

电子行业点评报告

Switch 芯片研究框架（一）：GPU-GPU 互连，从 Scale-Up 到 Scale-Out 的格局如何？

增持（维持）

2025 年 09 月 30 日

证券分析师 陈海进

执业证书：S0600525020001

chenhj@dwzq.com.cn

研究助理 李雅文

执业证书：S0600125020002

liyw@dwzq.com.cn

投资要点

■ **Switch 芯片厂商有哪些？**总体看，NVIDIA 以私有协议占据主要市场份额，博通以基于以太网的 SUE 协议逐渐掌握开源生态话语权，Alab 则侧重 PCIe/UALink；Marvell、瑞昱也各自发展差异化技术路线。具体来看，**Scale-Up 方面**，NVSwitch 目前仍是业界带宽最高、部署最成熟的私有方案。博通在云端数据中心交换机市场市占率为九成，推出基于以太网络的 SUE 架构。ALAB 是唯一一家同时掌握交换芯片+延长线+软件管理平台的厂家，且 Scorpio X 已通过 NVIDIA Blackwell 平台验证。**Scale-Out 方面**，博通有诸如 Tomahawk 5、Jericho3-AI 等布局。Marvell 提供 51.2Tbps 的 Teralynx 10 交换芯片。瑞昱主攻消费级市场，速率涵盖百兆至万兆。以上都是国际知名的主流 Switch 芯片厂商。

■ **如何实现 Switch 芯片的国产替代？**独立交换芯片厂商方面，国内玩家主要有盛科通信、数渡科技和澜起科技，整体格局为盛科通信领跑，25.6T 交换芯片已量产并导入主流设备商供应链；数渡科技和澜起科技聚焦 PCIe，形成多点突破。具体来看，盛科通信主要以太网芯片产品覆盖 100Gbps-25.6Tbps 的交换容量以及 100M 到 800G 的端口速率，在企业网络、运营商网络、数据中心网络等领域得到了广泛应用。数渡科技的 PCIe 5.0 交换芯片领域率先进入客户导入和小批量阶段，2025 年底有望实现国产替代的突破。澜起科技以 Retimer 为切入点，正逐步构建完整的 PCIe/CXL 互联产品族。

■ **大厂自研芯片如何助力 Switch 芯片的国产替代？**大厂自研芯片方面，海光、华为、中兴和新华三等服务器产业链厂商均有布局。其中，华为、中兴和新华三已推出自研交换芯片，并配套自家交换机及超节点方案；海光与华为则通过建设开放生态协议来抢占产业生态制高点。具体看，海光信息已面向 GPU、IO、OS、OEM 等产业全栈，正式宣布开放 CPU 互联总线协议（HSL）。华为已发布单芯片 51.2T 数据中心盒式液冷交换机。中兴通讯的自研交换芯片亦达到业界一流水平。新华三也有自主研发自主研发网络交换芯片。虽然与海外龙头有一定的距离，但国内 Switch 芯片厂商也逐渐呈现百花齐放的态势。

■ **投资建议：**重点推荐海光信息、盛科通信，建议关注中兴通讯、万通发展、澜起科技等。

■ **风险提示：**AI 应用进展不及预期，技术发展不及预期，市场竞争风险。

行业走势



相关研究

《AI 驱动 PCB 全面升级：材料、工艺与架构革新引领产业新周期》

2025-09-28

《模拟反倾销调查启动，多维拆解影响几何？》

2025-09-14

表 1：重点公司估值

代码	公司	总市值 (亿元)	收盘价 (元)	EPS (元/股)			PE			投资评级
				2024A	2025E	2026E	2024A	2025E	2026E	
688041	海光信息	5,946.59	255.84	0.83	1.23	1.85	307.96	208.00	138.29	买入
688702	盛科通信-U	520.70	127.00	-0.17	-0.04	0.08	-762.77	-3,175.00	1,587.50	增持

数据来源：Wind，东吴证券研究所预测

注：收盘价截至 2025/9/29

内容目录

1. Switch 芯片厂商有哪些？	4
1.1. Scale-Up	4
1.1.1. NVSwitch: NVIDIA GPU Scale-Up 专用高速交换芯片的多代演进	4
1.1.2. 博通: 面向 AI 的开放以太网 Scale-Up 互连规范与芯片方案	4
1.1.3. Astera Lab: 基于 PCIe/UALink 双模的 Scorpio X-Series Scale-Up 交换平台	6
1.2. Scale-Out	7
1.2.1. 博通: 三大交换芯片系列覆盖全面应用场景	7
1.2.2. Marvell: 面向高端数据中心与 AI 驱动网络服务	8
2. 如何实现 Switch 芯片的国产替代？	9
2.1. 独立交换芯片厂商	9
2.2. 大厂自研交换芯片	11
3. 风险提示	14

图表目录

图 1: NVSwitch 六代演进方案.....	4
图 2: 多家公司主流芯片产品性能对比.....	4
图 3: 专为 Scale-Up 设计的博通 SUE 框架, 竞争对手为 UALink.....	5
图 4: 博通 Tomahawk Ultra 与 Tomahawk 6 性能对比.....	5
图 5: “芯片+延长线+软件”三位一体示意图	6
图 6: UALink 的 Scale-Up 组网方式	7
图 7: 博通三大交换芯片系列概况.....	7
图 8: 博通交换芯片技术路线图 (部分)	8
图 9: Marvell 交换芯片技术路线图 (部分)	9
图 10: 盛科通信主要以太网交换芯片产品系列.....	10
图 11: 盛科通信 2020-2025H1 研发投入与交换芯片业务营业收入	10
图 12: 数渡科技产品矩阵.....	11
图 13: 海光信息产品矩阵.....	11
图 14: 华为 2025 全连接大会产品路线图.....	12
图 15: 中兴通讯交换芯片与 NP 芯片技术路线图	13
图 16: 新华三智擎 660 网络处理芯片结构图.....	13

