DELL Technologies

Dell Networking

Fabrics Ethernet para Inteligencia Artificial Generativa con SONIC

Gonzalo de Antonio

Dell Technologies Networking Sales Executive
Western Europe South

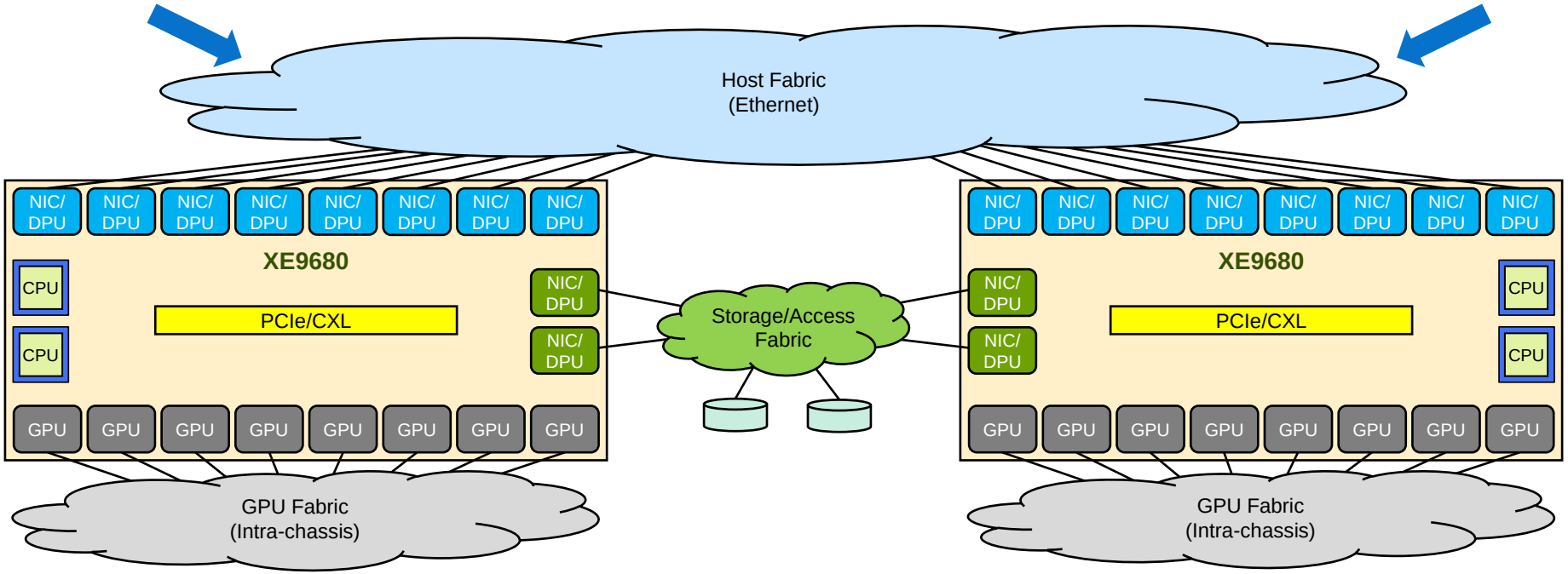
AI workload types and design considerations

	Building the AI model		Using the AI model
Workload	**Training** • Massive consumption of data to create a foundation of general knowledge • Processing can take days or weeks of elapsed time • Often the realm of “AI factories” or service providers given the size and scale of the infrastructure requirements	**Fine Tuning** • Refines base models with specialized knowledge for specific applications • Often run on-prem to ensure data security and protection of internal IP • Smaller clusters are in reach of enterprise customers	**Inferencing** • Leveraging the models for business applications • Often runs on-prem given the criticality of the information and results to a business • Must deliver results in a conversational flow, requiring excellent system responsiveness
Requirements	Highest Scale, thousands of GPU’s • Bursty Elephant Flows running iteratively over many thousands of cycles • Any congestion or delay will drive extended tail latencies – dramatically extending job completion time • **Fabric needs highly optimized topologies and load balancing strategies**	Lower Scale, dozens to hundreds of GPU’s • Derived from a general base model as a starting point • Similar workload pattern to training, but smaller data sets and fewer iterations required • **Less sensitive to BW / Latency in relation to overall job completion time**	Lowest Scale, can be CPU or GPU • Least demanding load on the fabric • Ability to serve requests and provide rapid response • **Most inferencing environments can execute in a single host**

Copyright © Dell Inc. All Rights Reserved.

