



SRv6 Deployments in Scale-Out Networks

Clarence Filsfils, Cisco Systems

Guohan Lu, OpenAI



SRv6 @ OCP

Networking Track - Thursday, Oct 15, 2026

10:30 AM	10:45 AM	SRv6 Powered Backbone Networks	Andrey Golvanov (Meta) and Jasmeet Bagga (Meta)
1:15 PM	1:30 PM	SRv6 Deployments in Scale-Out Networks	Clarence Filsfils (Cisco) and Guohan Lu (OpenAI)
01:35 PM	1:50 PM	SRv6 Traffic Engineering in AI DCI/WAN for AI Traffic	Clarence Filsfils (Cisco) and Yuanchao Su (Alibaba)
02:45 PM	03:05 PM	Clustermapper: Deterministic path diagnosis for AI backend networks using SRv6	Ryder Lewis (OpenAI) and Guohan Lu (OpenAI)
03:10 PM	03:30 AM	Beyond Traditional BGP: Decoupling Route Distribution for Large-Scale AI Clusters at Alibaba	Eddie Ruan (Alibaba) and Dennis Cai (Alibaba)

Legacy protocol stack

- A myriad of per-domain shim layers to enrich IP
 - InfiniBand in AI/HPC
 - VxLAN in the DC
 - MPLS in the core
- Proprietary solutions & vendor lock-in
- Low scale
 - No summarization in MPLS
- Poor End-to-End Services and High Operation Cost
 - Neither TE nor stateless service chaining in VxLAN
 - Translation gateways between domain (VxLAN/MPLS)

Programmable IPv6

- 128-bit DA becomes:
 - Ordered list of SRv6 uSID instructions (uSID)
 - SRv6 uSID instruction bound to anything .. Underlay TE, overlay, service chaining
 - 6 SRv6 uSID instructions in DA
 - All current deployments require less than 6 uSID
- SRv6 uSID
 - Outperforms IPv4 shim layers: MPLS, VxLAN, GTP
 - Delivers Any Service, End-to-End, Across domains
 - Standard, Open, Interoperable

End-to-end control from the application

- SRv6 empowers applications to select their path
- Deterministic path for GPU-to-GPU
 - OpenAI MRC AI backend [[Public Reference](#)]
- Stateless firewall service insertion
 - Nebius AI Frontend [[Public Reference](#)]



OCP
GLOBAL
SUMMIT

2026

Scaling Innovation for the AI Era.