1. Switch 芯片厂商有哪些？

1.1. Scale-Up

1.1.1. NVSwitch: NVIDIA GPU Scale-Up 专用高速交换芯片的多代演进

NVSwitch 是 NVIDIA 专为 AI 服务器 Scale-Up 场景设计的 GPU-to-GPU 高速交换芯片，经历了 Pascal→Volta→Ampere→Hopper→Blackwell→Rubin 的六代演进，目前仍是业界带宽最高、部署最成熟的私有方案。

图1: NVSwitch 六代演进方案

GPU架构	首发年份	NVLink 代数	NVSwitch 芯片代数	首发系统	单芯片双向带宽
Pascal (P100)	2016	NVLink 1.0	——	DGX-1	160 GB/s
Volta (V100)	2017	NVLink 2.0	1st Gen NVSwitch	DGX-2	300 GB/s
Ampere (A100)	2020	NVLink 3.0	2nd Gen NVSwitch	DGX A100	600 GB/s
Hopper (H100)	2022	NVLink 4.0	3rd Gen NVSwitch	DGX H100	900 GB/s
Blackwell (B200)	2024	NVLink 5.0	4th Gen NVSwitch	GB200 NVL72	1.8 TB/s
Rubin (R100)	2026 (规划)	NVLink 6.0	5th Gen NVSwitch	Rubin-Next	未公布

数据来源：公司官网，东吴证券研究所

图2: 多家公司主流芯片产品性能对比

特性	NVIDIA NVL72	NVIDIA NVL576	Astera Labs Scorpio X-Series	Broadcom Tomahawk Ultra	Broadcom Tomahawk 6
端口速率	57.6 Tb/s	基于NVL72架构多层互联	——	51.2 Tbps	102.4 Tbps
制程	TSMC 4NP (5nm)	TSMC N3P (3nm)	——	TSMC 5nm	TSMC 3nm
协议支持	NVLink 5.0 (机柜内) InfiniBand (跨机柜) PCIe 6.0 (连接网卡/DPU)	NVLink (机柜内及机柜间) InfiniBand (更大规模扩展)	PCIe Gen5/Gen6 UALink over Ethernet	扩展以太网 (SUE) 传统以太网	扩展以太网 (SUE) 传统以太网
纵向扩展支持加速器数	72个B200 GPU (单机柜内全互联) 最高576个GPU	——	——	最高1024个加速器	——

数据来源：各公司官网，东吴证券研究所

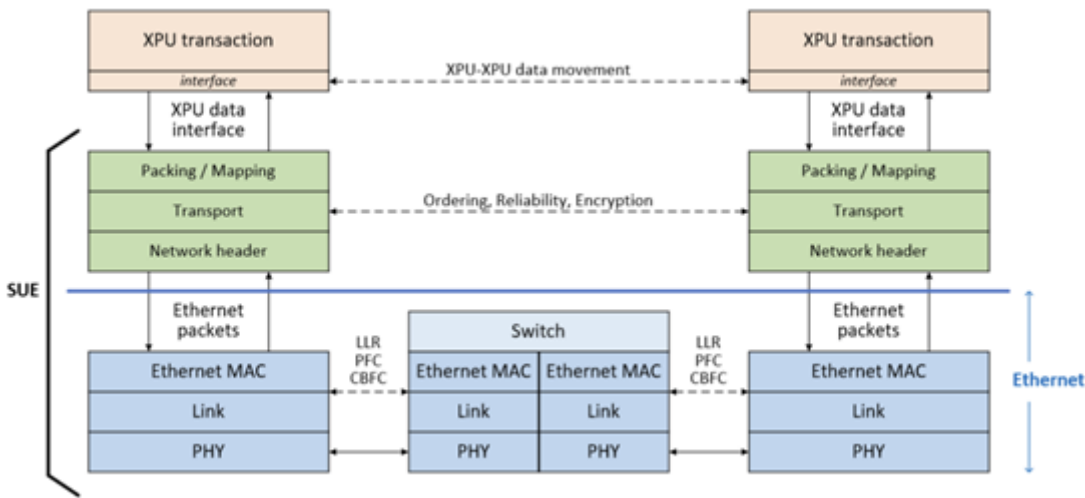
1.1.2. 博通：面向 AI 的开放以太网 Scale-Up 互连规范与芯片方案

博通凭九成市占率称霸云端数据中心交换器市场，推出基于以太网络的 SUE 架构，主打开放相容与低延迟性能，力守 AI 时代交换器主导权。51.2Tbps 的 Tomahawk Ultra 与 102.4 Tbps 的 Tomahawk 6 共同构成了统一以太网架构的基础：支持面向 AI 的纵向扩展以太网，以及面向 HPC 和分布式工作负载的横向扩展以太网。

SUE 并非具体芯片，而是一套开放接口规范，涵盖从加速器直连交换机的协议、光

模块到控制栈。博通在 2025 年 OCP 峰会上宣布免费开放 SUE 规范，希望行业共同定义“开放的 Scale-Up 互连”。例如，SUE 规定如何使用 200G/400G 直连铜缆连接 GPU 和 Tomahawk6，以及拥塞控制、流量管理机制，使以太网在小范围内达到类似 NVLink 的效果。这与 NVIDIA 的专有 NVLink 形成对比，博通试图把 Scale-Up 互连标准化为业界共有，从而打破封闭生态。博通押注“AI 网络终将回归以太网”这一逻辑。他们通过 Tomahawk6 把交换芯片带宽提升到 NVSwitch 级别，通过 Jericho3-AI 提升以太网在 All-Reduce 等操作上的效率，通过 SUE 建立开放互连规范。这代表了 AI 基础设施领域的另一条技术路线：更加开放、供应商多元化且强调标准化互通。