AI Fabrics Model

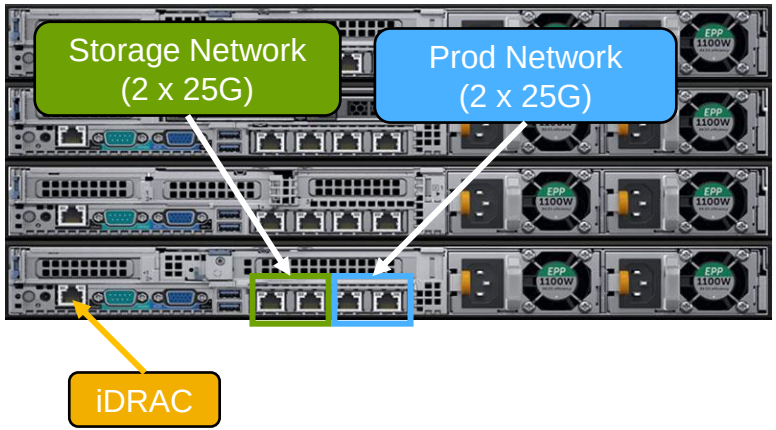
Other names used in the industry
- Host Fabric: Front-End Fabric
- GPU Fabric: Back-End Fabric



Copyright © Dell Inc. All Rights Reserved.

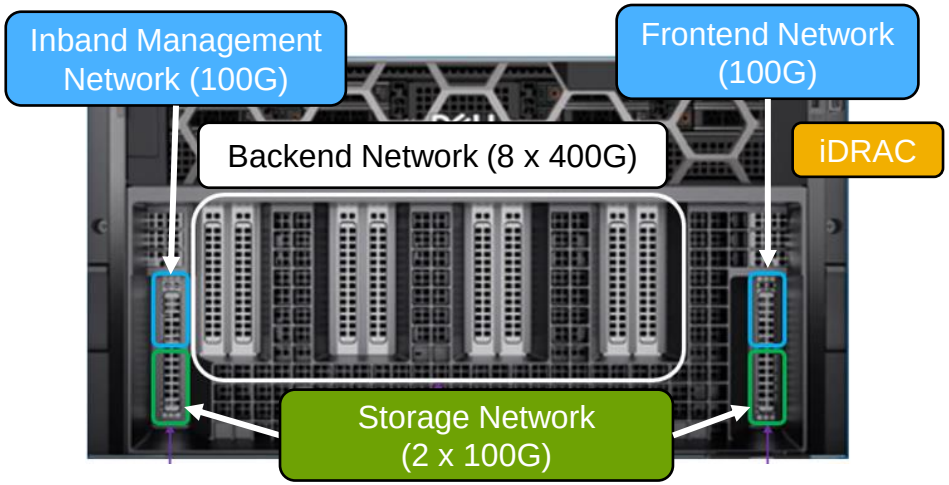
How Networking in Gen IA is different ?

DC “Classical” Network



- ❑ 2 Network Fabric @25G
- ❑ Design with oversubscription

Gen IA Network



- ❑ 3 Network Fabric @400G & 100G
- ❑ Design without oversubscription

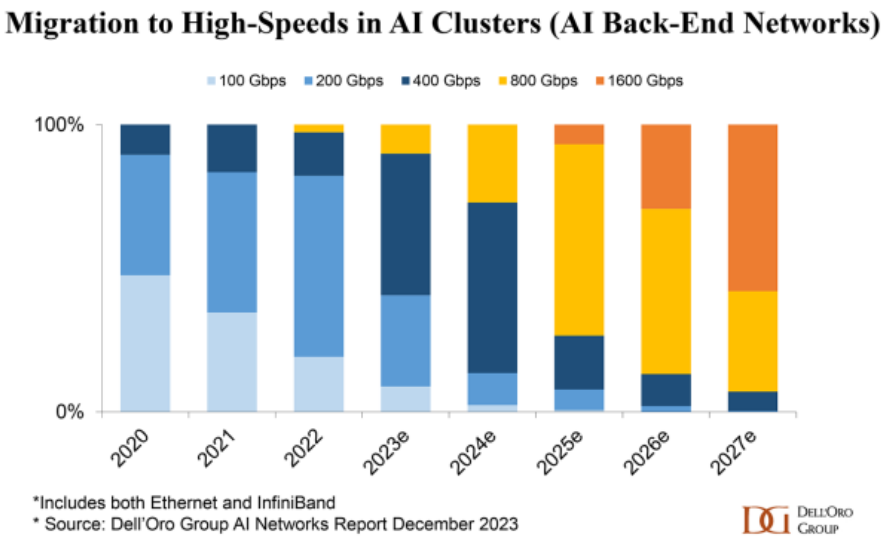
DELL Technologies

© Copyright 2021 Dell Inc.

