



Mellanox Ethernet Cloud Fabric (ECF)

Build Your Data Center Network Like the Cloud Titans

Data center networks are growing at a tremendous pace in size, scale, and complexity. The number of infrastructure elements and the pace of changes have gone up significantly. At the same time, there is a business imperative to accelerate application deployment, improve uptimes and reduce costs.

Over the last decade, cloud titans such as Amazon, Google and Microsoft have successfully built large-scale infrastructures by following and honing certain web-scale design principles. Mellanox Ethernet Cloud Fabric (ECF) democratizes this web-scale capability as a shrink-wrapped solution that both enterprises and tier-2 clouds can consume and deploy on foundational silicon technology.

ECF is a platform for building lock-in free, scalable, high performance private and public cloud infrastructures. It builds on the trend of moving away from box by box configuration towards datacenter-wide automation. ECF is open, automated, easy to operate and offers high performance.

ECF is based on four primary principles:

- Open networking
- Advanced network telemetry
- Software Defined Networks (SDN)
- Consistent, predictable, high-performance hardware

Open Networking

Legacy vendors offer closed networking solutions. Even after purchasing the hardware, customers are forced to license the most useful features and can only use the proprietary software and expensive re-labeled optics from that one vendor. The architectural choices are limited by the pace, priority, and capacity of the vendor to add features. Even worse, most legacy vendors build their systems with the same commodity silicon with crippling limitations. Such closed solutions are driven more by supply-chain rather than by better technology and innovation. ECF combines a variety of “fit for purpose” operating systems with high performance Mellanox Spectrum® Ethernet switches. ECF supports a selection of operating systems including Mellanox Onyx™, Cumulus Linux, Linux Switch and Microsoft SONiC. With ECF, customers have the flexibility to choose the architecture and operating system that best fits their specific needs.

Legacy vendors force prescriptive methods to orchestrate and automate infrastructure. This approach not only limits choices but also creates vendor lock-in. To help customers and to cement its position as a hardware vendor, Mellanox has contributed heavily to open-source software development. As of 2019, Mellanox is the only switch vendor among the list of top 10 that is contributing to open-source Linux.

ECF allows customers to build heterogeneous networks with equipment from multiple vendors. Open and vendor-agnostic tools such as Ansible, SaltStack, and OpenStack help to abstract vendor and device-specific configurations, orchestrate and automate the data center network while avoiding lock-in.

ECF GUIDING PRINCIPLES

Open Networking

- No vendor lock-in
- Fit-for-purpose NOS choices
- Custom and simplified operations

Advanced Network Telemetry

- Faster troubleshooting
- Switch-based accelerated telemetry
- Open and works in multi-vendor environments

Software Defined Networks

- Single-pass VXLAN bridging/routing
- 10X better VXLAN scale
- 100X better ACL scale
- Programmable blocks for future proofing

High-performance Hardware

- Consistent and low cut-through latency
- Predictable and fair bandwidth sharing
- Robust RoCE data path
- 1GbE out-of-band management port

Advanced Network Telemetry

Cloud titans have dramatically improved their network uptimes and improved resource utilization by leveraging advanced network telemetry. Traditional network monitoring approaches like SNMP polling, sFlow sampling, and streaming telemetry, use a centralized collector for telemetry data storage and problem identification. As data speeds increase, these traditional approaches present a growing challenge to the transportation, processing, and storage of telemetry data. Even a small network can generate petabytes of data per day. Some networks use powerful server clusters to cope with collecting and analyzing the data. Even with this expensive infrastructure, it is hard to pinpoint the root cause of problems without the appropriate network context.

The difficulty with commodity Ethernet switches is that they are unable to provide the critical telemetry data needed for root cause problem identification and issue resolution. This forces operators to process massive amounts of telemetry data centrally, creating an artificial bottleneck to problem resolution.

With What Just Happened (WJH)[™] telemetry, Mellanox Spectrum Ethernet switches provide rich, contextual and actionable insights on a variety of topics including Layer-1 through Layer-4, ACLs and buffer occupancy. A WJH[™] agent collects events and visibility insights locally from Mellanox Spectrum Ethernet switches to provide instant answers to when, what, why, who and where – critical questions in problem solving. WJH is available on every Mellanox Spectrum platform running any network operating system (NOS) including Mellanox Onyx, Cumulus Linux, Linux Switch and Sonic. WJH can be integrated and extended with both third party and open-source tools.