图3：专为 Scale-Up 设计的博通 SUE 框架，竞争对手为 UALink



数据来源：新浪财经，东吴证券研究所

Tomahawk Ultra 是以太网第一次以专用级低延迟和千级 XPU 纵向扩展能力改写 AI fabric 游戏规则。性能对比上，Tomahawk Ultra 在部分指标上已与专用协议接近或超越。其单跳可连接的 XPU 数量为 NVLink 交换机的 3 倍以上，支持至少 1024 个加速器的纵向扩展系统，而英伟达 NVLink 技术虽宣称支持 576 个加速器，但实际部署中尚未突破 72 个 GPU 插槽规模。

以太网交换 ASIC—Tomahawk 6 是业界首颗 102.4 Tbps 带宽的交换芯片，博通将其定位为同时支持 Scale-Out 和 Scale-Up 的统一方案。Tomahawk6 采用 SUE 框架，服务器/加速卡厂商可以按 SUE 规范设计接口，与 Tomahawk 交换机高效直连，从而实现类似 NVLink 的紧耦合通信，但其仍然沿用通用以太网 PHY。

图4：博通 Tomahawk Ultra 与 Tomahawk 6 性能对比

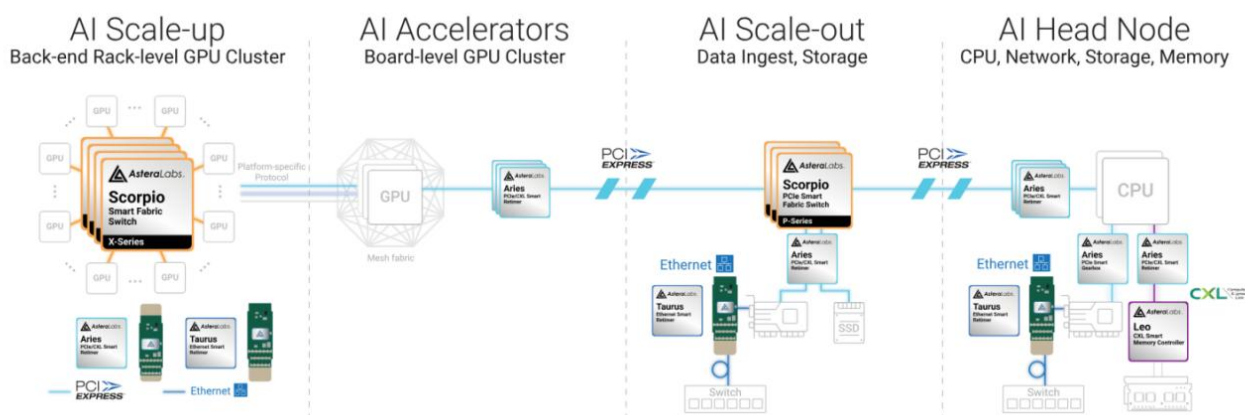
芯片	发布年份	带宽	端口速率	延迟	最大 XPU	头部开销
Tomahawk Ultra	2025年7月	51.2 Tbps	512×200G PAM4 或 1,024×100G PAM4	250 ns	≥1 024	10 B
Tomahawk 6	2025年6月	102.4 Tbps	512×200G PAM4 或 1,024×100G PAM4	—	512 (Scale-Up) 100 000+ (Scale-Out)	—

数据来源：财经头条，腾讯云，公司官网，东吴证券研究所

1.1.3. Astera Lab: 基于 PCIe/UALink 双模的 Scorpio X-Series Scale-Up 交换平台

ALAB 作为 Switch 端首批入局者，公司产品链完整，是唯一一家同时掌握交换芯片（Scorpio X-Series）+延长线（Aries SCM/AEC）+软件管理平台（COSMOS Suite）的厂家。ALAB 从 PCIe Retimer 起步，到 Scorpio X-Series 芯片放量，再到 2027 年全面拥抱 UALink 1.0（由 AMD、AWS 等推动），形成机柜内 GPU-to-GPU 内存语义交换的完整解决方案。

