HiWire™ SWITCH AEC and SONiC® Dual TOR Webinar

Presented by Credo with Microsoft

Hosted by

Today's Host

HiWire Switch AEC and SONiC Dual TOR Webinar

650 GROUP
MARKET INTELLIGENCE RESEARCH

Alan Weckel | 650 Group

is a Technology Analyst/Co-Founder at 650 Group, where he is in charge of Networking and Cloud research. Alan's expertise has been quoted in CIO Today, Wall Street Journal, and Fierce Telecom. He has presented at a wide variety of industry and finance events. His work at previous companies including Cisco Systems and Raytheon provides a foundation for his deep knowledge of the industry and its supply chain.

Microsoft SONiC 650 GROUP MARKET INTELLIGENCE RESEARCH CREDO we connect. HiWire Active Electrical Cable

650 Group Overview

HiWire Switch AEC and SONiC Dual TOR Webinar

650 GROUP
MARKET INTELLIGENCE RESEARCH

What We Track

Cloud
- IaaS, SaaS, Colo, Search and Social
- CAPEX, Equipment Trends

Telecom Equipment
- Broadband Access, Telecom Core
- NFV, Mobile RAN, SP Routing, Optical Transport

Data Center Equipment
- Data Center Switching, Servers, Storage
- Merchant Silicon, DCI, Security

Enterprise Networking Equipment
- Switching, WLAN, Security
- Enhanced NAC, Unified Access, SD-WAN

Semiconductors and FABs
- Campus, DC, Cloud, IoT Semis
- ASIC trends in Switching, WLAN

Who Uses Our Research

Buy Side / Sell Side

Cloud Providers

Telco SPs

System Vendors

Semiconductor Suppliers

Component Manufacturers

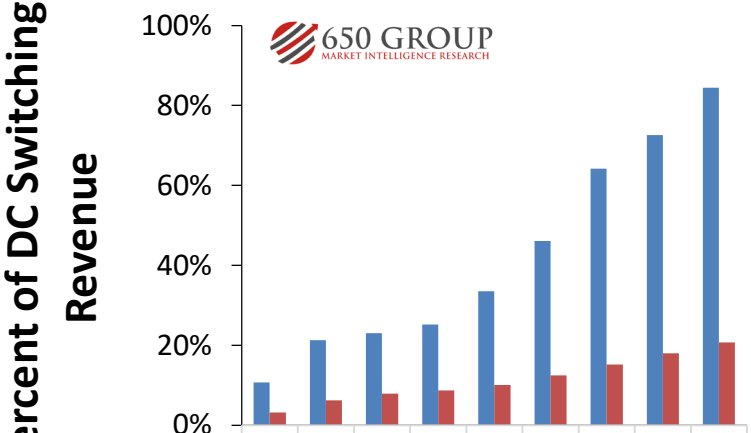
FABs

Microsoft SONiC 650 GROUP MARKET INTELLIGENCE RESEARCH CREDO we connect. HiWire Active Electrical Cable

Ethernet Switch – Data Center: SONiC Projections

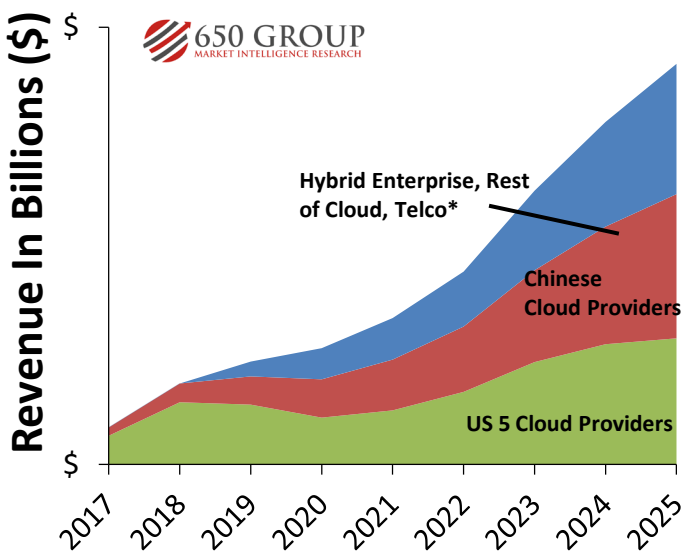
HiWire Switch AEC and SONiC Dual TOR Webinar

SONiC Capable Ports



	2017	2018	2019	2020	2021	2022	2023	2024	2025
SONiC Capable	11%	21%	23%	25%	34%	46%	64%	73%	84%
Running SONiC	3%	6%	8%	9%	10%	12%	15%	18%	21%

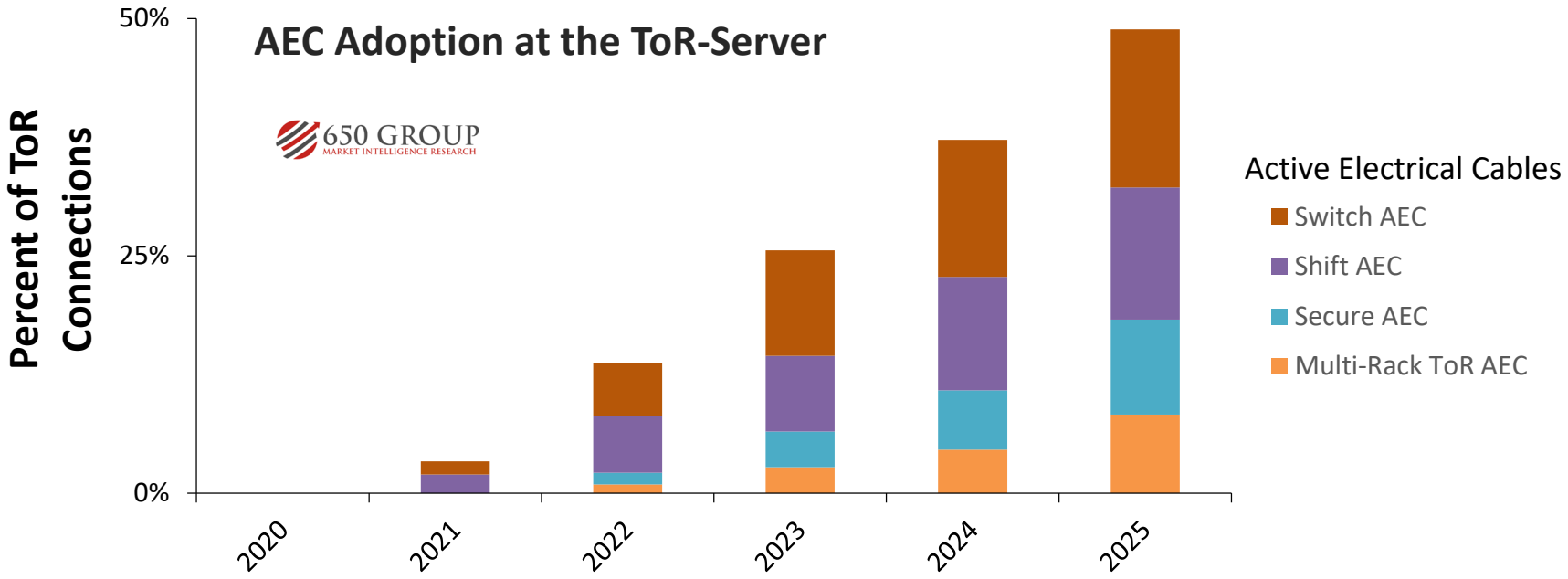
SONiC Revenue by Customer Type



*Routing is not included in these projections; especially impactful in Telco Cloud.