Gen AI Fabrics - Market Insights – TAM

Data Center Ethernet Switch Projections

- Addressable TAM is more than **$100 B** in Cumulative Ethernet Data Center Switch Sales **Over Next Five Years**[1]
- **AI Fabric** (backend switching for GPU-to-GPU connectivity) forecasted to grow from **$1.2B (2022) to $15.2B (2027), five-year CAGR of 65%** [2]
- 20 Percent of Ethernet Data Center Switch Ports Will Connect to AI Servers By 2027



Back-end AI Network Projections

- InfiniBand and Ethernet will coexist
- Expect Ethernet to gain share driven by
  - Improvement on Ethernet
  - Customers desire to multi-source
- Mix of Ethernet vs InfiniBand will depend on
  - Customers distinct choices
  - Size of clusters
  - Nature of workloads

Dell'Oro Group

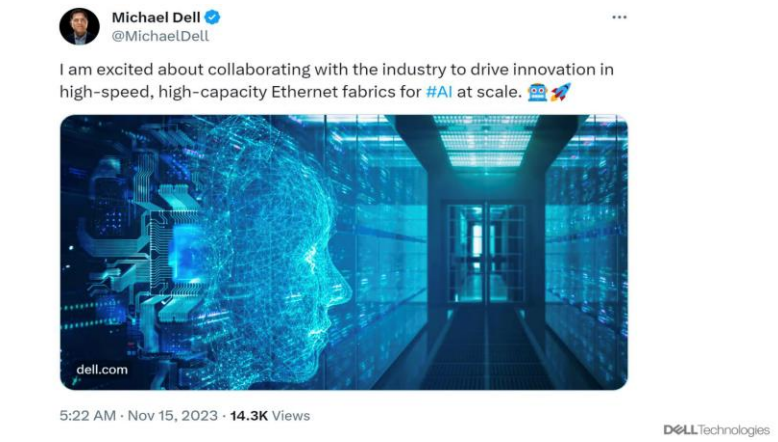
20 Percent of Ethernet Data Center Switch Ports Will Connect to AI Servers By 2027, According to Dell'Oro Group - Dell'Oro Group (delloro.com) [1]
Advanced Research Report - Networks For AI Workloads, Dell ORO Market Forecast [2]

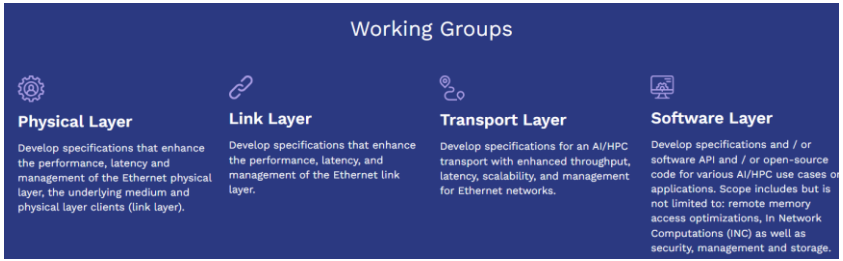
Copyright © Dell Inc. All Rights Reserved.

DELL Technologies

Ultra Ethernet Consortium

- Formed under the Linux Foundation July 2023
  - based on a private consortium created previously
- https://ultraethernet.org/





UEC Public Statements Summary

"*The UEC specifications offer significant improvements by addressing the following critical networking requirements for the next generation of applications*":

- Average message completion time & "Tail latency" should both be minimized.
- Maximize concurrency
- fully utilizing the AI network
- Avoid causing congestion or head-of-line blocking or final hop Incast
- Minimize use of centralized traffic engineering or algorithm specific tuning

By using…

- Multi-pathing and packet spraying
- Flexible delivery order
- Modern congestion control mechanisms
- Efficient and fair rate control algorithms
- End-to-end telemetry
- Larger scale, stability, and reliability (is 10^6 endpoints too many? ☺)
- Physical layer of 800G, 1.6T, and faster Ethernet speeds
- End to end security

DELL Technologies

Copyright © Dell Inc. All Rights Reserved.