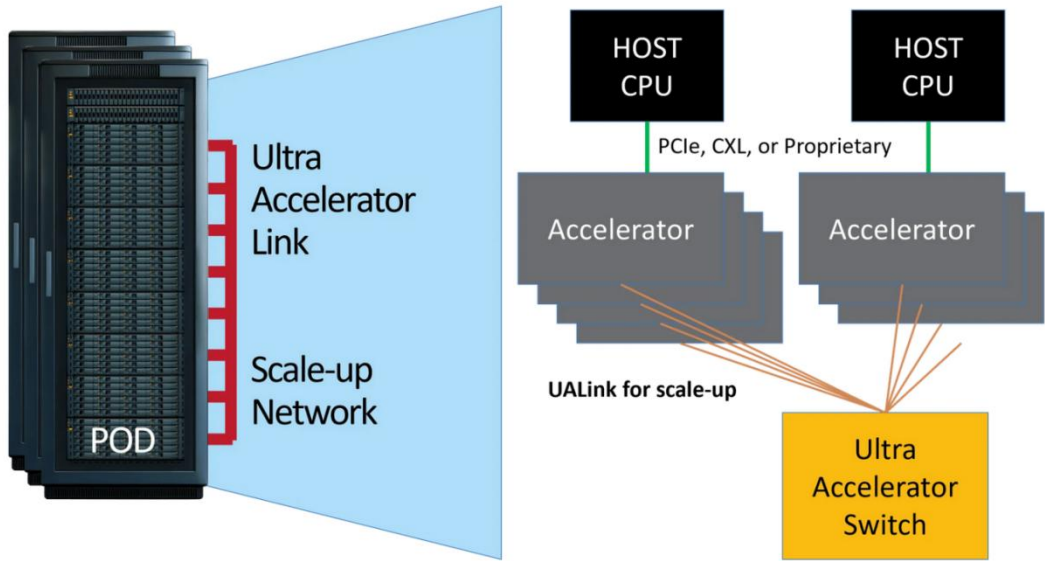
图5: “芯片+延长线+软件”三位一体示意图



数据来源：公司官网，东吴证券研究所

UALink 作为一种可扩展的 AI Fabric，可部署用于 AI 训练和 AI 推理解决方案以支持广泛的 AI 模型。UALink 是由 AMD、AWS、博通、思科、谷歌等公司组成的超级加速器链接促进会发布的开放行业互连标准。对于 AI 训练，UALink 将使可扩展域增加到数百 GPU，以满足未来 LLM 和 Transformer 模型的需求。UALink 的更高带宽和更低延迟将使大型基础模型的 GPU 训练性能继续扩展，但它也使较小的模型能够更快地训练，从而更频繁地训练。UALink 实现的高达 40% 的功耗改进，提供了将可用数据中心功率最大化用于 GPU 计算并减少训练新模型所需能量的机会。UALink 200G 1.0 规范支持每通道最高 200GT/s 的数据传输速率，能在 AI 计算舱内实现最多 1024 个加速器的每通道 200G 扩展连接，延迟低于 1 微秒，具有低功耗、以太网兼容性好、安全与管理性强等特点，可打破英伟达 NVLink 技术垄断，为下一代 AI 集群性能提供开放标准互连。随着部署的 AI 模型种类不断增加，AI 推理架构需要不断发展，以提高 AI 服务器的总体拥有成本（TCO），帮助提供商最大化其投资回报率（ROI）。UALink 带来的 TCO 效率改进将直接惠及 LLM 和推荐系统部署的 ROI。

图6: UALink 的 Scale-Up 组网方式



数据来源：公司官网，东吴证券研究所

ALab 最早布局 PCIe 领域，定义了 Redriver 和 Retimer 技术。当前，ALab 已研发 5nm PCIe Gen6 交换机，正在研发 PCIe Gen7。ALAB 凭借 PCIe 协议优势切入 UALink 生态，获得 AWS 和 AMD 支持。其双模交换机芯片（PCIe Gen7 + UALink 128G）支持 Trn 4 机群互联。实现 PCIe Gen7 与 UALink 128G 等速匹配，未来向 UALink 200G 演进。其 Scorpio X 已通过 NVIDIA Blackwell 平台验证，双方共同定义 UALink 标准，ALAB 是唯一首批硬件伙伴。ALAB 预计 UALink 和部分基于 PCIe 的 Scale-Up 部署将在 2026 年开始起量，而 UALink 将在 2027 年开始更大规模推广，管理层预计以 Scorpio X 系列为核心的 Scale-Up 业务将成为增长重点。目前的产品正在围绕 UALink 1.0 进行开发，涵盖加速器及其他设备间的互联 fabric。UALink 1.0 基于两个标准构建：上层协议采用 PCIe，物理层采用基于 200G 以太网的标准。

1.2. Scale-Out

1.2.1. 博通：三大交换芯片系列覆盖全面应用场景

博通是全球领先的商用交换芯片制造商，拥有数十年的技术积累和丰富的产品线。其在交换芯片产品上布局完善，主打三大交换芯片系列：高端的 Tomahawk、中端的 Trident 和低端的 Jericho，覆盖各种应用场景。