Cisco Confidential

Scale out

*SRv6 for Deterministic place placement to achieve
Optimal Load-balancing*



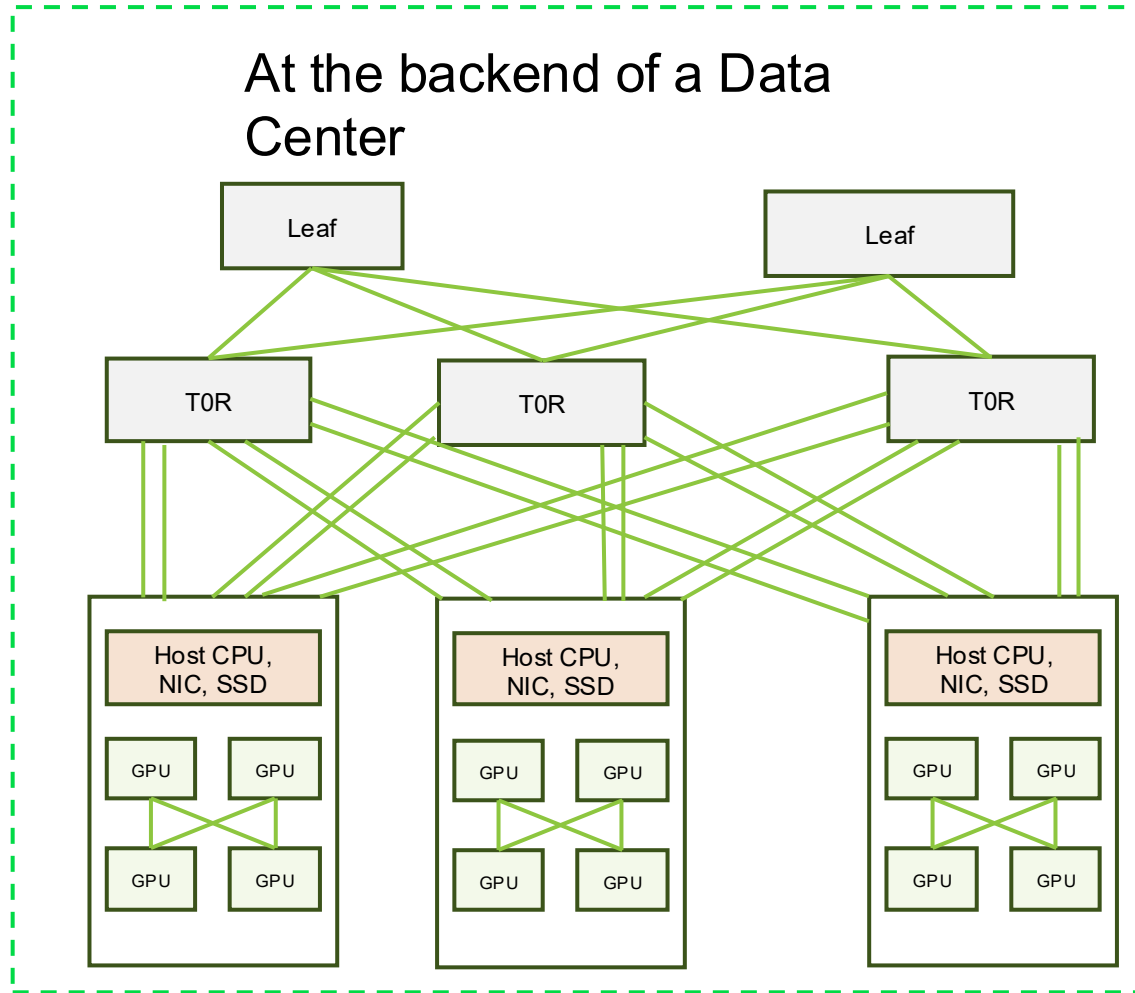
OCP
GLOBAL
SUMMIT

2026

Scaling Innovation for the AI Era.

