存储位置的信息）将接收到的数据按照原顺序存放。针对 AI 计算中显著增加的“大象流”，自适应路由通过动态监控各端口传输负荷并按此分配路径，均衡负载，解决长尾问题。

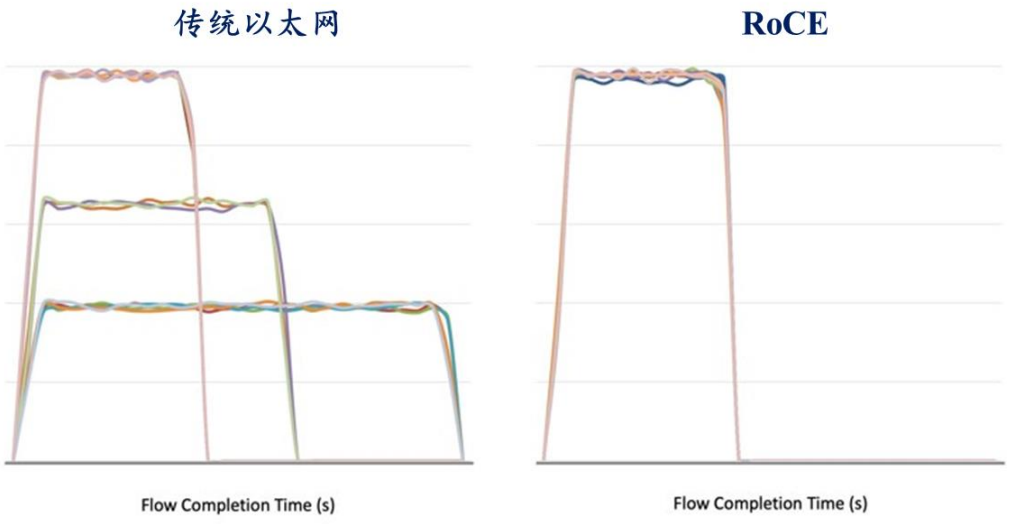
图12：自适应路由原理图





数据来源：英伟达，东吴证券研究所

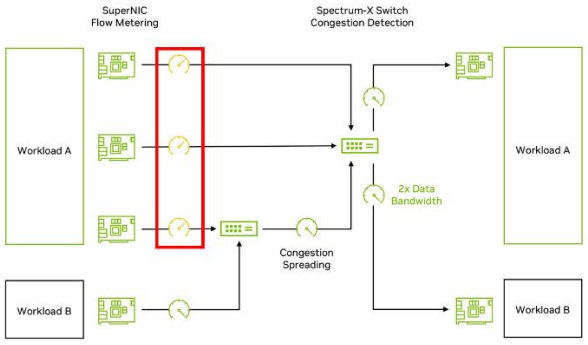
图13：RoCE 通过自适应路由减少“长尾效应”的效果明显



数据来源：英伟达，东吴证券研究所

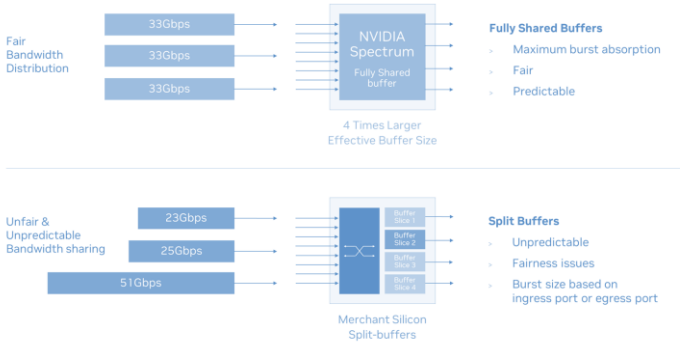
交换机拥塞控制算法+缓存池化实现性能隔离。1）各节点交换机实时监控传输速率及拥塞程度，由交换机芯片接收处理该节点及相邻节点的检测数据，并基于拥塞控制算法调节各相关交换机的传输速率；**2）**交换机将物理缓存池化，根据不同端口的接收、传输速率分配缓存。