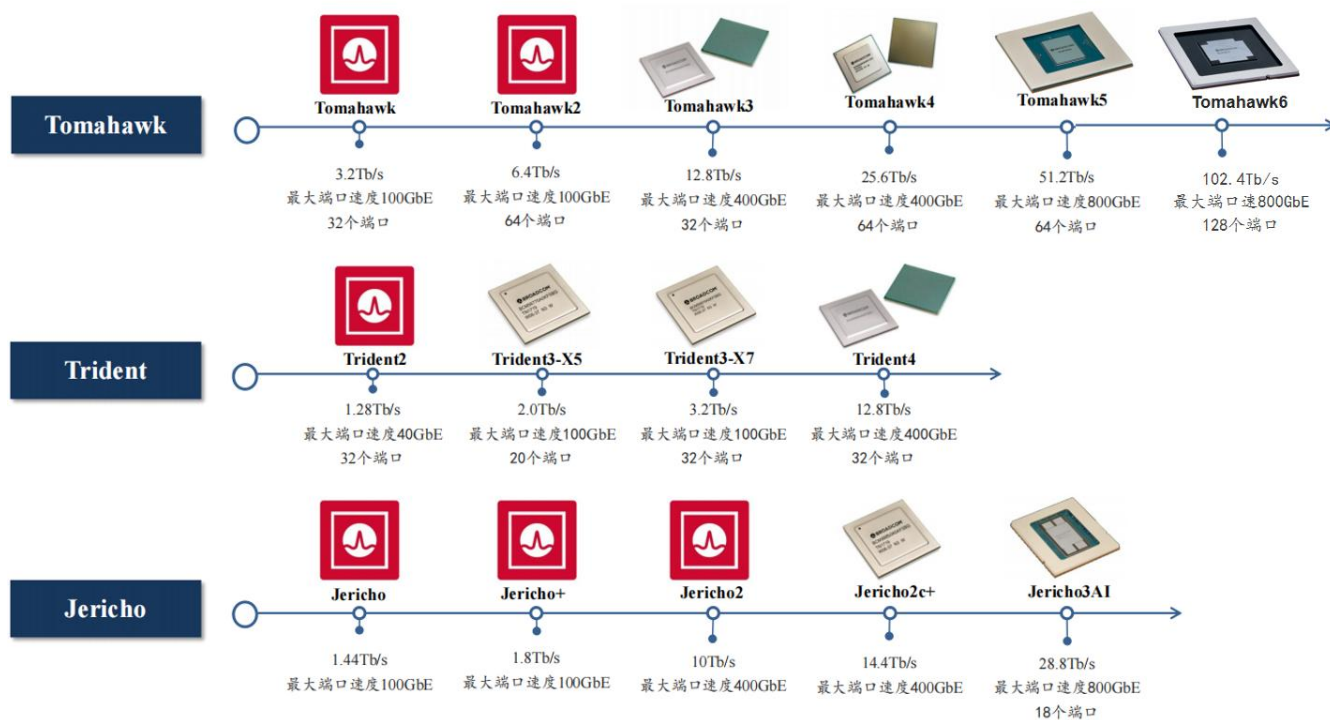
图7: 博通三大交换芯片系列概况

交换芯片系列	特点	应用场景
Tomahawk	超高带宽	带宽最高，适用于超大规模数据中心
Trident	多功能性	多功能，可编程，适用于企业和云
Jericho	可扩展性	扩展性强，可编程，高缓存，适用于服务提供商

数据来源：芯世相，东吴证券研究所

博通目前最具代表性的交换芯片是 51.2Tbps 的 Tomahawk 5，最高支持 800Gbps 的速率。针对超大规模企业和云构建者商用交换机和路由器芯片市场，2022 年 8 月博通发布了业内首款商用 51.2Tbps 容量的交换芯片 Tomahawk 5，采用 5nm 制程，由台积电代工生产。该芯片支持多种配置的交换机，包括 64 端口 800Gbps、128 端口 400Gbps 和 256 端口 200Gbps。应用场景包括数据中心、云计算、AI 和 ML 集群组网，主要面向于小规模 AI 集群数据调度。从 2010 年 Trident 的 640Gbps 增长到 2022 年 Tomahawk 5 的 51.2Tbps，Tomahawk 实现了 80 倍带宽提升，并实现了超过 90% 能耗的降低。每一代的带宽翻倍减少了所需交换机数量的同时，还降低了端口成本。如 6 个 25.6Tbps 的 Tomahawk 4 芯片互联组成的 Leaf/Spine 交换机网络，才能提供与 Tomahawk 5 芯片本身相同的端口数。同时其数据交换性能是 Tomahawk 4 芯片的两倍，能充分满足数据中心不断增长的网路需求。

图8：博通交换芯片技术路线图（部分）



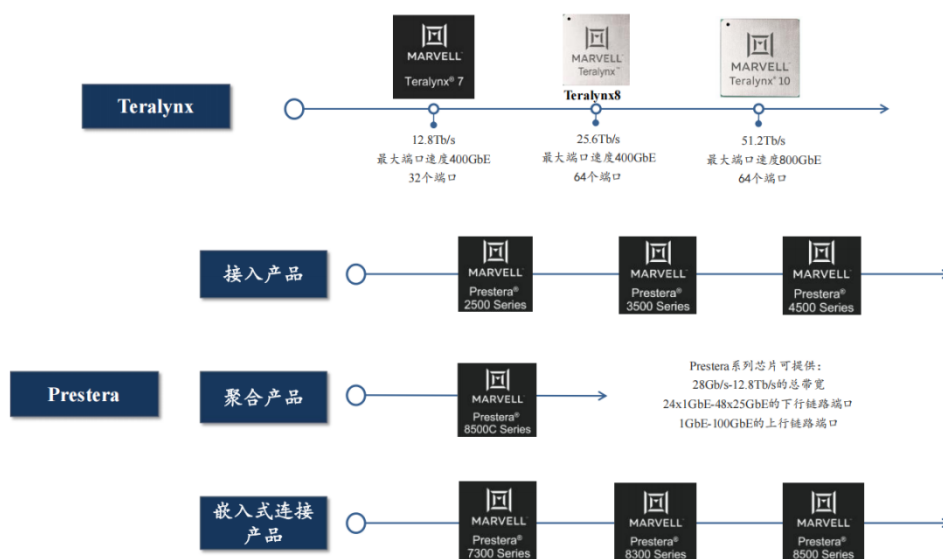
数据来源：博通官网，东吴证券研究所

1.2.2. Marvell: 面向高端数据中心与 AI 驱动网络服务

Marvell 在交换芯片市场中是高端数据中心与 AI 驱动网络的核心供应商，速率配置灵活，最高可达 800Gbps。Marvell 当前的技术能力主要来自于在 2021 年对 Innovium 的收购。完成收购后，Marvell 的交换芯片产品面向了完整的高中低市场。其中公司原本的 Prestera 系列定位中低端，主要用于企业网络和边缘市场，而 Innovium 的 Teralynx 系列产品定位中高端，主要面向数据中心和 AI 场景。Marvell 目前已实现 12.8T 交换芯片 Teralynx 7 的大规模量产，且数据中心交换的研发增加了 2.5 倍。2023 年 3 月 Marvell 推出了 51.2T 性能的 Teralynx 10 交换芯片。该可编程 5nm 交换机芯片可提供 64 端口

800Gbps 或者 512 端口 100Gbps，且延迟低至 500ns，运行功耗低至 1W/100G 端口，可适用于下一代数据中心网络中脊叶互联，以及 AI 和高性能计算（HPC）。