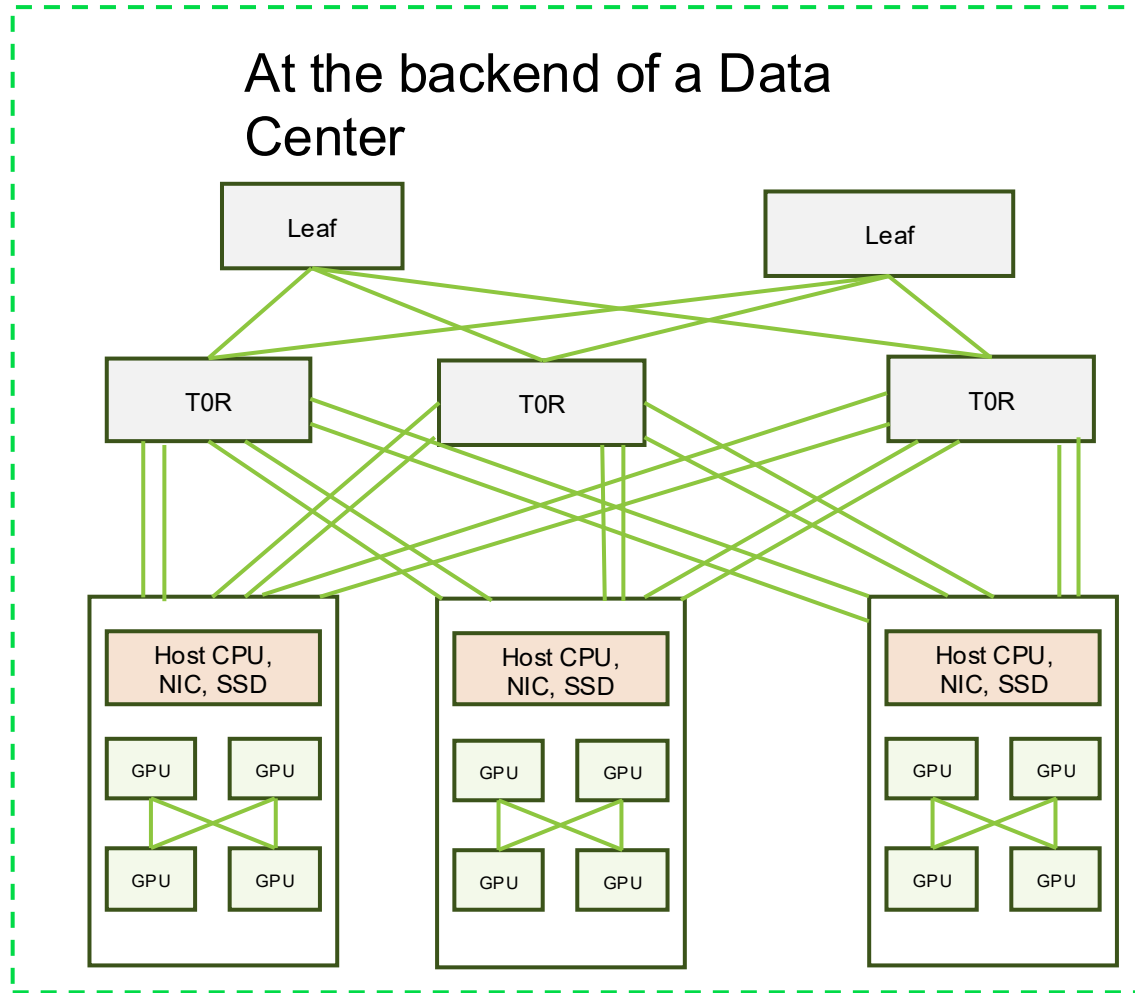
Cisco Confidential

AI - Raising the Bar for Hyperscalers Networks



- New Traffic Pattern:
 - Small number of large flows
 - Periodic bursts of data sent synchronously

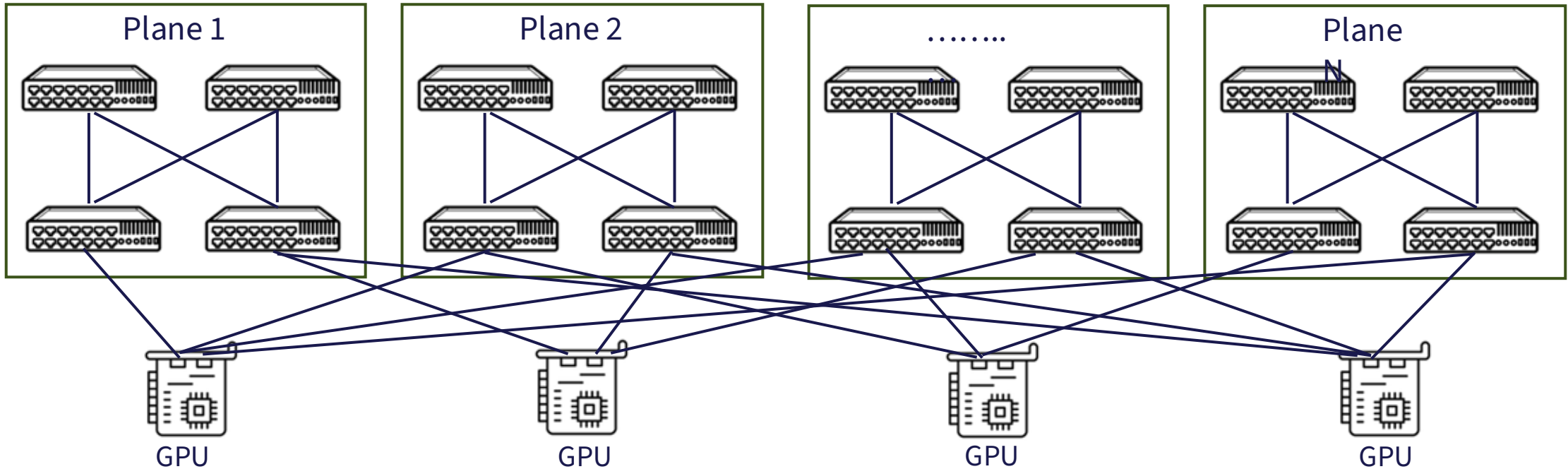
AI - Raising the Bar for Hyperscalers Networks