Ethernet Switch – Data Center

HiWire Switch AEC and SONiC Dual TOR Webinar



Microsoft SONiC 650 GROUP MARKET INTELLIGENCE RESEARCH CREDO we connect. HiWire Active Electrical Cable

Today's Speakers

HiWire Switch AEC and SONiC Dual TOR Webinar

Microsoft

SONiC

CREDO we connect.

HiWire™ Active Electrical Cable

Dr. Lihua Yuan | Microsoft

is currently a Partner Software Engineering Manager in Microsoft Azure Networking. He leads the network platform team supporting the data center network connecting major businesses including Azure, Microsoft 365, and Bing.

Don Barnetson | Credo

is the VP of Product for HiWire AECs at Credo with a focus on product definition and customer engagement for this new, exciting category.

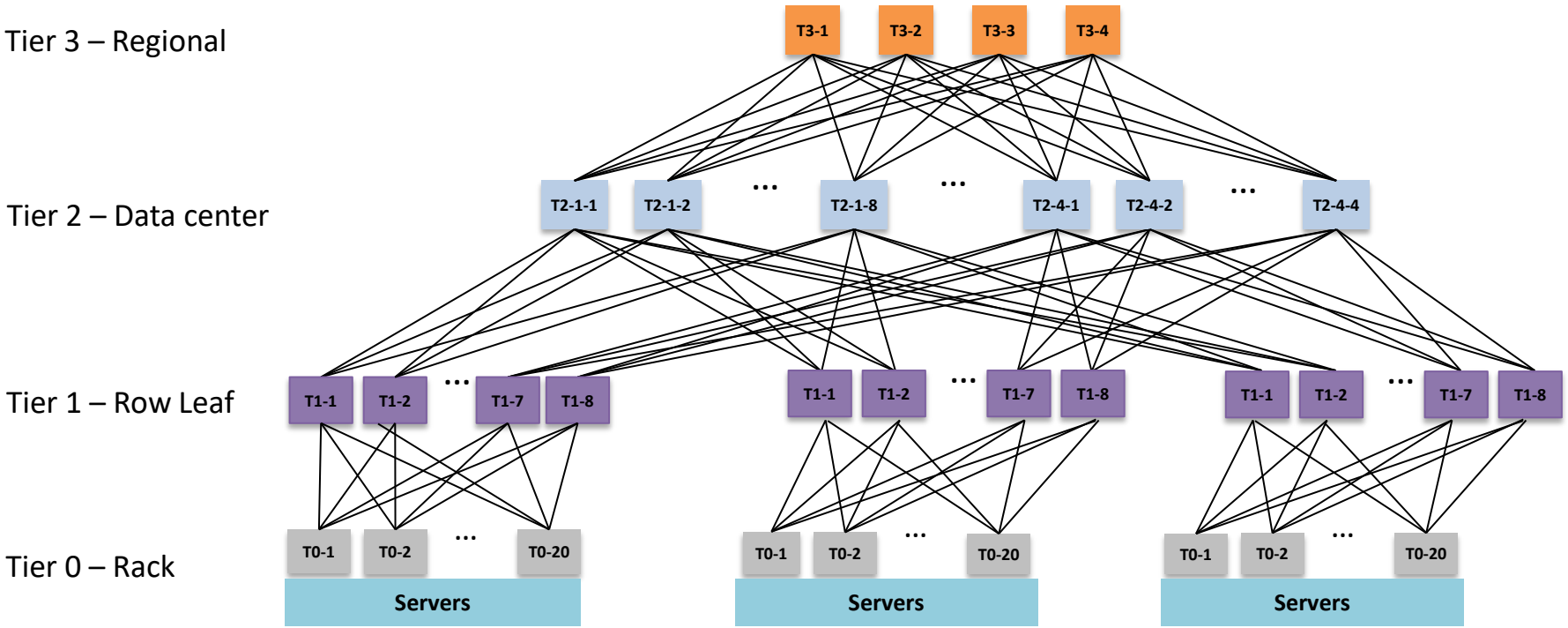
Please ask questions in the Question Box on the right
Q&A at the end of the Presentation

Microsoft SONiC 650 GROUP MARKET INTELLIGENCE RESEARCH CREDO we connect. HiWire™ Active Electrical Cable

Dual ToR Technology and Challenges

Data Center Network with CLOS

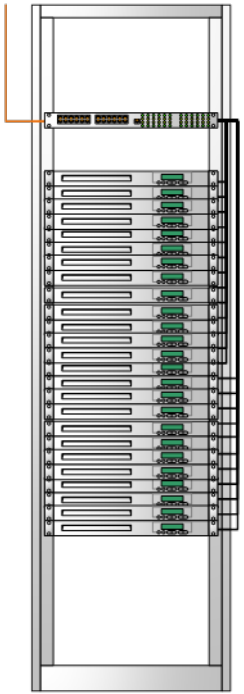
HiWire Switch AEC and SONiC Dual TOR Webinar



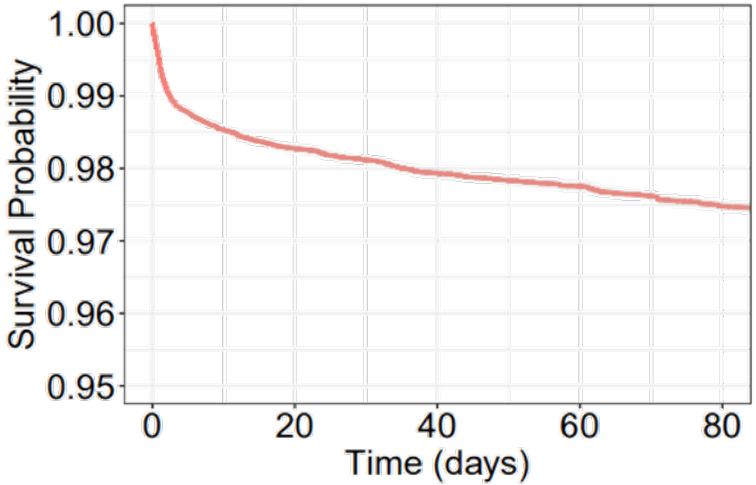
Microsoft SONiC 650 GROUP MARKET INTELLIGENCE RESEARCH CREDO we connect. HiWire Active Electrical Cable