图14：拥塞控制算法调节相关节点交换机速率



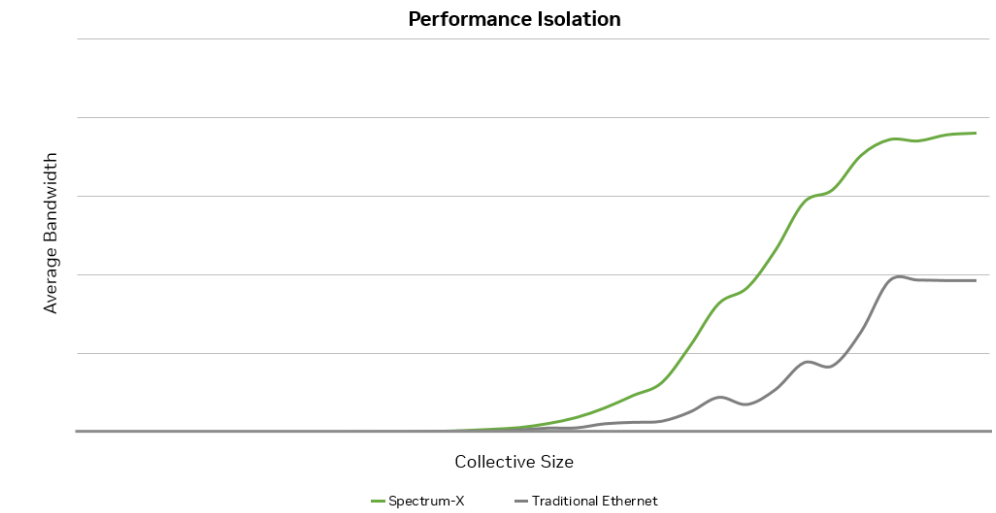
数据来源：英伟达，东吴证券研究所

图15：交换机缓存池化



数据来源：英伟达，东吴证券研究所

图16：Spectrum-X 进行拥塞控制的网络平均带宽是传统以太网两倍



数据来源：英伟达，东吴证券研究所

3. RoCE 的渗透将带来哪些产业变化？

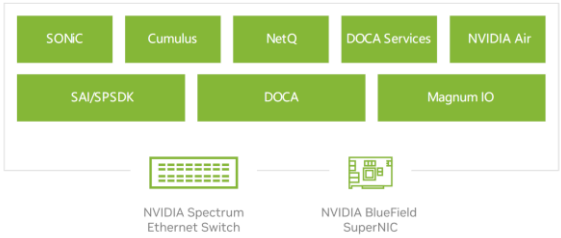
3.1. 交换机集成更多功能，白盒交换机获更多发挥空间

芯片支持容量提升，增加 RoCE 配套功能。交换机芯片支持的容量迭代提升是必然趋势，博通 Tomahawk 5 总容量达 51.2T，支持 64 个端口单口带宽达 800G，相比上代翻倍，英伟达 Spectrum-X800 交换机总容量 51.2T、端口 64 个，分别是上一代的 4 倍和两倍；同时前一章中提到 RoCE 实现的自适应路由、拥塞控制及缓存池化分配等功能均需要交换机、网卡软硬件支持。