- Challenges:

- Traditional passive ECMP-based load balancing mechanisms suffered from low entropy problem.
- Failures of communications in LLM training are costly
- An epoch of training is blocked until the synchronized collective communications of last epoch finishes
- GPU hours are expensive resources because of tight supply
- If an ongoing training job crashes, all progress since the last checkpoint may be lost.

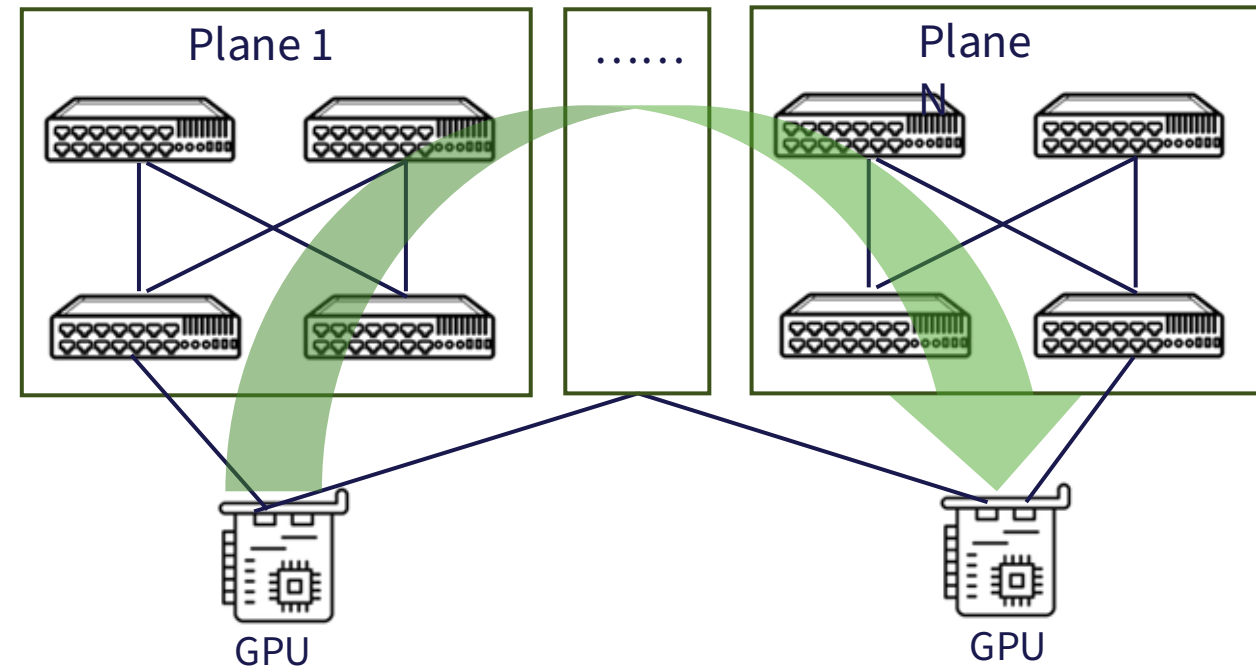
Microsoft Fairwater DC



Multi-Plane design allowing up to 524,288 GPUs!