Top of Rack Switch : Single Point of Failure for a Rack

HiWire Switch AEC and SONiC Dual TOR Webinar



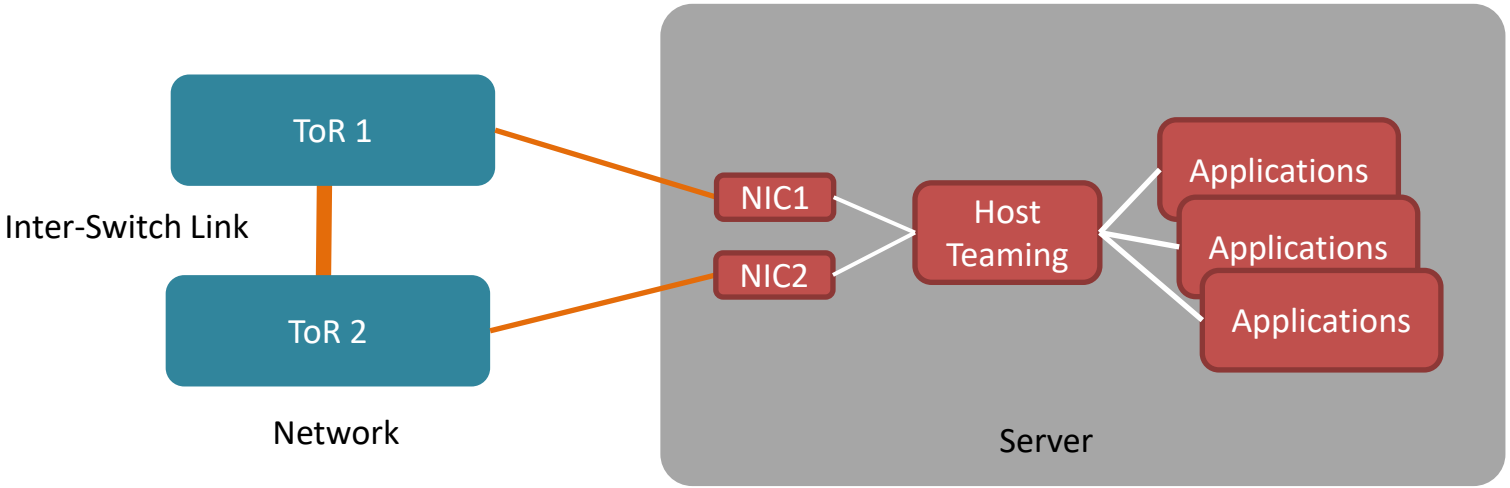
- ToR is a Single-Point-of-Failure (SPOF) for full rack of servers
- And TORs do fail
  - ~2% of switches fail in first 3 months
  - 32% due to hardware failures
  - 27% due to power failures



Microsoft SONiC 650 GROUP MARKET INTELLIGENCE RESEARCH CREDO we connect. HiWire Active Electrical Cable

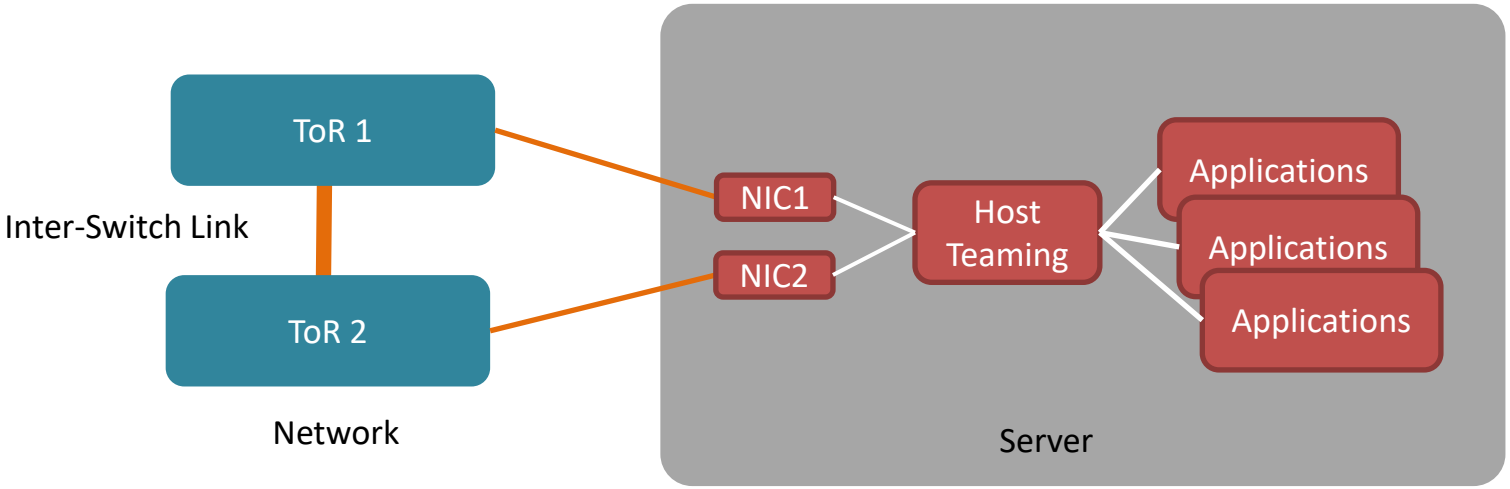
Classic Solution: **Multi-Chassis LAG with Dual Uplink**

HiWire Switch AEC and SONiC Dual TOR Webinar



Network: **Inter-Switch Link**

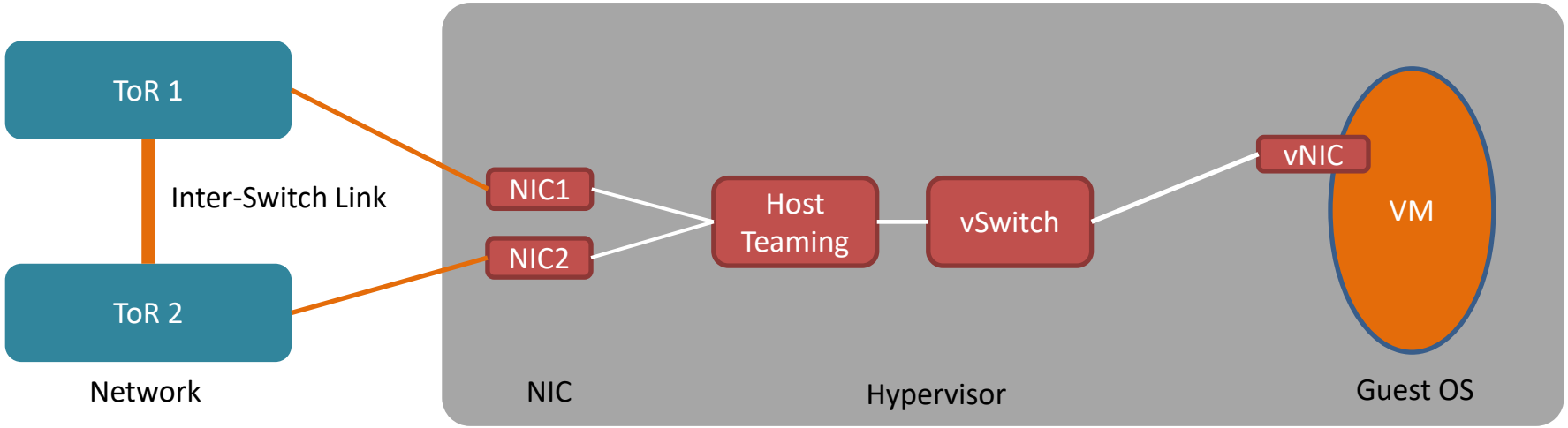
HiWire Switch AEC and SONiC Dual TOR Webinar



- Inter-Switch Link requires custom design for capacity planning
- Requires complex state sync between ToRs
  - Creates split-brain problem when ISL fails