图9：Marvell 交换芯片技术路线图（部分）



数据来源：Marvell 官网，东吴证券研究所

2. 如何实现 Switch 芯片的国产替代？

2.1. 独立交换芯片厂商

盛科通信：国内以太网交换芯片市场的领先企业，12.8T/25.6T 交换芯片已进入客户推广阶段。公司成立于 2005 年，凭借近 20 年的研发经验，已经开发出一系列覆盖接入层到核心层的以太网交换产品。其交换芯片产品覆盖 100Gbps-25.6Tbps 的交换容量以及 100M 到 800G 的端口速率，在企业网络、运营商网络、数据中心网络和工业网络等领域得到了广泛应用。如 TsingMa.MX 系列芯片交换容量为 2.4Tbps，支持 400G 端口速率，具备智能网络可视化技术和确定性网络技术；GoldenGate 系列芯片交换容量为 1.2Tbps，支持 100G 端口速率，特有的“灵云”设计为网络虚拟化应用提供了极具竞争力的硬件加速方案。其 TsingMa 系列芯片在中低端市场具有较强优势；在高端产品上，即数据中心领域，TsingMa.MX 系列芯片及 GoldenGate 系列芯片均已导入国内主流网络设备商并实现规模量产。且 TsingMa.MX 系列芯片还供货于新华三、锐捷网络和迈普通信，切入了国内主流设备商供应链。根据公司 2025 年中报，公司面向大规模数据中心和云服务需求，交换容量为 12.8Tbps 及 25.6Tbps 的高端旗舰芯片在客户处进入市场推广和逐步应用阶段，该产品支持最大端口速率 800G，搭载增强安全互联、增强可视化和可编程等先进特性。

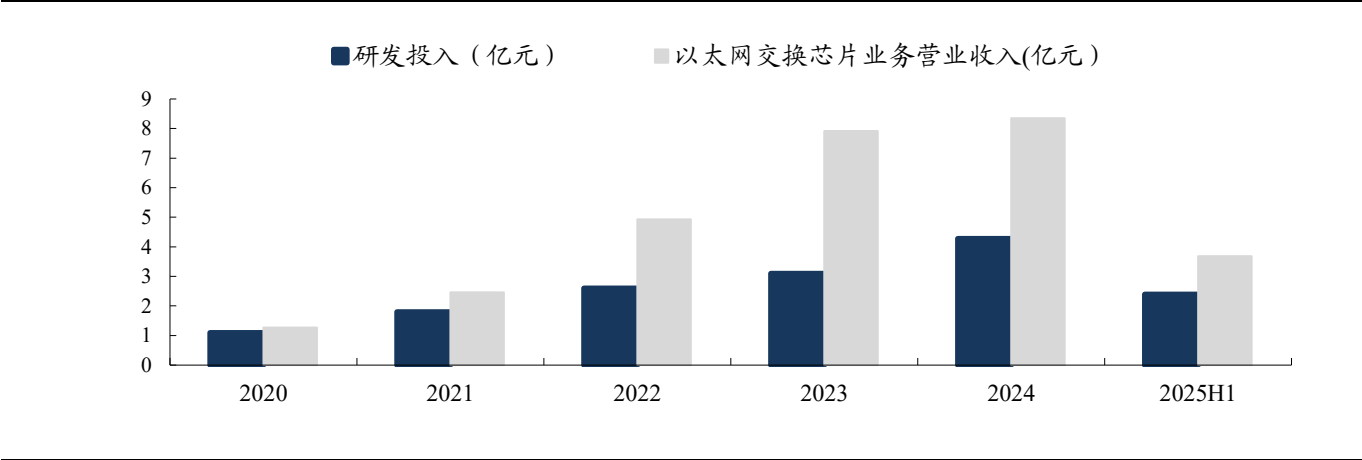
图10：盛科通信主要以太网交换芯片产品系列

交换芯片	型号	图例	交换容量	最大端口速率	应用场景
-	-	-	25.6Tbps	800G	大规模数据中心和云服务
-	-	-	12.8Tbps	800G	大规模数据中心和云服务
TsingMa.MX	CTC8180		2.4Tbps	400G	企业网和边缘计算
TsingMa	CTC7132		440Gbps	100G	企业网、边缘计算和IP RAN
TsingMa.AX	CTC2118		28Gbps	1G	工业接入、SOHO/SMB 路由器接口扩展
TsingMa.CX	CTC5118		180Gbps	10G	企业网、边缘计算和工业网
GoldenGate	CTC8096		1.2Tbps	100G	架顶（ToR）/核心交换机 企业网汇聚/核心
GreatBelt	CTC5160		120Gbps	10G	城域网接入/汇聚 PTN/IP RAN 企业网接入/汇聚

数据来源：盛科通信官网，盛科通信公司公告，裕泰电子官网，东吴证券研究所

盛科通信持续投入研发资源抢占国产替代空间。营业收入上，在 2019-2024 年间，营业收入高速增长，从 1.92 亿元上升至 10.82 亿元。同时以太网交换芯片业务在总营收中的占比也在不断提升。以太网交换芯片行业具备较高的技术壁垒、客户及应用壁垒和资金壁垒，因此当前行业整体国产程度低，国内参与厂商较少。盛科通信的以太网交换芯片在国内具备先发优势和市场引领地位，为我国数字化网络建设提供了坚实的芯片保障。

图11：盛科通信 2020-2025H1 研发投入与交换芯片业务营业收入



数据来源：公司年报，东吴证券研究所

数渡科技：PCIe 5.0 交换芯片领域率先进入客户导入和小批量阶段，2025 年底有望

实现国产替代的突破。公司 PCIe 5.0 交换芯片被视为解决国产 AI 芯片互联性能瓶颈的关键器件，目前国内同行尚未实现批量供货。数渡科技产品已获得多家客户导入验证，其中 3 家客户签署了小批量采购协议，9 家客户完成制板，并已有企业投入商用。若进展顺利，预计 2025 年底起可逐步批量供货，有望缩短与国际主流产品差距，率先切入 AI 基础设施互联市场。