Ethernet evolving to be the preferred choice for backend AI fabrics

- Market inflection points for Ethernets powered by AI fabrics
  - **Availability** of High Radix switching with next-Gen silicon technologies – 64x400G (25.6T), 64x800G(51.2T), 102.4T…
  - Improved **congestion monitoring, flow control, and Transport (RoCEv2)** protocol availability in NOS
  - Community effort to drive Ethernet Standards – **Ultra Ethernet Consortium**
  - Desire for **no-vendor lock-in** infrastructures
  - **Silicon and supplier diversity**
  - Lower **Total Cost of Ownership (~3x lower)**
  - **Latency improvements** with next Gen Silicon **from 800ns to 200ns** and **Cut-Through** switching

© Copyright 2020 Dell Inc.

Dell Networking

Delivering Ethernet Solutions across all use cases within AI Fabrics

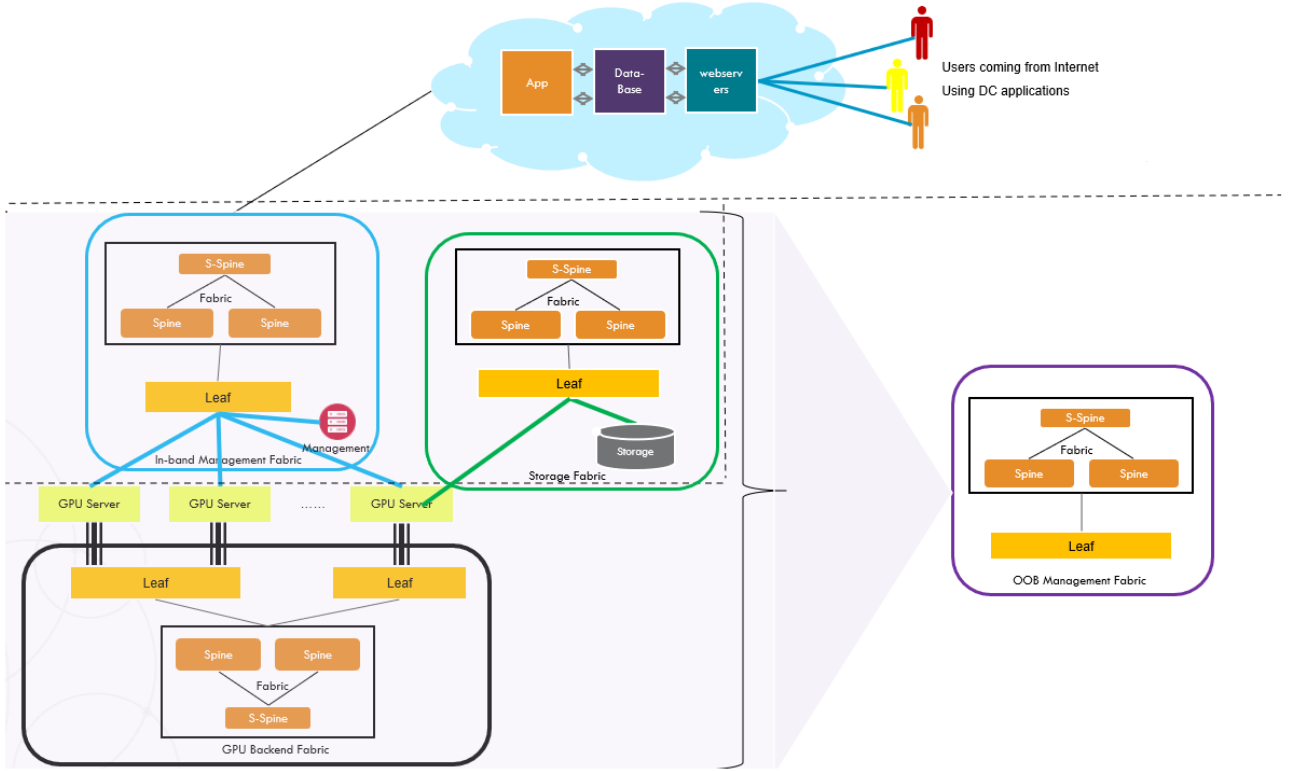
AI fabric

Bringing it ALL together

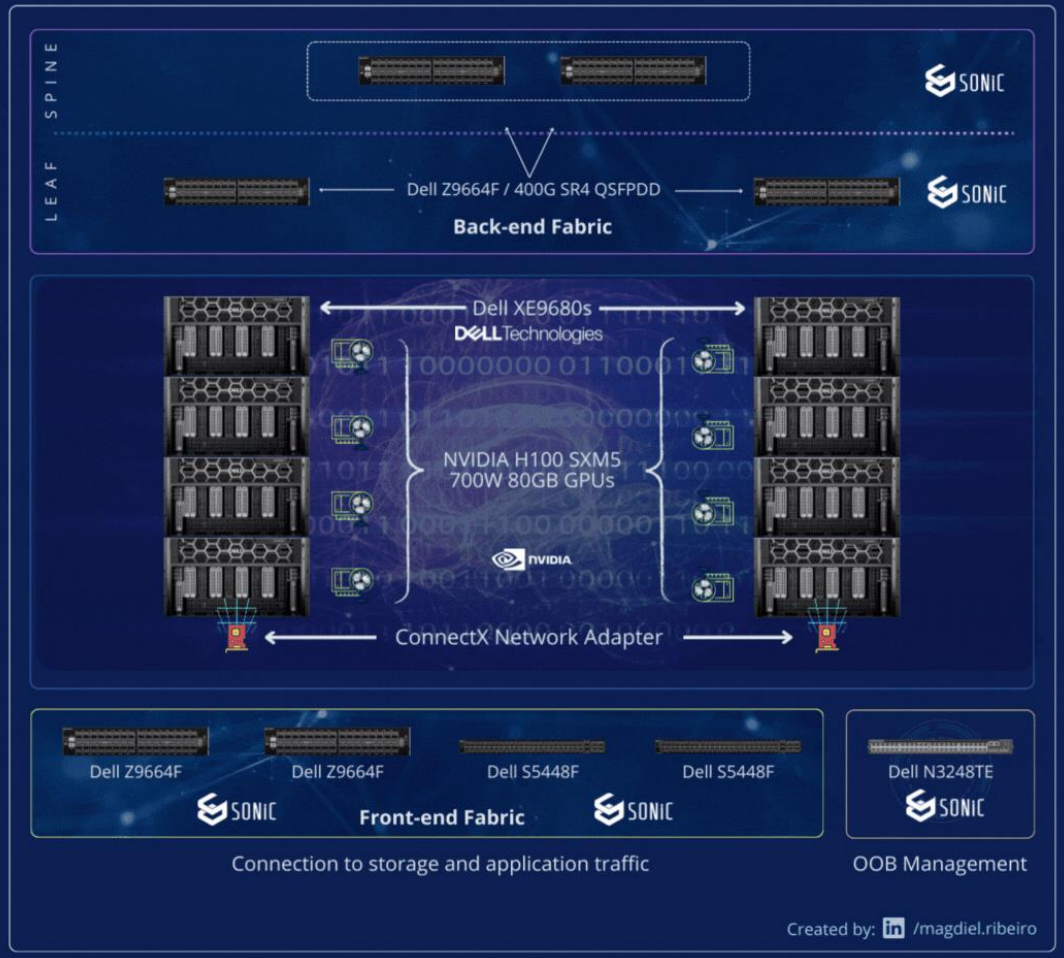
Back-End (GPU Fabric) has most demanding requirements for raw performance, lossless attributes and lowest latency

Front-End fabrics support application traffic, storage access and connection to the general network

OOB Mgmt Network for administration and fabric management



Copyright © Dell Inc. All Rights Reserved.



Dell Enterprise SONiC and GenAI use case

Copyright © Dell Inc. All Rights Reserved.

DELL Technologies

Dell Technologies World 20-23 May

Dell Networking for AI Solutions

High-bandwidth, low-latency advanced Ethernet fabrics

Z9864F-ON

2U form factor for twice the performance of the current generation PowerSwitch

SONiC

Enterprise SONiC 4.4 with SmartFabric Manager

Simplified management, validated blueprints and better resiliency with advanced AI Ethernet features

Broadcom Thor 2 NIC

2x End-to-End Bandwidth

400 Gbps bi-di throughput

Low Latency

Next-gen Ethernet fabrics for modern workloads like GenAI

2X	64	8K
density of 400GbE switching ports[1]	800GbE ports	supported GPU nodes[2]

1 vs. the previous generation 400GbE switch, the Z9664-ON
2 Two-tier fabric

High performance Ethernet for modern AI workloads

Adaptive Routing and RoCE congestion control

Detailed performance insights with advanced analytics and monitoring

Copyright © Dell Inc. All Rights Reserved.