HyperVisor: **Performance Limit**

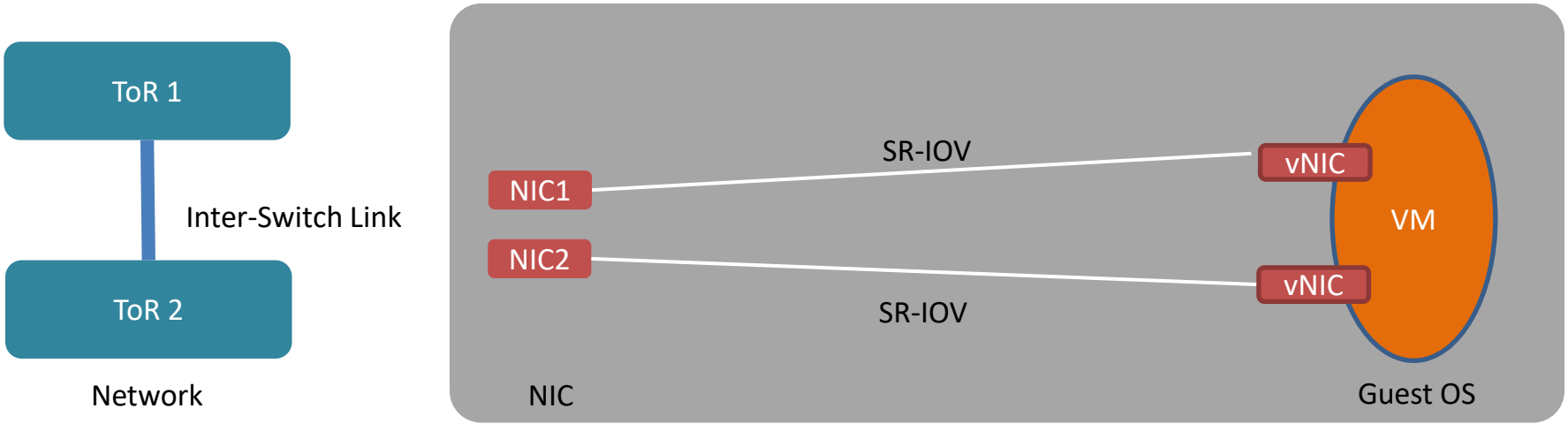
HiWire Switch AEC and SONiC Dual TOR Webinar



- A few Gbps/core
- Unstable latency
- Won't work for RDMA

SR-IOV: **VM Complexity**

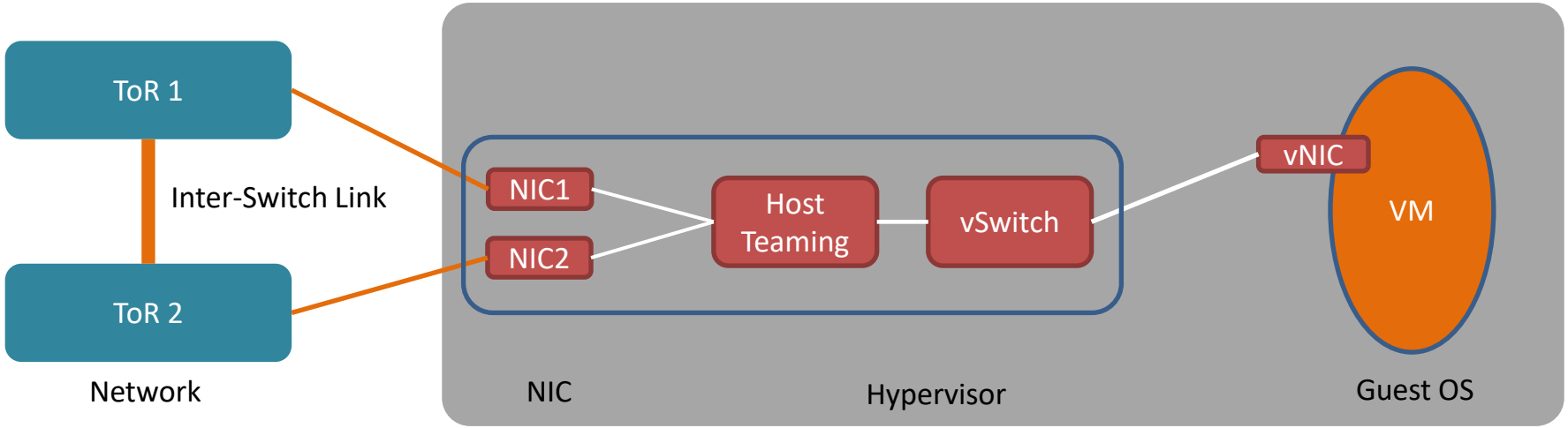
HiWire Switch AEC and SONiC Dual TOR Webinar



- VM sees failures and must handle
- Hardware-dependent vNIC driver

vSwitch Offloading: **Vendor Specific**

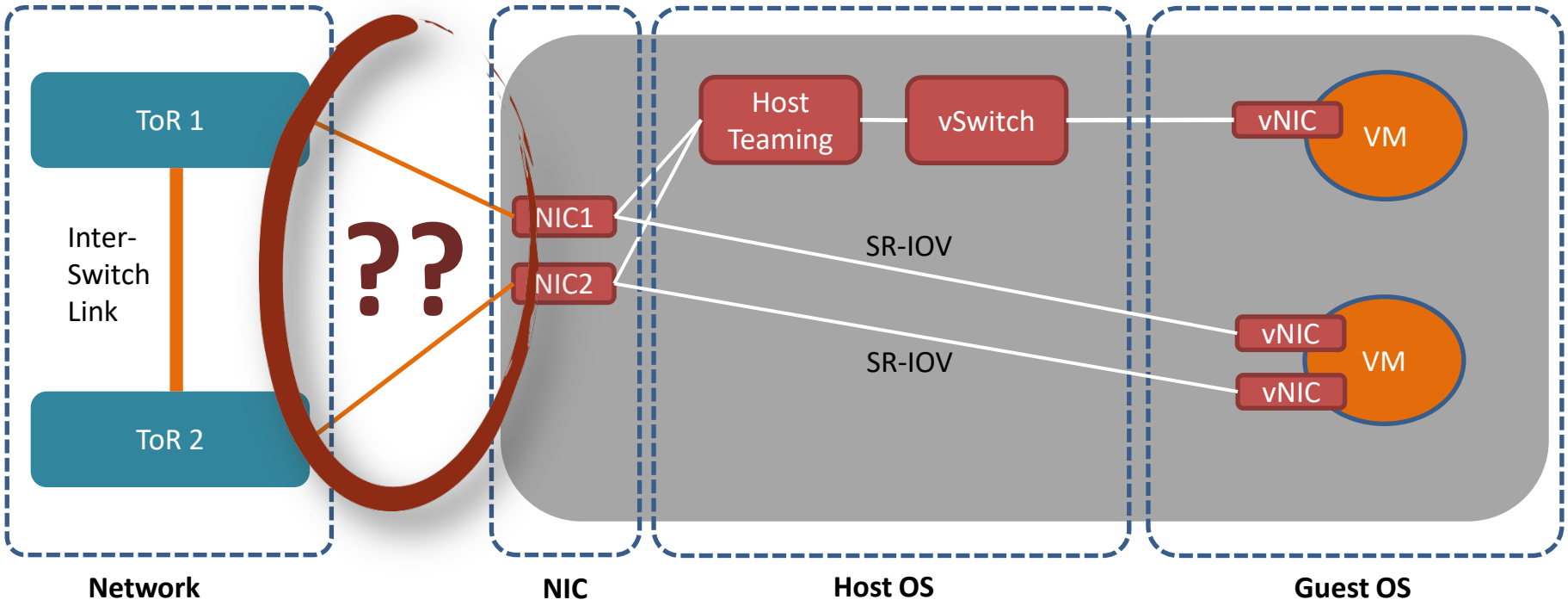
HiWire Switch AEC and SONiC Dual TOR Webinar



- Variety of solutions to offload Teaming and vSwitch to Smart NICs
- Capabilities and Implementations vary by NIC Vendor
- Require NIC vendor drivers in VM -> HW Dependency