With WJH, traffic inspection, filtering, and issue identification are performed by the switch platform where the network context is readily available. As a result, only issue-relevant data is streamed out. For example, when packets are dropped, WJH implements the appropriate drop counter as well as captures the relevant header data from the packet, for more thorough analysis. WJH also helps network operators by dramatically reducing mean time to innocence (MTTI) or issue resolution. Additionally, WJH provides insights to help improve resource utilization and capacity planning.



Software Defined Networks

Software-defined networks are revolutionizing modern data centers with the promise of dynamically changing the underlying fabric that connects servers and storage, containers and virtual machines. Software-defined networking demands agility and programmability from the underlying physical infrastructure. Some vendors have taken the “large hammer” approach by building switching platforms that are completely programmable. They have thrown-out dedicated hardware functions that have been perfected over years of evolutionary improvements, with the downside that even well-understood traditional network functions need to be programmed into these systems. Such platforms trade performance for programmability. They support radical programmability by substantially reducing the packets rate and throughput of the switch.

Mellanox FlexFlow packet processing technology provides high performance without compromising programmability and flexibility. The FlexFlow-based pipeline is comprised of optimized functional blocks that are streamlined in hardware to deliver the highest performing traditional functionality, and programmable blocks that provide a flexible packet processing capability. With this hybrid pipeline approach, there are no mandatory/static packet lookups. Advanced users can customize the lookups as needed. Traditional packet flows are optimized for low latency by performing only those lookups that are pertinent to that flow.

With Mellanox FlexFlow, the Spectrum family of switches delivers the highest performance at scale and at the same time supports programmability. Spectrum switches have comprehensive support for VXLAN overlays including single pass VXLAN routing and bridging, as well as centralized VXLAN routing for brownfield environments. The Spectrum hardware pipeline also supports other overlay protocols including VXLAN, VXLAN-GPE, MPLS-over-GRE/UDP, NSH, NVGRE and MPLS/IPv6-based segment routing. The desired network virtualization capabilities can be enabled with the fit-for-purpose operating system.

Consistent, Predictable, High-performance Hardware

Modern data center infrastructures comprise of a myriad of high-performance storage and compute endpoints. All-flash storage arrays, multi-core virtualized servers, and GPUs are pushing performance to new highs. The data center network is the critical portion of the infrastructure that interconnects these high-performance endpoints.

Mellanox Spectrum Ethernet switches, with their best-in-class buffer architecture, support a robust high-throughput data path. Spectrum switches not only provide high-performance interconnections but also share bandwidth resources fairly across competing ports and traffic flows. Spectrum switches implement smart and standards based congestion management mechanisms which help streamline RDMA over Converged Ethernet (RoCE) traffic that powers NVMe over Fabrics (NVMe-oF) storage and Machine Learning applications.

Conclusion

Mellanox Ethernet Cloud Fabric combines industry proven web-scale principles with high-performance Spectrum Ethernet switch silicon technology to deliver enterprise-ready cloud infrastructure solutions. Combined with WJH, ECF dramatically reduces time to issue resolution, simplifies network operations and improves network uptimes. ECF enables a variety of Software Defined Networking and overlay technologies by leveraging the unique Mellanox Spectrum FlexFlow technology. Spectrum Ethernet switches also support a robust and high bandwidth data path.



About Mellanox

Mellanox Technologies (NASDAQ: MLNX) is a leading supplier of end-to-end Ethernet and InfiniBand intelligent interconnect solutions and services for servers, storage, and hyper-converged infrastructure. Mellanox intelligent interconnect solutions increase data center efficiency by providing the highest throughput and lowest latency, delivering data faster to applications and unlocking system performance. Mellanox offers a choice of high performance solutions: network and multi-core processors, network adapters, switches, cables, software and silicon, that accelerate application runtime and maximize business results for a wide range of markets including high performance computing, enterprise data centers, Web 2.0, cloud, storage, network security, telecom and financial services. More information is available at: www.mellanox.com



350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com