650 GROUP
Market Intelligence Research

RDMA Networking and AI
Research Report

June 24, 2024

RDMA Networking and AI

Table of Contents

What is Remote Direct Memory Access (RDMA)

RDMA is a technology that allows two servers to read and/or write to each other's memory without going through either the server's processor, cache, or operating system. RDMA frees up the CPU, lowers latency, and provides faster data transfer by working around these resources. Network, storage, and compute applications benefit from this technology. RDMA is implemented in each server's Network Interface Card (NIC). By bypassing the operating system and networking kernel, networking performance and data exchange between the two servers occurs much faster. RDMA was initially designed for high-performance computing (HPC) clusters for massively parallel computing. Massive parallel computing is how Artificial Intelligence / Machine Learning (AI/ML) works, and RDMA is a vital component of those deployments. RDMA was developed initially with InfiniBand networks.

What is RDMA over Converged Ethernet (RoCE)

RDMA over Converged Ethernet allows the same direct memory exchange to occur over Ethernet networks. In early versions of RoCE, you needed Converged Ethernet, but in later versions, RoCE runs over standard Ethernet Networks. RoCE runs on top of Ethernet and requires support in the Ethernet NIC to implement. At an industry level, significant effort is being put into Ethernet to improve congestion control and reduce traffic loss. It is important to note that the Ethernet switch installed base in the data center has in excess of 400 M ports. Ethernet will play a growing role in AI/ML networks due to its ubiquitous nature, and a significant amount of RDMA will occur over Ethernet as we advance.

Introduction

Before 2022, the primary use case for RDMA was HPC. Most HPC buildouts were focused on supercomputing projects and didn't have much applicability in Cloud and Enterprise data centers. That quickly changed in late 2022 as AI/ML became a focal investment, and data center spending quickly shifted to AI/ML and away from traditional computing. The surge was unprecedented, with RDMA base networks exiting 2023 at a rate higher than all of 2022 and 2021 combined. The surge caused RDMA to become mainstream and critical to the expansion of AI/ML.

RDMA Networking and AI

Change in Server Market

The server market changes significantly as customers move from general-purpose servers to AI/ML servers. In this transition, AI/ML servers will grow from 1 M units in 2023 to over 6 M units in 2028 (Figure 1). At the same time, AI/ML servers will approach nearly $300 B in spend by 2028 (Figure 2). The data center market has never seen such significant growth or rapid transition, and spending in the data center will double because of AI/ML. Most of the 6 M AI/ML servers in 2028 will have a back-end network or AI Fabric to interconnect the compute.

Figure 1: Server Market Units

The number of GPUs and AI ASICs per server within the server market will increase with each generation. Today, 8 GPUs is the most common deployment, but we can expect the market to expand to 16 and 32 GPUs by the decade's end. The amount of memory per GPU will continue to increase as training models scale from billions to trillions of parameters. With this type of scale, the ability to transfer data more efficiently between servers becomes a pivotal factor in scaling and achieving the ambitious goals of these training models. RDMA will play an essential role in these servers as the ability to access memory and resources on other servers is a crucial area to scale.

Job Completion Time (JCT) and Performance Metrics

The ability to access memory on other servers directly helps improve the overall performance of the AI model. RDMA also helps reduce JCT by quickly getting the data into the GPU. One of the most significant issues in early AI/ML clusters is
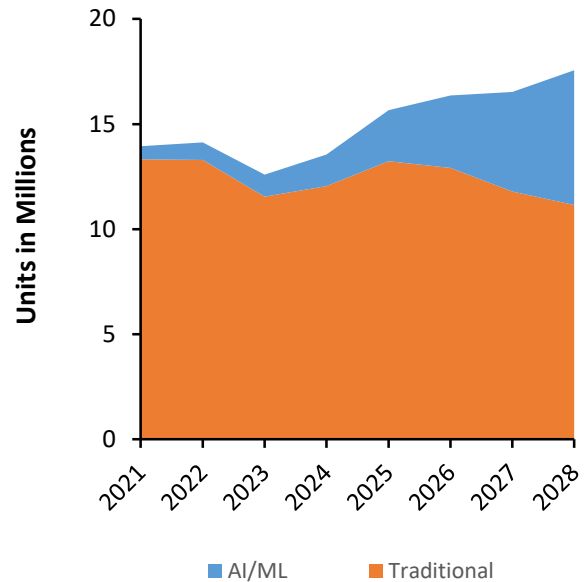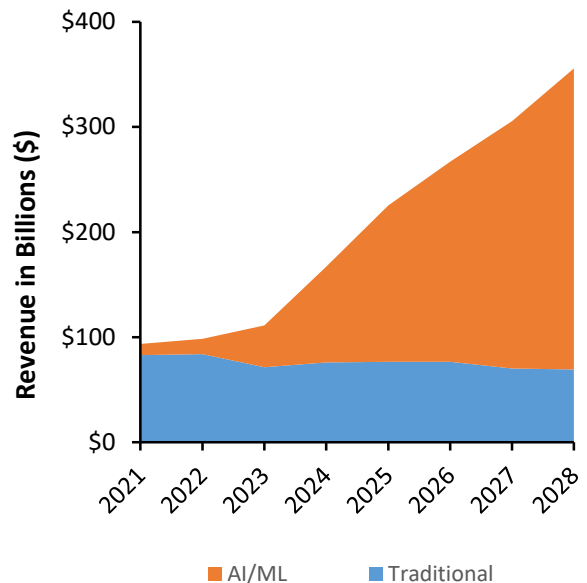

Figure 2: Server Market Revenue

idle GPU cores, where the whole AI/ML cluster can come to a screeching halt and become idle because of dropped packets or data that doesn't get to the GPUs quickly enough. RDMA helps solve this networking bottleneck and ultimately improves JCT and performance metrics in the AI/ML cluster. While there may be additional performance differences between Ethernet and InfiniBand, RDMA is a massive improvement over standard networks.