图 12：数渡科技产品矩阵

类别	产品名称	产品定位/说明
芯片	超节点智算服务器的高速交换芯片	突破国产智算服务器 8 卡限制，实现更大规模与更高性能
	通用服务器内部互连的 PCIe/CXL 高速扩展和交换芯片	连接 ASIC、FPGA、存储设备等，实现高速互连与算力/数据处理提升
	存储服务器内部扩展的 PCIe/CXL 高速扩展和交换芯片	扩展连接高性能存储阵列，提升读写性能和存储密度
	边缘算力平台内部互连的 PCIe 高速扩展和交换芯片	连接传感器、控制器、存储器、网络设备，实现高效数据传输
智能算力系统方案	高性能算力平台系统方案	通过系统层面优化，在算力不足时提升整体效能
	高度智算服务器系统方案	突破单服务器资源限制，实现算力资源整合
	纵向扩展智算服务器模型	支持单台服务器纵向扩展至数十 GPU/NPU
	高性能集中式存储服务器系统方案	支持 PB 级容量，RAID 冗余，自动修复与安全功能
	全国异构边缘服务器系统方案	实现内存、计算、存储和网络资源池化与调度

数据来源：数渡科技公司公告，东吴证券研究所

澜起科技：以 Retimer 为切入点，正逐步构建完整的 PCIe/CXL 互联产品族，覆盖 Retimer、MXC 到 Switch，定位 AI 互联底层关键器件供应商。2025 年 1 月公司已向客户送样 PCIe 6.x/CXL 3.x Retimer，技术指标达到国际先进水平，并在 CXL MXC 上通过合规认证，进一步夯实了在互联领域的地位。同时，公司明确布局 PCIe Switch，并强调自研 SerDes 能力，意味着其产品线正从信号中继延伸到互联调度。Retimer、MXC 与 Switch 共同构成 CPU↔GPU、GPU↔GPU 互联的关键基础器件组合，广泛应用于 AEC 有源线缆、背板与机箱级 Scale-Up 场景，未来有望成为国产 AI 算力底座的重要受益者。

2.2. 大厂自研交换芯片

海光信息：依托自研处理器与协处理器平台，在 Chiplet 互联和高速 I/O 方向深度布局，逐步构筑 Scale-Up 互联的核心能力。公司在 2024 年年报中披露，已开展先进封装与高带宽低时延 Chiplet 互联研发，并通过 ComboPHY 支持处理器间互连总线、PCIe 及 CXL 等高速 I/O 接口，能够扩展片上网络（NoC）并强化 QoS 与低时延特性。2025 年 9 月 13 日，“海光系统互联总线协议开放生态研讨会”在京举办。会上，海光面向 GPU、IO、OS、OEM 等产业全栈，正式宣布开放 CPU 互联总线协议（HSL）。我们认为同时布局 CPU、GPU（DCU）、Switch 互联，已初步形成算力基础设施全覆盖。随着 AI 大模型训练/推理需求的增长，公司产品有望形成多点联动，卡位优势显著。

图 13：海光信息产品矩阵



数据来源：海光信息公司公告，东吴证券研究所

华为：超节点集群布局领先，开源 Scale-out 互联协议抢占生态高地。华为于 1990 年开启交换芯片自研，1999 年自研 Solar 系列交换芯片，并于 2016 年发布 Solar 5.0 交换芯片。该芯片采用 16nm 制程，架构上的持续优化使之较上代版本提升了 4 倍的吞吐量。2024 年 9 月 19 日，华为在全联接大会上发布了单芯片 51.2T 数据中心盒式液冷交换机 CloudEngine XH9230。2025 年 9 月 18 日，华为举办全连接大会，发布新一代 AI 芯片、超节点与集群产品，同时开放“灵衢”互联协议。其中最新超节点产品 Atlas950 SuperPoD 和 Atlas 960 SuperPoD 分别支持 8192 及 15488 张卡，在卡规模、总算力、内存容量、互联带宽等关键指标上实现了全面领先。基于超节点，华为同时发布了超节点集群产品 Atlas950 SuperCluster 和 Atlas960 SuperCluster，算力规模分别超过 50 万卡和达到百万卡。会上，华为还宣布将开放面向超节点的互联协议灵衢（UnifiedBus）2.0 技术规范，邀请产业界伙伴基于灵衢研发相关产品和部件，共建灵衢开放生态。我们认为华为在 AI 算力基础设施从芯片、互联到整机集群整体布局，从产品路线图看已实现全栈打通，拥有持续翻倍迭代的能力。

图14： 华为 2025 全连接大会产品路线图

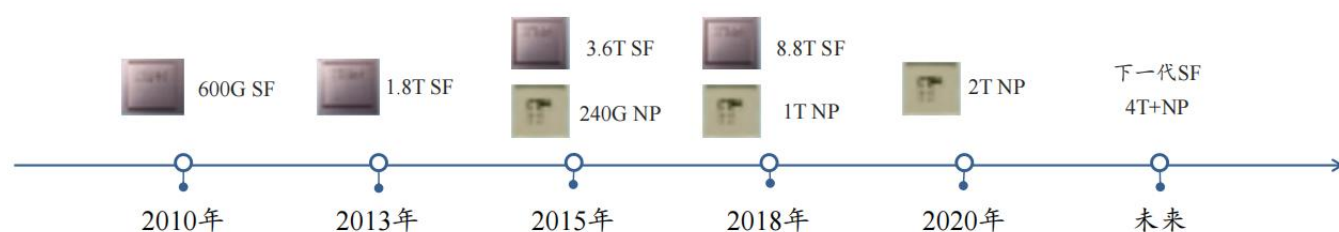
产品	上市时间	核心定位与亮点
Ascend 950PR / 950DT AI芯片	950PR: 26Q1; 950DT: 26Q4	新一代昇腾芯片，支持FP8/MXFP8/MXFP4/HiF8数据格式，互联带宽提升至2TB/s，分别面向推理Prefill与推荐场景（950PR）及推理Decode与训练场景（950DT）。
Ascend 960 / 970 AI芯片	960: 27Q4; 970: 28Q4	在算力、内存带宽、互联带宽等全面翻倍升级，960支持HiF4精度，970计划在FP4、FP8算力与互联带宽再翻倍。
Atlas 950超节点	26Q4	基于Ascend 950DT打造，支持8192卡，FP8算力8 EFLOPS，FP4算力16 EFLOPS，互联带宽16PB/s，训练性能比Atlas 900提升17倍。
Atlas 960超节点	27Q4	基于Ascend 960打造，支持15488卡，FP8算力30 EFLOPS、FP4算力60 EFLOPS，内存容量4460TB、互联带宽34PB/s，性能再度翻倍。
Kunpeng 950处理器	26Q1	96核/192线程与192核/384线程版本，支持通用计算超节点，安全新增四层隔离，首颗机密计算数据中心处理器。
TaiShan 950超节点	26Q1	全球首个通用计算超节点，基于Kunpeng 950，最大16节点、32颗处理器、48TB内存，支持内存/SSD/DPU池化，可平滑替代大型机和小型机。
灵衢（UnifiedBus）2.0互联协议	2025	面向万卡超节点的新型互联协议，实现高可靠全光互联、2.1微秒超低时延、TB级带宽，并首次向产业界开放2.0规范。
Atlas 950 SuperCluster集群	26Q4	64个Atlas 950超节点组成，50万+卡，FP8总算力524 EFLOPS，支持UBoE与RoCE协议。
Atlas 960 SuperCluster集群	27Q4	百万卡级集群，FP8总算力2 ZFLOPS，FP4总算力4 ZFLOPS，静态时延更低、可靠性更高