图17：博通 TH 系列路线图

数据来源：飞速，东吴证券研究所

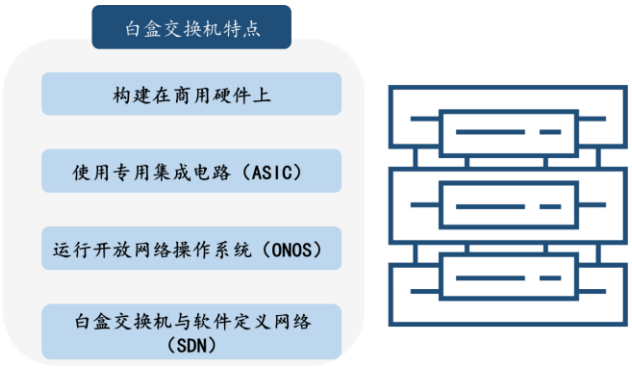
图18：Spectrum-X800 适配的软件



数据来源：英伟达，东吴证券研究所

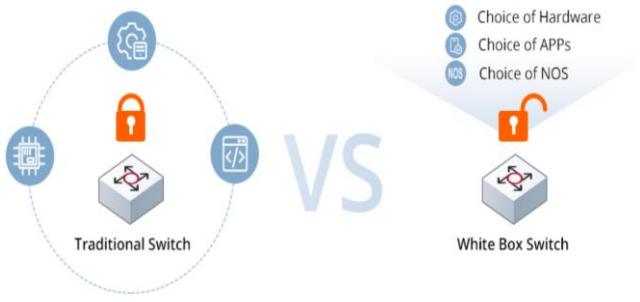
RoCE 带来更多软件客制化可能，白盒交换机有望进一步渗透。白盒交换机采用开放式网络交换架构，将商用硬件与开源软件操作系统相结合，以实现更灵活的网络配置和管理。RoCE 网络中的硬件升级以实现自适应路由、拥塞控制等功能，同时云厂商亦可根据自身硬件特性、需求和痛点自行开发相应功能的算法及软件，白盒交换机在软硬件上的发挥空间进一步扩展。

图19：白盒交换机自身特点



数据来源：华为，东吴证券研究所

图20：传统交换机和白盒交换机架构对比



数据来源：SDNLAB，东吴证券研究所

3.2. 英伟达及云厂商一前一后拉动光模块需求

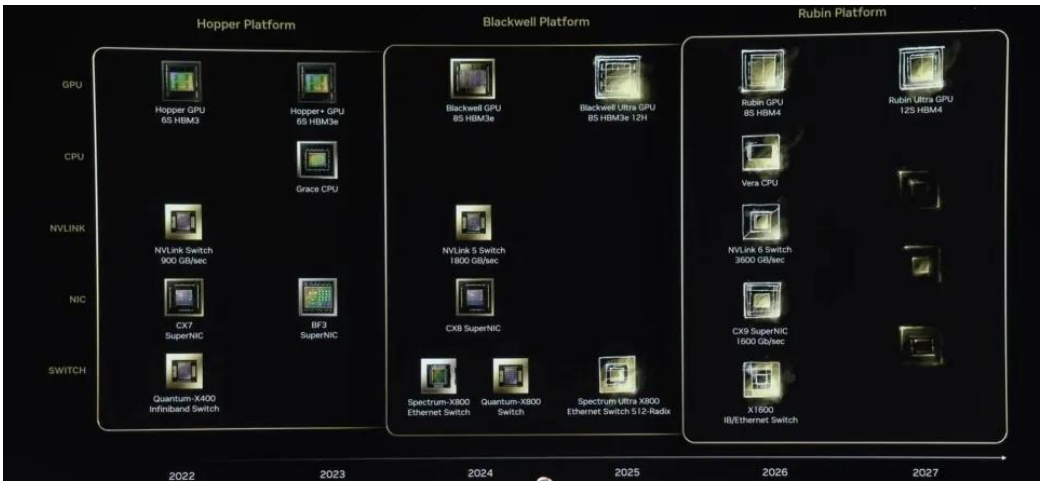
推理需求增长+RoCE 技术到位，云厂自建加速带来 800G 光模块新增量。各大云厂均有深厚的以太网算力集群投资建设经验，在之前的算力建设阶段，训练及训推一体算力投资占比高，基本以 IB 为主，以太网均为 400G 架构，后期推理需求陆续释放，且自建 RoCE 可实现 800G 带宽，云厂有望以 IB 会以更大力度投资建设，800G 光模块等产品需求得以加速增长

云厂不构成对英伟达总需求的分流:

一方面，英伟达的客户结构越来越多样化，大型云厂客户在公司数据中心业务收入占比由上个季度的 50%以上降低至 45%左右，同时后续主权 AI、企业计算、自动驾驶等垂类均将增长至贡献数十亿美元收入，客户、需求结构还会进一步优化;