NIC Market

All InfiniBand NICs support RDMA, but not all current Ethernet NICs support RDMA/RoCE. Traditional Ethernet NIC vendors that want to participate in AI/ML must add RoCE to their offerings to participate in the market and compete with vendors that already support RoCE. As NIC speeds transition to 400 Gbps and above, most Ethernet NIC offerings will likely support RoCE. Increased functionality and higher port speeds will drive Ethernet ASPs higher for NICs.

Vendors will have different RoCE capabilities in their Ethernet NICs as the type of processor, additional offload engines, and engineering expertise will create different performance tiers across all the NIC vendors. As each product generation comes out, we expect vendors to fine-tune their offerings and RoCE performance metrics to become more consistent between vendors. Consistency will ultimately lead to better interoperability and more customer choice.

AI/ML Back-End Networks

Most AI/ML servers utilize a back-end network. The back-end network is separate from the rest of the data center and can utilize InfiniBand or Ethernet. The back-end network focuses solely on the AI/ML cluster and connects each server. Back-end networks focus on GPU-to-GPU or GPU-to-Memory connectivity. It is important to note that back-end networks are incremental to the existing networks, increasing the number of ports per server and significantly increasing the networking revenue opportunity.

AI/ML can have multiple back-end networks that focus on different tasks. For example, RDMA runs on Ethernet and InfiniBand, but there can be additional networks that each GPU provider or AI ASIC uses to create higher-performance solutions.

Revenue ($B)	2021	2022	2023	2024	2025	2026	2027	2028
Ethernet	$0.2	$0.3	$0.9	$2.1	$6.1	$8.7	$12.0	$13.7
InfiniBand	$0.9	$1.5	$5.9	$10.2	$8.3	$8.9	$8.9	$8.8
Total	$1.0	$1.8	$6.9	$12.3	$14.5	$17.6	$20.9	$22.5

Table 1: RDMA Market Forecast by Technology

Revenue ($B)	2021	2022	2023	2024	2025	2026	2027	2028
NICs	$0.4	$0.6	$2.2	$3.6	$3.5	$3.8	$3.8	$3.9
Switching	$0.6	$1.2	$4.7	$8.7	$11.0	$13.8	$17.1	$18.6
Total	$1.0	$1.8	$6.9	$12.3	$14.5	$17.6	$20.9	$22.5

Table 2: RDMA Market Forecast by Device Type

Size of the Market

The RDMA capable market ranged from $400-$700 M a year before 2021, led primarily by the HPC use case. The demand exceeded $6 B in 2023, led by the surge in AI/ML deployments, and should exceed $22B in 2028 Tables 1 and 2 and Figures 3 and 4 next page). As operators increase their AI/ML CAPEX spend and future guidance, RDMA projects will likely revise upwards in the coming years.

The RDMA market segments into two distinct categories. The first is on technology. Today, RDMA is mainly deployed with InfiniBand, but as we advance, there will be more RDMA networks in Ethernet. The second segmentation relates to the NIC and Switch. On the InfiniBand side, this purchase occurs simultaneously, but on the Ethernet side, procurement of the switch often occurs with different teams and at a different cadence than the NIC. This server team and networking team silo approach will still happen in AI/ML because organizations silos will remain despite the converged nature of AI/ML. Therefore,

Figure 3: RDMA Market Revenue by Technology
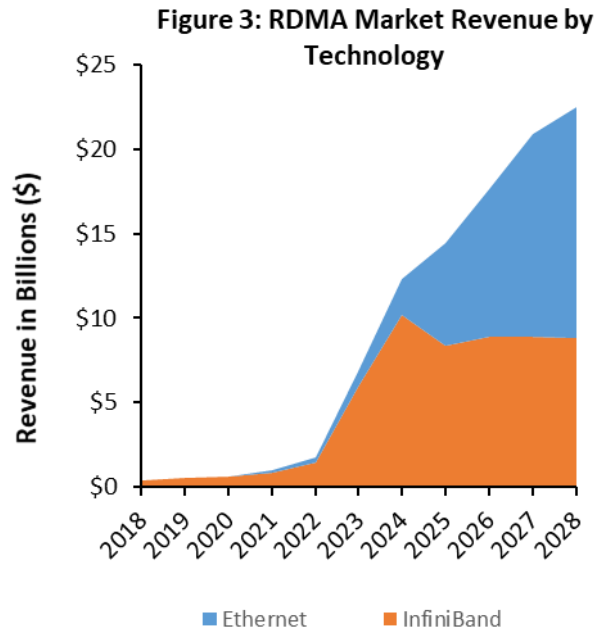
we expect each customer to develop a unique preference for NIC and switch suppliers in their AI/ML networks and that potential AI/ML will have different NIC and switch suppliers than a customer's traditional compute deployments.

Conclusion

RDMA and RoCE play a critical role in AI/ML networking. AI/ML could not scale as quickly as the customer demands without them. As the server market quickly transitions to AI/ML and away from traditional computing, the market opportunity for RDMA and RoCE will explode over the next several years.

Figure 4: RDMA Market Revenue by Device Type

While there will be customer and vendor preference between technologies, RDMA will thrive. Ethernet and InfiniBand will coexist, and the market/customers should not view these as binary events. Most customers will use RDMA on both InfiniBand and Ethernet networks and across a mix of GPUs. Very few customers will be exclusive to one type of network or supplier. AI/ML workloads span multiple workloads, from foundational training to reinforcement learning to inference. Having RDMA work across numerous technologies helps customers scale and focus on the AI/ML workloads, not the underlying networks.