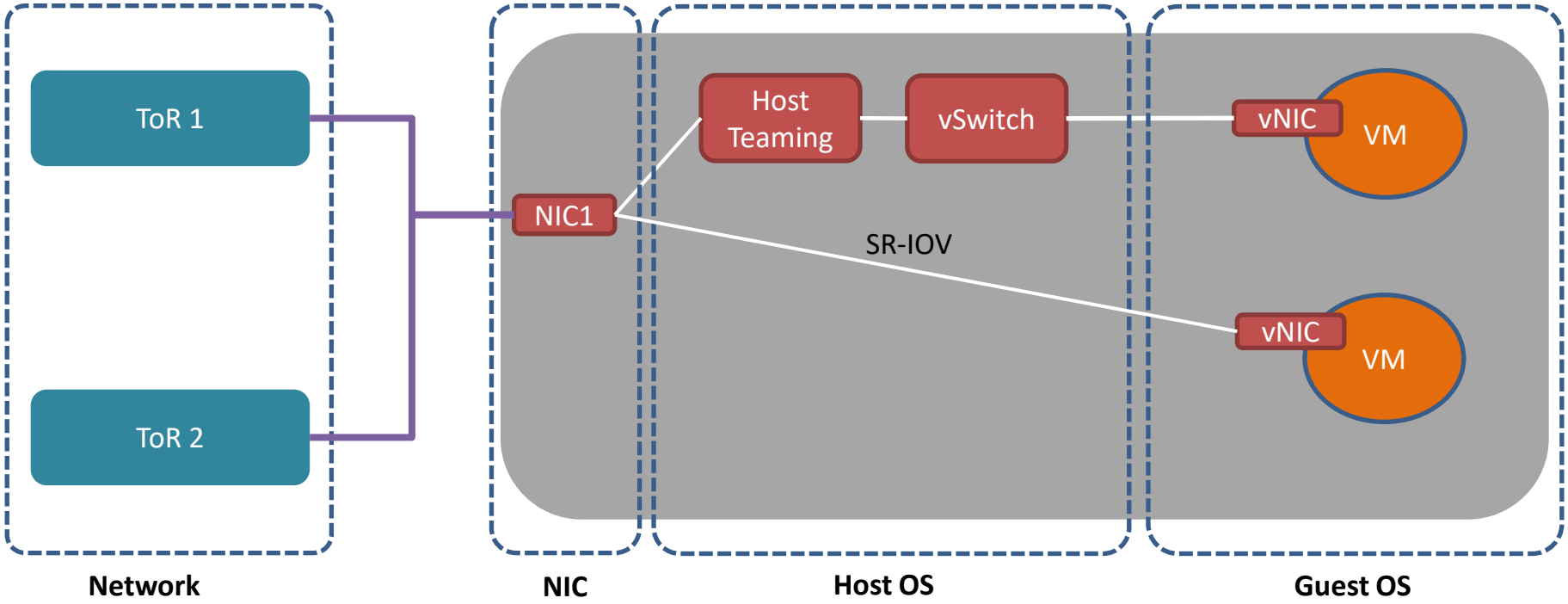
The Solution Space - Options to Manage Redundancy

HiWire Switch AEC and SONiC Dual TOR Webinar



The Solution Space – Introducing the Switch AEC

HiWire Switch AEC and SONiC Dual TOR Webinar



A New Approach:
HiWire™ SWITCH AEC +
SONiC Dual TOR Management Container

What are HiWire™ Active Electrical Cables (AECs)

HiWire Switch AEC and SONiC Dual TOR Webinar

AECs are industry-backed solutions to solve the connectivity bottleneck

- Integrate active components : retimer, gearbox, L2 Switch into Copper Cables to facilitate secure, high-integrity plug and play connectivity and mode/speed conversion.
- Compared to DACs
  - Up to 75% less weight and Volume
  - Tighter bend radius
  - Longer reach
- Compared to AOCs
  - Up to 50% less power
  - Up to 50% lower cost
  - 2.5x longer life

LP SWITCH/SWITCH AEC

Designed to enable Network Managed NIC to dual TOR connectivity in an Active/Standby configuration.

LP CLOS AEC

Designed to replace backplanes in chassis with front panel interconnect used in Distributed, Disaggregated Chassis (DDC) applications.

LP SHIFT / SHIFT AEC

Deliver plug and play connectivity between PAM4 ports and NRZ NICs with speed shifting and FEC termination in-cable.

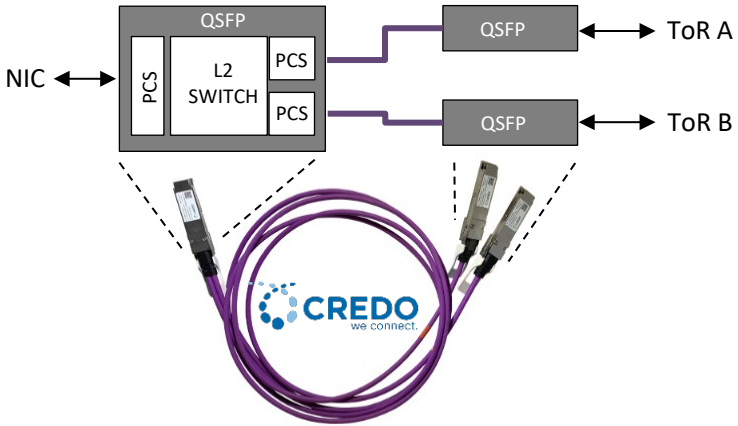
LP SPAN/SPAN AEC

Replace AOCs with low power plug and play AECs for port-to-port and rack-to-rack connectivity.

Microsoft SONiC 650 GROUP MARKET INTELLIGENCE RESEARCH CREDO we connect. HiWire™ Active Electrical Cable

New Solution in Collaboration with Microsoft / Azure

HiWire Switch AEC and SONiC Dual TOR Webinar



Dual TOR Management Container

SONiC

Microsoft Azure

HiWire SWITCH AEC: 50G/100G/200G QSFP AEC Cable

- Active/Standby Layer 2 switch – switching in <1 µs
- Common control plane on all 3 ends
- Fully ToR managed, works in standard QSFP NICs

Dual TOR Management Container on SONiC

- Manages the SWITCH AEC
- Manages convergence in failover conditions using standards based ARP/BGP/Encap & forward
- Free and Open Source

Microsoft SONiC 650 GROUP MARKET INTELLIGENCE RESEARCH CREDO we connect. HiWire Active Electrical Cable

