MARVELL

Product Brief

Teralynx® T100 Data Center Switch

102 Tbps Ethernet switch series for AI data center networks
MV-TX93251 | MV-TX93225 | MV-TX93151

Overview

As artificial intelligence (AI) workloads continue to grow in complexity and scale, the infrastructure supporting AI models must evolve to meet performance, efficiency, and scalability demands. These complex workloads require massive data transfers between XPUs, CPUs, and storage. A XPU-to-XPU backend network is designed to facilitate high-speed, deterministic low-latency communication between multiple XPUs for parallel processing and intensive computational tasks.

The Marvell® Teralynx® T100 is designed to be Ultra Ethernet Consortium (UEC) compliant, targeting large-scale XPU scale-out networks such as multi-thousand-XPU AI training clusters and rail-optimized backend fabrics. It supports UEC's Ultra Ethernet Transport (UET) for XPU-to-XPU RDMA, enabling packet spraying and out-of-order packet delivery across multiple parallel paths to maximize link utilization and eliminate flow-level congestion hot spots. Combined with advanced, standards-based congestion control and multi-pathing, the T100 helps ensure consistent low-latency, high-bandwidth performance for collective operations (all-reduce, all-to-all, all-gather) across the largest AI training and inference clusters.

The Teralynx series of Ethernet switch silicon has been architected from the ground up to provide the most optimized Ethernet solutions for AI networks. It offers the industry's lowest deterministic latency, unmatched visibility through fine-grained telemetry, the industry's best performance-per-watt power efficiency, large buffers and rich programmability. Teralynx offers a broad range of data center switches with full software/SDK compatibility across the entire series.

Example Use Cases:

- **Scale out.** Rail-optimized backend fabrics connecting multi-thousand-XPU AI training clusters, enabling UEC-compliant, non-blocking spine and leaf connectivity across large GPU/XPU domains with deterministic low latency and lossless RDMA transport.
- **Scale up.** High-radix, low-latency interconnect for tightly-coupled XPU pods and rack-scale domains, supporting dense all-to-all connectivity within a node or rack to maximize collective-operation performance (all-reduce, all-gather, all-to-all) before traffic crosses into the scale-out fabric.

Packaging Options:

- **BGA** (ball grid array) package for traditional pluggable-optics or passive-copper system designs.
- **CPC** (co-packaged copper). Copper-integrated package that shortens the electrical path to the faceplate, reducing power and improving signal integrity for high-radix, short-reach connectivity.
- **CPO** (co-packaged optics). Optical engines integrated in-package with the switch silicon, minimizing SerDes reach and power for maximum bandwidth density.

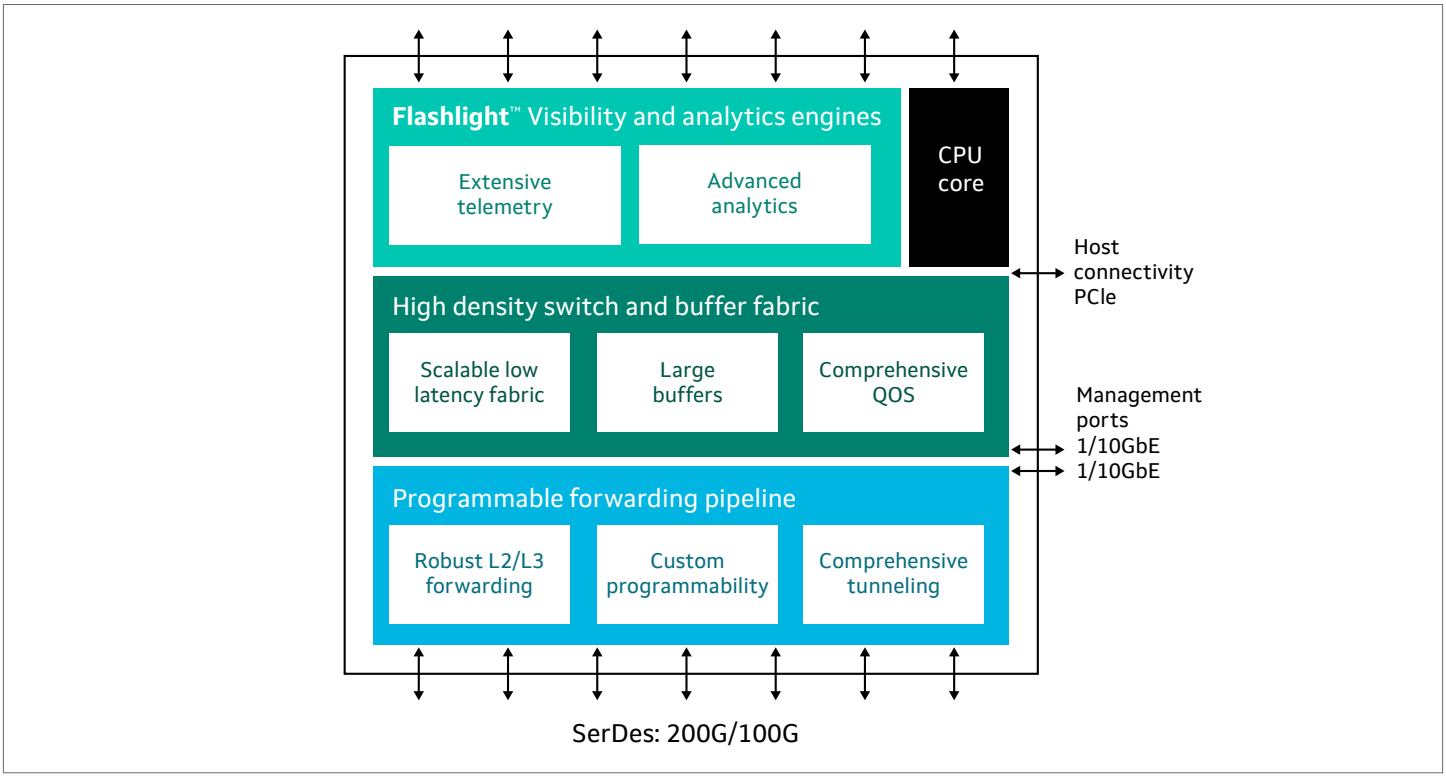
Key Benefits:

Groundbreaking performance and scale enable customers to deploy transformative solutions with reduced power consumption, latency and cost.

- Innovative 200 Gbps PAM4 SerDes lowers the cost per bit and enables higher scale I/O such as 800GbE and 1600GbE while maintaining backward compatibility with 50G and 100G PAM4.
- Substantial reduction in power consumption due to monolithic design eliminating the overhead of die-to-die communication.
- Industry-leading deterministic low latency helps mitigate issues caused by tail latency, reducing job completion times for training and fine-tuning workloads, while simultaneously improving tokenomics – tokens/sec/XPU - the key efficiency metric for inference clusters serving real-time AP applications.
- Enhanced buffer sharing capability provides robust packet handling in a low entropy, high bandwidth AI network minimizing congestion and latency. The same capability is equally critical for disaggregated inference architectures, where incast traffic patterns emerge between prefill servers, KV-cache tiers, and decode servers, demanding precise buffer management to sustain throughput and response SLAs.
- Rich and real-time analytics capability with sub second accuracy that can easily integrate with end user applications using REST APIs.
- Programmable forwarding pipeline enables support of custom and new standard protocols without requiring spins to future-proof the network.

- **512 radix along with superior 200 Gbps SerDes performance** enables large clusters to be designed with passive copper resulting in massive TCO savings.
- **Teralynx switches support standards-based SAI** for seamless integration and programmability in high capacity networks.

Block Diagram



Key Features

Features	Details
Switch capacity up to 102.4 Tbps with large buffers	• Delivers the highest throughput in the industry, enabling massive AI/ML clusters and next-generation data center spine deployments
Up to 512 SerDes supporting 50G, 100G, 200G, 400G, 800G, and 1600G I/O	• Provides unmatched port density and flexibility to support a wide range of optical and copper interconnects across generations
Industry-leading power efficiency via monolithic chip design	• Eliminates die-to-die overhead, reducing total power consumption and lowering data center operating costs
Programmable forwarding pipeline	• Supports proprietary, new, and emerging protocols without ASIC re-spins, future-proofing network investments
Teralynx® Flashlight™ visibility and telemetry, including support for P4 in-band network telemetry (INT) along with critical extensions	• Delivers breakthrough real-time analytics including P4-INT, path tracing, queue drop forensics, and IBM Watson Analytics Interface for sub-second accuracy
Industry-leading deterministic low latency (cut-through and store-and-forward)	• Minimizes tail latency and job completion times, directly improving AI/ML training efficiency
Advanced QoS with 12 queues and 4 lossless CoS, RoCE and DCB support	• Ensures reliable, lossless transport for GPU-to-GPU RDMA traffic in high-band-width AI backend networks
Comprehensive PTP (IEEE 1588v2) and SyncE support	• Enables precise timing and synchronization required for time-sensitive distributed AI workloads

Features	Details
PCIe 5 x2 host connectivity with two on-chip Arm cores	• Provides fast DMA transfers to the host CPU and enables embedded application processing
Flexible packaging options	• BGA for standard chip to module configuration, CPC to enable retimer-less in-rack connectivity, and CPO to support optical scale-up architectures.

Target Applications

- AI and cloud data center frontend networks
- AI data center backend networks
- AI scale-up networks

Switch System Examples

102.4T Systems

- 1RU: 512 × 200G (backplane — CPC version only)
- 2RU: 64 × 1600G (OSFP1600)
- 4RU: 128 × 800G (QSFP224)

51.2T Systems

- 2RU: 64 × 800G-R8 (OSFP800 / QSFP-DD800)
- 1RU: 32 × 1600G-R8 (OSFP1600 / QSFP-DD1600)
- 2RU: 64 × 800G-R4 (QSFP224)

Ordering Information

Characteristics	P/N: MV-TX93251	P/N: MV-TX93225	P/N: MV-TX93151
Capacity	102.4 Tbps	51.2 Tbps	51.2 Tbps
SerDes Lanes	512 × 200 Gbps	256 × 200 Gbps	512 × 100 Gbps
50G / 100G ports	512	256	512
200G ports	512	256	256
400G ports	256	128	128
800G ports	128	64	64
1600G ports	64	32	—

MARVELL™

To deliver the data infrastructure technology that connects the world, we're building solutions on the most powerful foundation: our partnerships with our customers. Trusted by the world's leading technology companies for over 30 years, we move, store, process and secure the world's data with semiconductor solutions designed for our customers' current needs and future ambitions. Through a process of deep collaboration and transparency, we're ultimately changing the way tomorrow's enterprise, cloud and carrier architectures transform—for the better.

Marvell_Teralynx_MV-TX93xxx_PB Revised: 08/26