另一方面，对于上游光模块供应链而言，目前英伟达更多扮演前沿产品迭代引领者及新产品需求推动者的角色，英伟达 IB 架构的交换机容量及支持带宽相对最新 RoCE 产品保持一代领先，我们认为至少在后续一到两代产品中，英伟达主导训练及训推算力，云厂自建算力以推理为主，如明年英伟达出货 GB200 搭配 1.6T 光模块，以训练、训推一体为主，云厂商采用 H100 等搭配 800G 光模块自建，以推理为主。因此当前节点针对英伟达的需求应更多关注明年上量的 1.6T 光模块。

图21：英伟达目前规划的产品路线图



数据来源：英伟达，东吴证券研究所

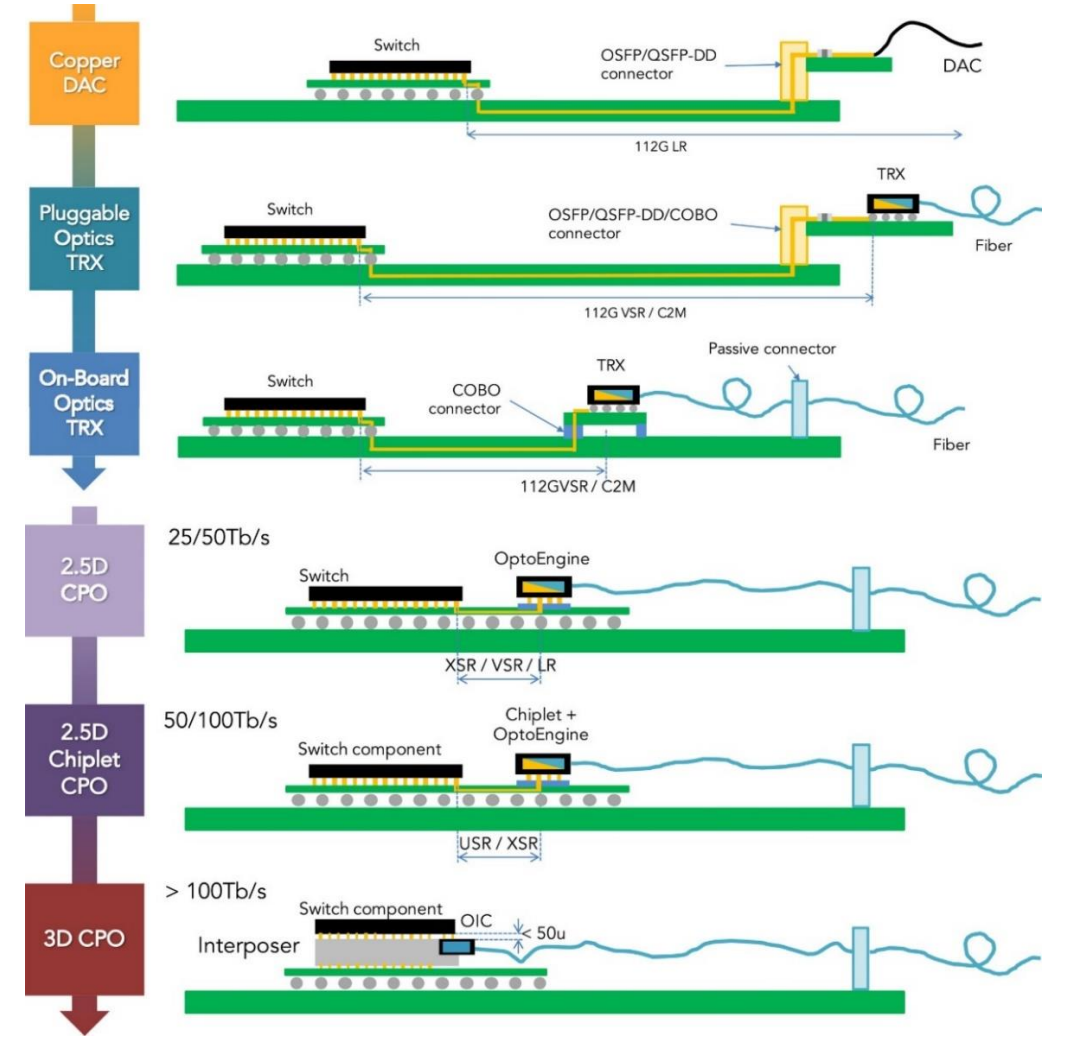
产业逻辑充分，硅光模块有望加速渗透:

1、 补上 EML 芯片潜在缺口。云厂商采用的为单模光模块，明年 EML 芯片的实际产能仍不确定，出于保证供给考虑，硅光模块的验证有望在今年开展并逐渐落地，并在明年渗透;

2、 向潜在技术路线靠拢，为硅基 2D、2.5D、3D 光电共封装布局；

3、 成本更低。

图22：共封装方案演进路线图



数据来源：IET Optoelectronics，东吴证券研究所

4. 投资建议

我们认为 RoCE 的渗透将有效刺激 AI 算力互联产业链的需求增长、产品技术迭代，行业头部厂商有望以更稳固的份额保持出货增长，国产芯片厂商有望实现技术产品、客户突破，推荐产业各环节领军者【中际旭创】及【天孚通信】，建议关注【新易盛】、【源杰科技】、【盛科通信】。

5. 风险提示

下游需求不及预期：若后续下游客户算力建设投入未达预期，或 RoCE 发展、渗透未及预期，各客户对于高速光模块的需求也将不及预期，公司业绩表现将收到影响；

客户开拓与份额不及预期：如果公司未如预期开拓潜在客户，或在客户处份额低于预期，公司业绩将受到影响；

产品研发量产不及预期：如果公司在具有潜在应用前景的硅光、1.6T、3.2T 光模块、LPO 的研发及量产应用上未达预期，将对公司业绩的长期表现造成影响；

行业竞争加剧：公司目前在全球或国内处于领先地位，如果行业竞争持续加剧，公司产品份额存在下降的可能。

东吴证券 SOOCHOW SECURITIES

免责声明

东吴证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本研究报告仅供东吴证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议，本公司及作者不对任何人因使用本报告中的内容所导致的任何后果负任何责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

在法律许可的情况下，东吴证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

市场有风险，投资需谨慎。本报告是基于本公司分析师认为可靠且已公开的信息，本公司力求但不保证这些信息的准确性和完整性，也不保证文中观点或陈述不会发生任何变更，在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。

东吴证券投资评级标准

投资评级基于分析师对报告发布日后 6 至 12 个月内行业或公司回报潜力相对基准表现的预期（A 股市场基准为沪深 300 指数，香港市场基准为恒生指数，美国市场基准为标普 500 指数，新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的），北交所基准指数为北证 50 指数），具体如下：

公司投资评级：

买入：预期未来 6 个月个股涨跌幅相对基准在 15%以上；

增持：预期未来 6 个月个股涨跌幅相对基准介于 5%与 15%之间；

中性：预期未来 6 个月个股涨跌幅相对基准介于-5%与 5%之间；

减持：预期未来 6 个月个股涨跌幅相对基准介于-15%与-5%之间；

卖出：预期未来 6 个月个股涨跌幅相对基准在-15%以下。

行业投资评级：

增持： 预期未来 6 个月内，行业指数相对强于基准 5%以上；

中性： 预期未来 6 个月内，行业指数相对基准-5%与 5%；

减持： 预期未来 6 个月内，行业指数相对弱于基准 5%以上。

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议。投资者买入或者卖出证券的决定应当充分考虑自身特定状况，如具体投资目的、财务状况以及特定需求等，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。

东吴证券研究所
苏州工业园区星阳街 5 号
邮政编码：215021
传真：（0512）62938527
公司网址： http://www.dwzq.com.cn