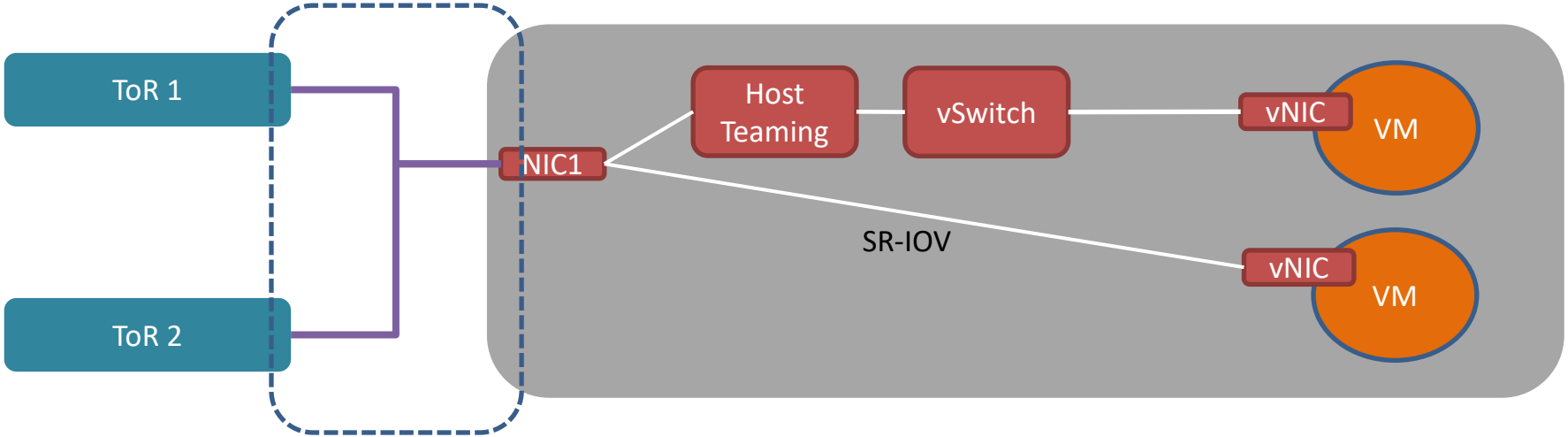
HiWire SWITCH AEC

HiWire Switch AEC and SONiC Dual TOR Webinar

Performance	• 50G – 200G NRZ or PAM4
Switching	• Automatic switching on loss-of-signal and Manually switching on I^2C command
Updates	• Hitless Firmware Updates
Functionality	• Gearbox, FEC translation
AEC Classification	• Networking device, not a DAC
Deployment	• Added at integration to preserve flexibility right up to deployment
Roadmap	• line rate encryption

The Solution Space – Simplicity & Performance

HiWire Switch AEC and SONiC Dual TOR Webinar



HiWire Switch AEC Interconnect

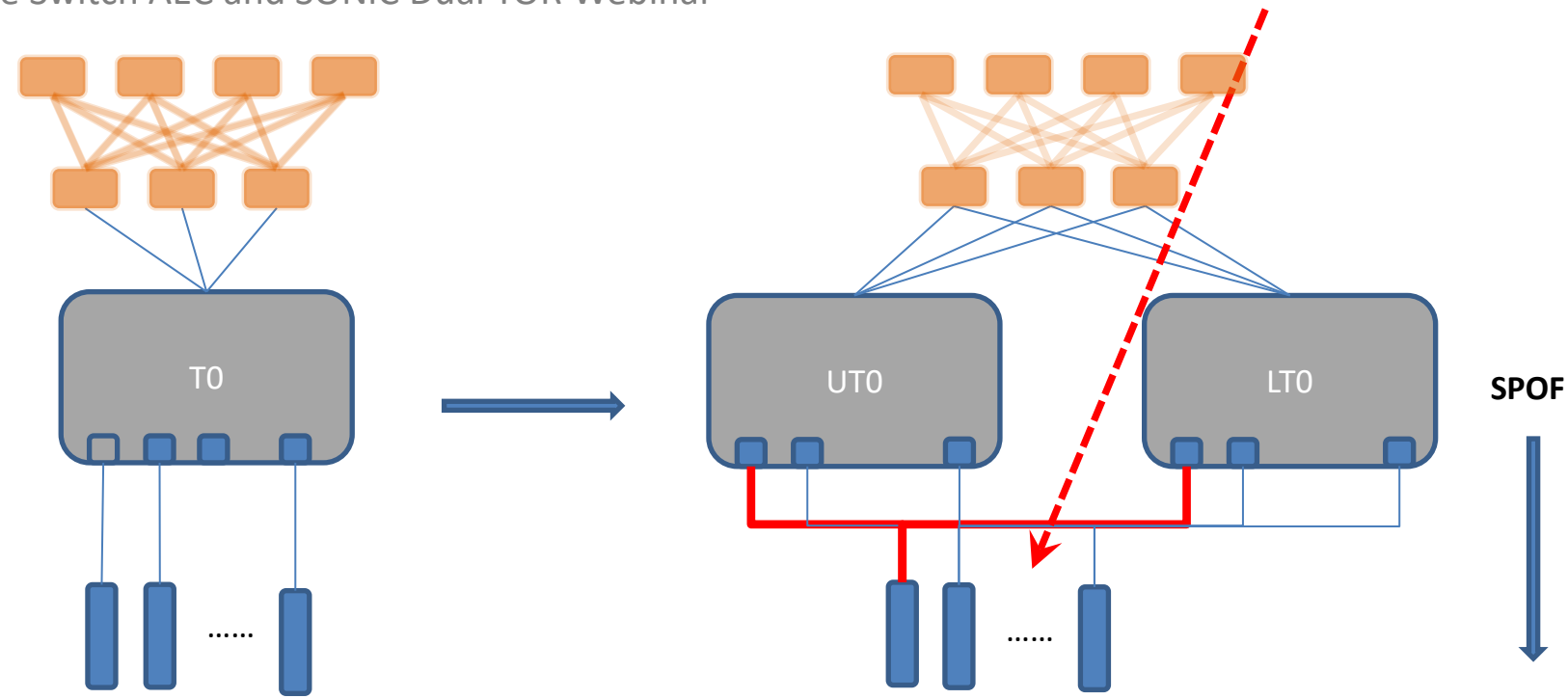
- Simpler than Link Aggregation
- Faster than Host Teaming
- Decouples Server & Network Link Management

Microsoft SONiC 650 GROUP MARKET INTELLIGENCE RESEARCH CREDO we connect. HiWire Active Electrical Cable

How do we deploy?

Dual ToR with Switch AEC

HiWire Switch AEC and SONiC Dual TOR Webinar



Single TOR : Failure Domain: Rack (20 – 40 Servers)

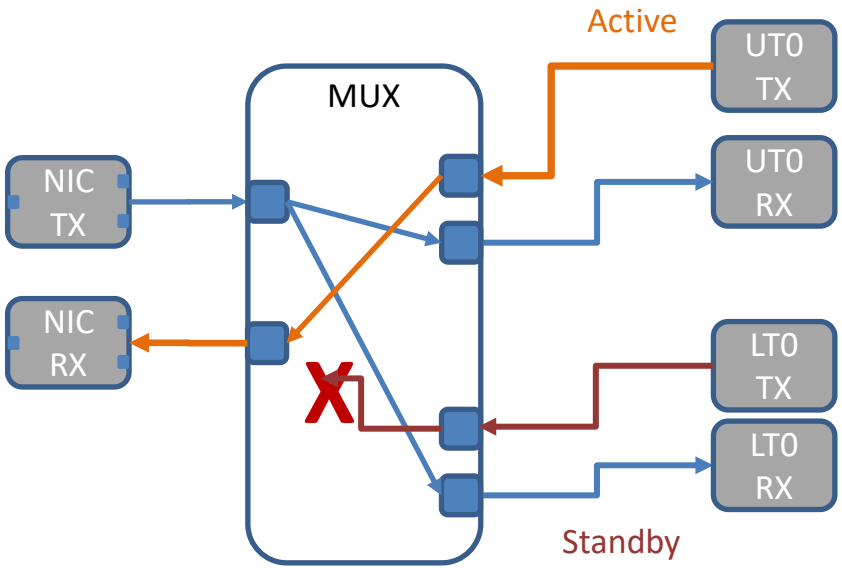
Dual TOR + Switch AEC : Failure Domain: Single Server