数据来源：华为全连接大会，重庆晨报，东吴证券研究所

中兴通讯：已形成从通用高性能交换芯片到面向 AI 超节点的自研交换/NP 芯片全

栈布局。在交换机芯片领域，中兴通讯在 2008 年启动交换芯片的自研，于 2011 年成功推出第一代自研交换网套片，并迅速在路由器等产品上成功应用。随后的几年，中兴通讯持续改进交换网技术，紧跟工艺革新的节奏，以 3 年一代的速度进行交换芯片的更新换代，陆续推出了 600Gbps、1.8Tbps、3.6Tbps 交换容量的 SF 系列交换芯片。并于 2018 年推出了 8.8Tbps 交换容量的第四代自研交换芯片，达到业界一流水平。2020 年，中兴通讯启动了第五代自研交换芯片的研发。此外，中兴通讯还于 2015 年推出了首款自研 NP 芯片 SSP-1，并于 2019 年初推出了业界首款集成 FlexE 和 TSN 功能的 NP 芯片，展示了中兴通讯在交换机芯片技术上的领先地位和创新能力。2025 年，中兴通讯推出了基于自研 AI 交换芯片的超节点方案，GPU 间通信带宽达到 400GB/s 至 1.6TB/s，能够支持上百至上千张算力卡的高效互联。公司在超节点、高性能交换机、AI 交换芯片等关键环节完成了技术卡位，形成了国内市场的稀缺性。

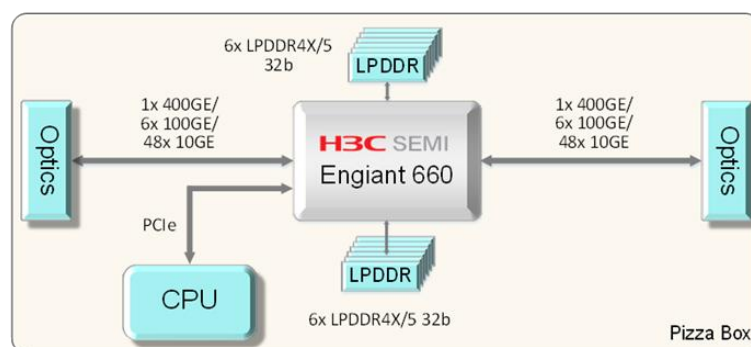
图15：中兴通讯交换芯片与 NP 芯片技术路线图



数据来源：中兴通讯官网，东吴证券研究所

新华三：自主研发智擎系列可编程 NP 芯片，接口吞吐能力 1.2Tbps。新华三于 2021 年 4 月正式发布其首款自研网络处理芯片智擎（Engiant）600/660，并于同年 7 月实现量产。该系列芯片基于 16nm 工艺制造，集成 256 个专用 CTOP 处理核心、支持 4096 硬件线程，提供 1.2Tbps 的接口吞吐能力，集成度超过 180 亿晶体管，搭载 12 路 LPDDR5 控制器。芯片支持 C 语言编程，覆盖 L2-L7 层网络业务处理，内置流量管理、查找引擎、报文管理等加速单元，可灵活满足路由、交换、防火墙、负载均衡、SDN/NFV 等多类网络设备需求。智擎 660 有望推动国产高端网络处理器在运营商核心网、数据中心及智能安全设备领域的应用落地。

图16：新华三智擎 660 网络处理芯片结构图



数据来源：半导体行业观察，东吴证券研究所

3. 风险提示

AI 应用进展不及预期。算力的长期需求是建立在 AI 应用逐步发展之上，在初期大模型训练带来大量算力需求之外，AI 应用带来的推理需求是长期维度上相关硬件设备市场空间增长的前提。如果 AI 应用进展不及预期，将对算力各环节需求产生影响。

技术发展不及预期。Scale-Up 互连最底层的 SerDes 技术是 AI 芯片互连环节的核心，该技术难度大、壁垒高，若下一代发展受阻，或影响 AI 芯片互连速率提升。

市场竞争风险。虽然在 Scale-Up 环节，英伟达当前处于一骑绝尘的竞争地位，但随着 UALink 联盟的崛起，以及技术的不断发展，尤其是海外云厂商自研芯片入局，市场竞争格局或将发生变化。

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。

本报告的版权归本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。经授权刊载、转发本报告或者摘要的，应当注明出处为东吴证券研究所，并注明本报告发布人和发布日期，提示使用本报告的风险，且不得对本报告进行有悖原意的引用、删节和修改。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

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>