DELL Technologies

SONiC

Open Source Networking Made Real

EVOLUTION OF NETWORKING

A move from closed & proprietary to flexible & open

Open-Source Networking

Traditional Networking

Open Networking

Open Source Networking

Containerized Apps

SONiC

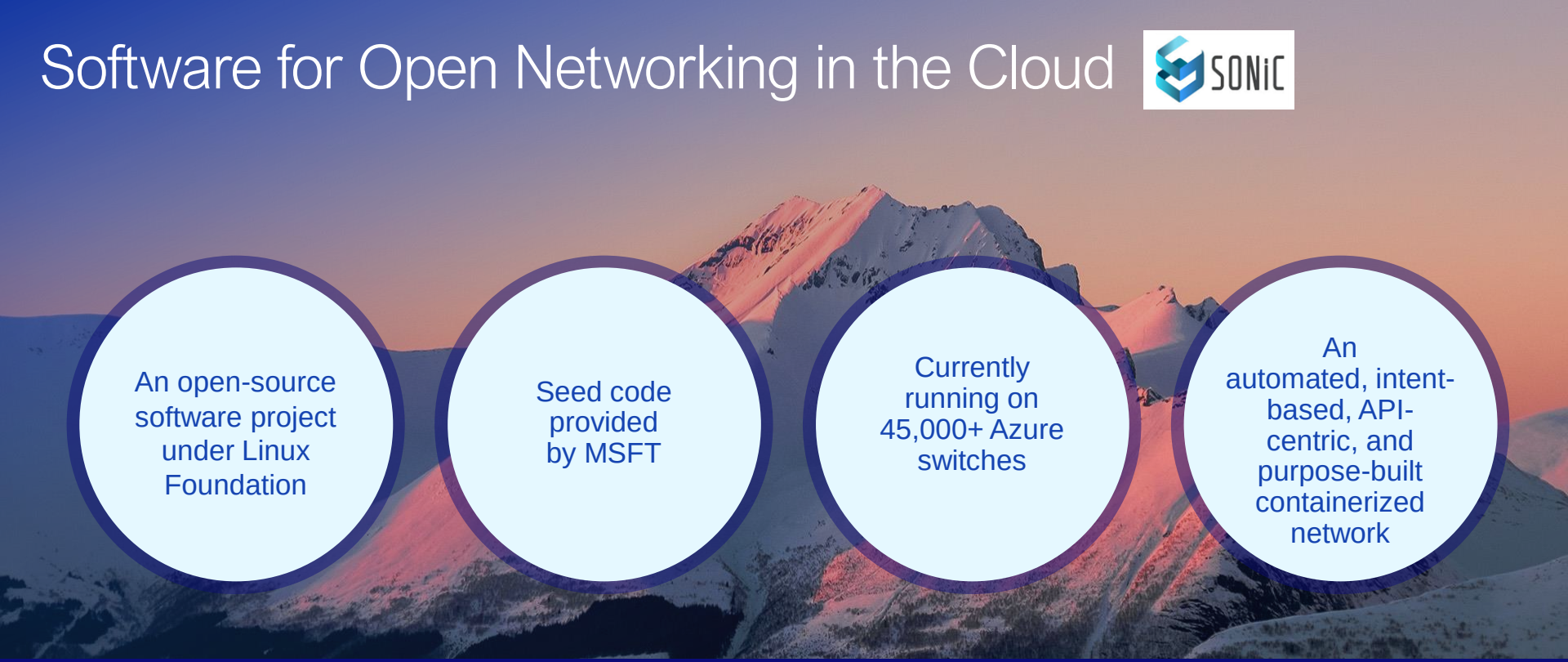
Open Networking HW Platforms

Before 2014

2014 to 2020

2020

Software for Open Networking in the Cloud



An open-source software project under Linux Foundation

Seed code provided by MSFT

Currently running on 45,000+ Azure switches

An automated, intent-based, API-centric, and purpose-built containerized network

"By 2025, 40% of organizations that operate large data center networks (more than 200 switches) will run SONiC in production environments" ~ Gartner

© Copyright 2023 Dell Inc.