Switch AEC Data Plane Behavior

HiWire Switch AEC and SONiC Dual TOR Webinar

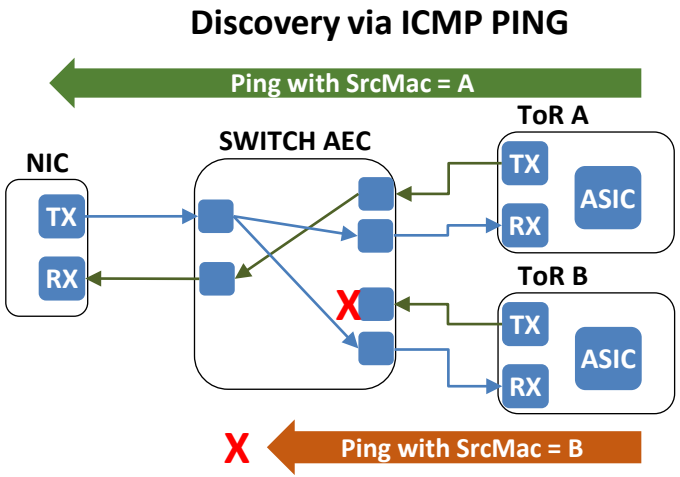
NIC TX is always broadcasted

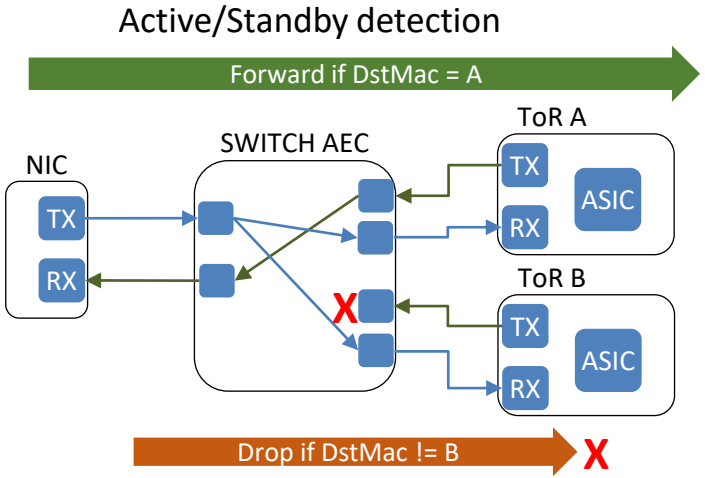
NIC RX will take active side and only Switch with Loss-of-signal or on-command over time < 1ms



Independent Active/Standby Design

HiWire Switch AEC and SONiC Dual TOR Webinar





Southbound Traffic is MUX'd

- Both ToRs both ping NIC,
- ToR A ping is forwarded to NIC
- ToR B ping is dropped by SWITCH AEC
- NIC learns ToR A as destination MAC

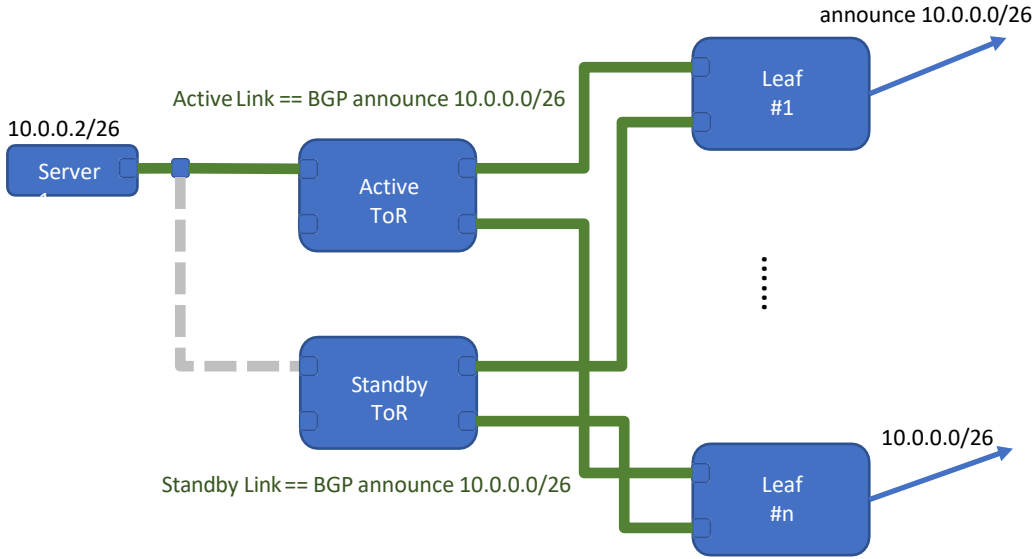
Northbound traffic broadcasted to both ToRs

- ToR A:
  - Forward north based on DstMac correctness
  - Verifies link integrity
- ToR B:
  - Drop packets due to wrong DstMac
  - Sniffs to verify ToR A's link integrity

Microsoft SONiC 650 GROUP MARKET INTELLIGENCE RESEARCH CREDO we connect. HiWire Active Electrical Cable

Routing Behavior – **Both ToRs announce via BGP**

HiWire Switch AEC and SONiC Dual TOR Webinar



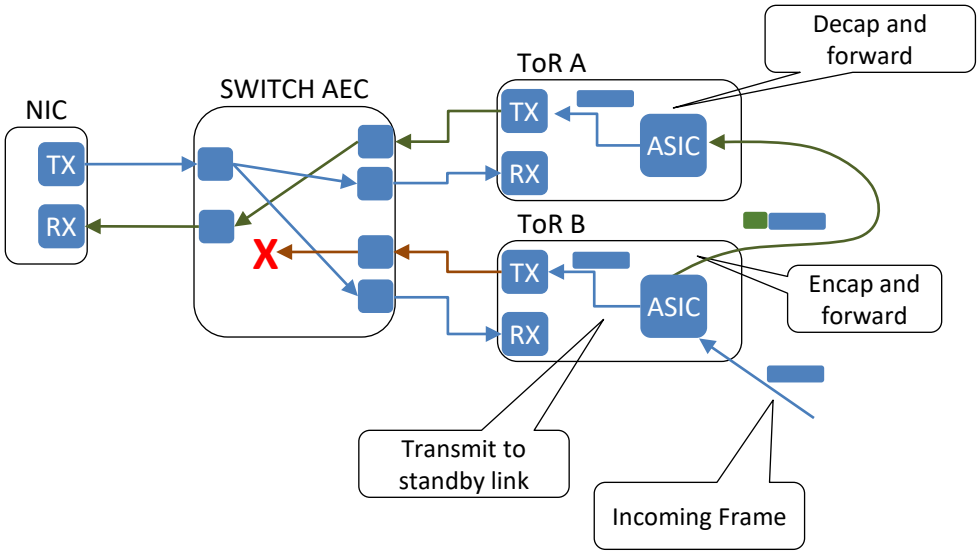
Both ToRs see Server (10.0.0.2/26) and announce to Leaf Nodes

All Leaf nodes see ToR announcements and announce to high level via standard BGP

Southbound traffic can arrive at either ToR

Southbound Packet Path

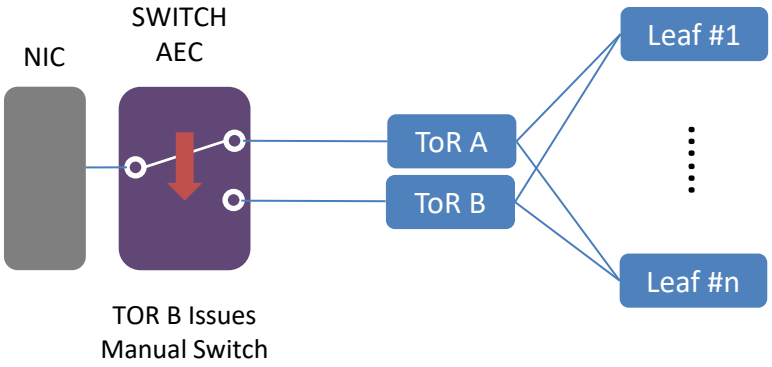
HiWire Switch AEC and SONiC Dual TOR Webinar



- Southbound traffic arriving on active ToR A is forwarded to NIC
- Southbound traffic arriving on Standby ToR B
  - ToR B encaps and transmits to ToR A
  - Tunneled through leaf instead of ISL
  - ToR A decaps and forwards to NIC
- Encap/decap handled by ASIC with minimum overhead

Failure Scenarios 1: **Planned Maintenance**

HiWire Switch AEC and SONiC Dual TOR Webinar



Scenarios:

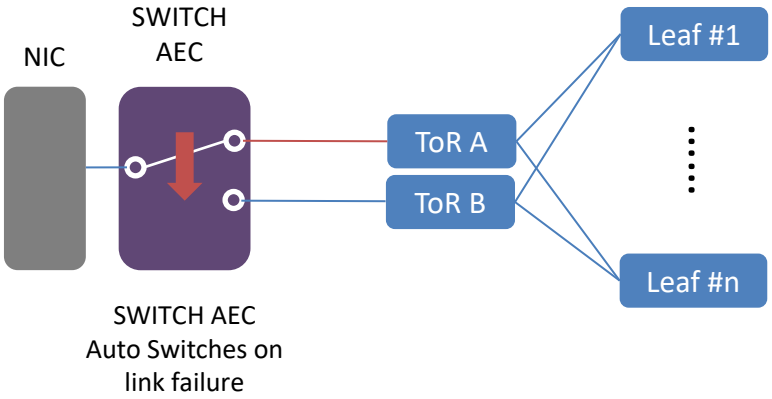
ToR OS upgrade, hardware replacement

1. ToR B proactively issues command to switch MUX to ToR B
2. ToR B ICMP Ping is now forwarded; ToR A is blocked
3. NIC updates DstMaC to ToR B
4. ToR B becomes active, ToR A become standby

Convergence time < 100ms

Failure Scenarios 2: **Hardware Failure for ToR and Link**

HiWire Switch AEC and SONiC Dual TOR Webinar



Deployment Scenario:

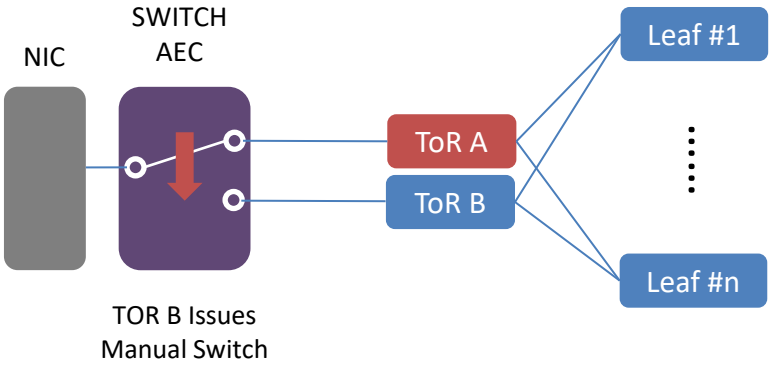
Cable cut, ToR port failure, ToR power failure, ToR

1. Cable detect Loss-of-sigal to ToR A
2. Cable auto switches on link failure in less than 1µs
3. ToR B ICMP is now forwarded; ToR A is blocked
4. NIC changes DstMac address
5. ToR assumes Active role

Convergence time < 100ms

Failure Scenarios 3: **ToR Forwarding Failure**

HiWire Switch AEC and SONiC Dual TOR Webinar



Deployment Scenario:

ToR grey failure impacts forwarding, but link remained up. Actually ~26 scenarios here – review SONiC Dual TOR Container docs

1. ToR B times out on sniffed ToR A Pings
2. ToR B Initiates Manual Switch of MUX
3. ToR B ICMP is now forwarded; ToR A is blocked
4. NIC changes DstMac address
5. ToR assumes Active role

Convergence in <100ms

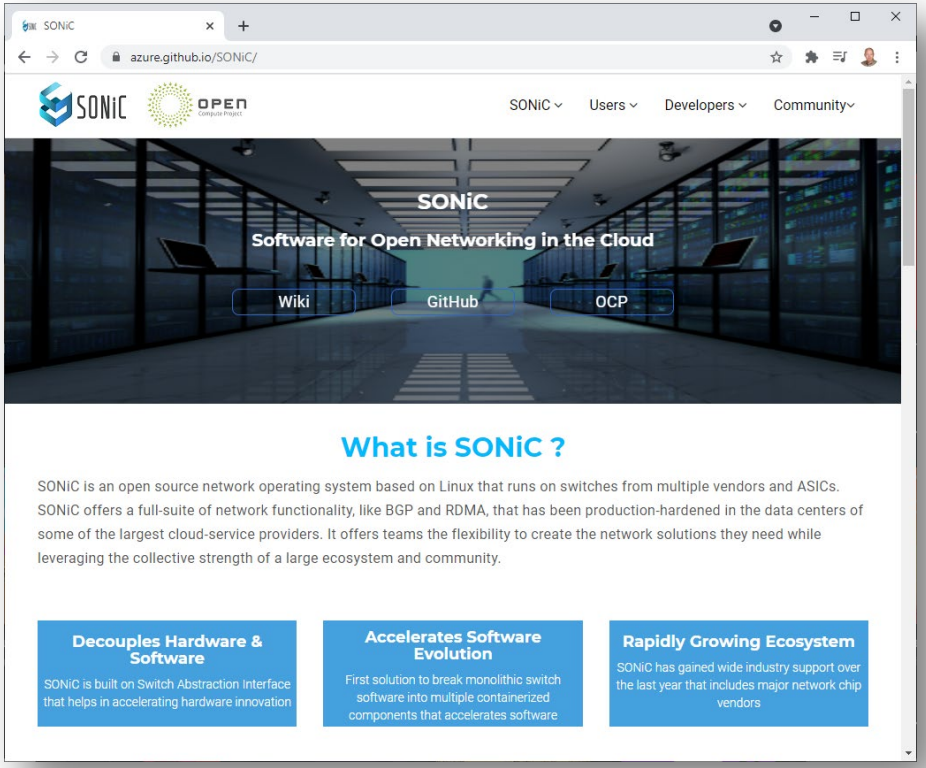
Demo Video

Terminal screen recording – end to end setup –
no packet lost in middle

SONiC Support for HiWire SWITCH AEC

HiWire Switch AEC and SONiC Dual TOR Webinar

- SWITCH AEC fully upstreamed and abstracted in SONiC
  - Support for any SONiC hardware
  - Hitless firmware updates from any cable end
  - Advanced Cable telemetry, loopback, BER and debug
- 50G and 100G HiWire™ SWITCH AEC GA now!
- SONiC Dual TOR Management Container will be released in master SONiC branch after July 21
- HiWire SWITCH AEC + Dual TOR Management Container work out of box with SONiC



Time for Q&A

Thank You