SONiC Provides the Foundation for Modern Network Fabrics

Open & Extensible

with standards-based protocols and access to Linux tools & 3rd party apps.

DevOps Ready

through Containerization, automation and API-centricity.

Scalable

Created for cloud operations at scale.

Innovative

Feature velocity, advanced telemetry and deep analytics.

Copyright © Dell Inc. All Rights Reserved.

Why SONiC ?

Avoid vendor lock-in

- **Open-source** solution to **improve innovation speed** and **code reliability**
- **Open-networking** to **chose** hardware and software and **optimize cost**
- **Easy integration** to various open tools (Automation, Monitoring, Backup, Analytics ...)

Resiliency and modularity through a containerized system architecture

- Increased **security** and **stability** by **siloing** of the different functionalities within each **container**
- Possibility to **deploy home-made container**
- **Adding features** made easy
- **One solution for all use cases** depending microservices deployed (**DC, Edge, SmartNic** ...)

Standardization that allows the sustainability of investment and skills:

- Standard **Protocols**
- Standard **Architecture**
- Standard **CLI & Automation**

SONiC – Innovation of the industry > Linux of Networking

SONiC, now a Linux Foundation Project

Software for Open Networking in the Cloud (SONiC) Moves to the Linux Foundation

Leading open source network operating system enabling dis-aggregation for data centers now hosted by the Linux Foundation to enable neutral governance in a software ecosystem

SAN FRANCISCO– April 14, 2021 – Today, the Linux Foundation, the nonprofit organization enabling mass innovation through open source, announced the Software for Open Networking in the Cloud (SONiC, an open source networking operating system), is now part of the Linux Foundation. The Linux Foundation provides a venue for continued ecosystem, developer growth and diversity, as well as collaboration across the open source networking stack

- Leading Network Operating System for Disaggregated Hardware, deployed in large scale enterprise and cloud data centers globally
- Global Community of Cloud Providers, Enterprises, ODMs and Silicon Vendors
- Collaboration with Open Compute Project on SAI specifications, & new Hardware-Software co-design strategy

THE LINUX FOUNDATION

SONiC

SONiC Ecosystem

Merchant Silicon | Switch Platform | Adoption | System/Service

THE LINUX FOUNDATION

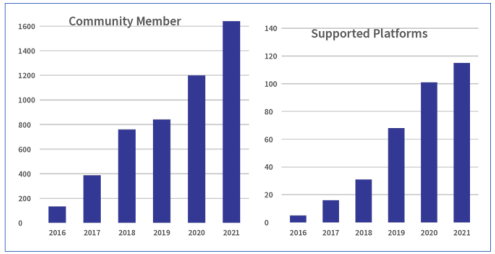
SONiC

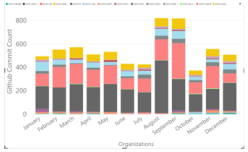
SONiC Community

2 releases/year	**400 ~ 800** commits/month
1710 active code contributors	**105** supported platform

New Subgroup on Special Topics

- PINS subgroup
- DASH subgroup
- IPSec SAI subgroup
- YANG model subgroup
- Kubernetes subgroup
- Chassis subgroup
- MPLS subgroup
- SONIC application subgroup



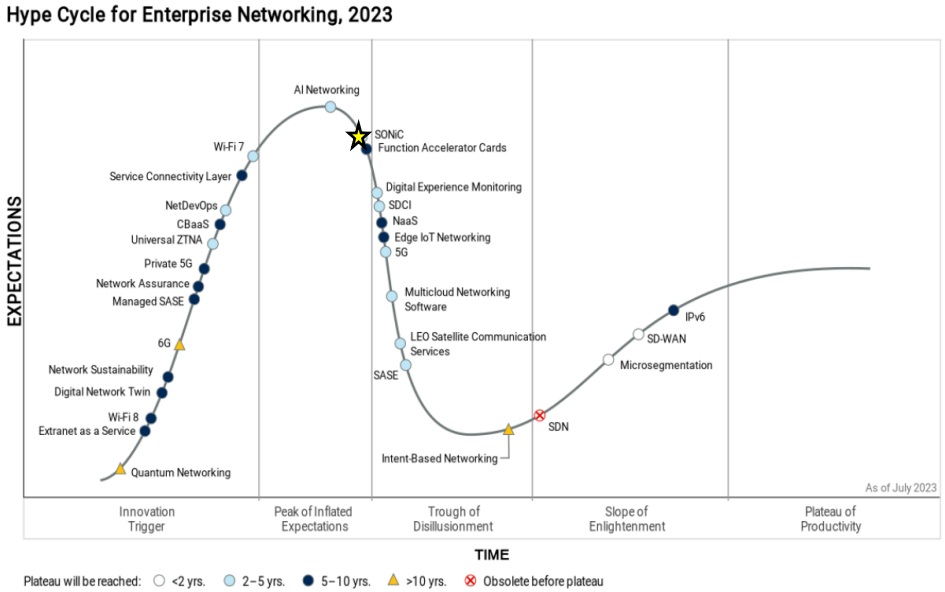


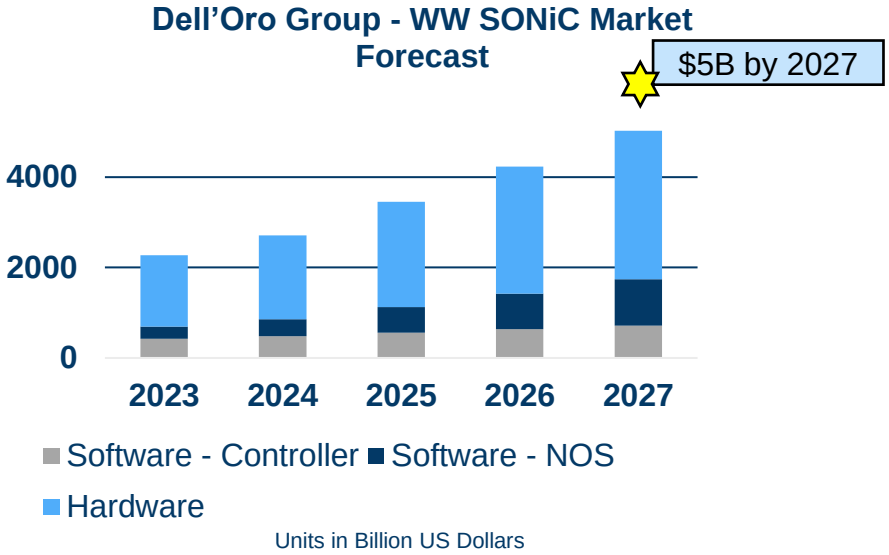
THE LINUX FOUNDATION

SONiC

Copyright © Dell Inc. All Rights Reserved.

Exponential Growth and Adoption of SONiC





Gartner

Disruption

SONiC is predicted to be a long-term market disruption driver

Innovation

SONiC promises innovation like Linux transformation of server OS Market, by removing ties w/ hardware vendors & standardizing NOS

Commercial Support Recommended

For mission-critical data centers, consider using commercially supported distributions of SONiC to ensure stability and reliability

Copyright © Dell Inc. All Rights Reserved.

Copyright © Dell Inc. All Rights Reserved.

Our Vision & Strategy for a Modernized Network Fabric

Result

Delivering customers full control and choice of the technology stack – HW, SW, and ecosystem tools, at optimal cost without vendor lock-in

Copyright © Dell Inc. All Rights Reserved.

SONiC Differentiators

Edge Use Case

E3248PXE : 48 x 10GbE PoE (90W) + 4x25GE + 2x100GE
E3248P : 48 x 1GE PoE (30W) + 4x10GE + 2x100GE

- Unique SONiC play at the Edge
- High-end access switches

Run 3rd party container

strongSwan, Fail2Ban, Asterisk, freeRADIUS, kea, SONiC

- Huge databased of containerized applications
- Extend use cases beyond switching …

Ecosystem

Grafana, telegraf, Prometheus, influxdb, ANSIBLE

Open source

BEYONDEDGE, metalsoft, dorado SOFTWARE, Augtera, racksnet

Commercial

- No software tools lock-in
- No hardware lock-in
- Consistency from Edge to DC

Summary

- We are entering the era of open source networking
- SONiC and its partners/community is leading the change: Question is not ‘IF’ but ‘**WHEN**’
- The risk is not in moving to SONiC. The risk is in **not moving** to SONiC.
- Dell Technologies is bridging the gap and bringing SONiC to the enterprise with open source innovation, enterprise and cloud data center and GenAI readiness and global support and services on Dell EMC PowerSwitch and Open HW platforms

DELL Technologies

This is the way.
Disney+