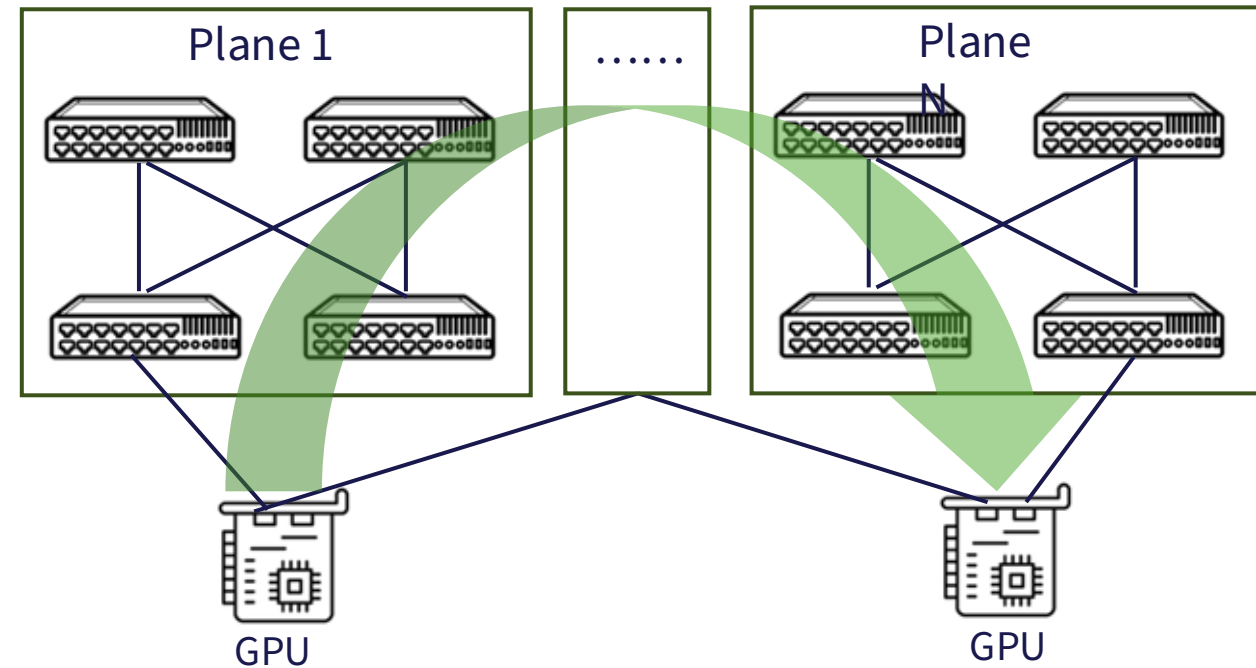
SRv6 in AI Backend DC

- Fined-grained control based on Source Routing
- Path enumeration for traffic management
- Integration with AI workloads flow scheduling provides optimal network performance
- Source quickly reroute upon path failures or congestion
- Reliability: No dependency on Routing Protocol
- Ship-in-the-night with traditional traffic
- Standard, Open, Multi-vendor
- Ease of operation and troubleshooting



SRv6 in AI Backend DC

- The uSID Stack is communicated to NIC
- Source NIC
 - Encapsulates the traffic with the uSID List
- Switch
 - Forward the packet based on the uSID List - Strict Source Routing
 - If drop, Packet Trimming
- Destination NIC
 - Decapsulates the uSID List and delivers the packet
- Both SRv6 uN and uA are supported.



SRv6 in SONiC for AI Backend

SRv6 Configs
IPv6 Static Routes

Config DB

SRV6_MY_LOCATORS Table
serves as a configuration source for SRv6 locators

SRV6_MY_SIDS Table
configures segment identifiers

segment-routing
srv6
static-sids
sid fcbb:bbbb:20::/48 locator loc1 behavior uN

SONiC

BGP docker

FRR

FRR config

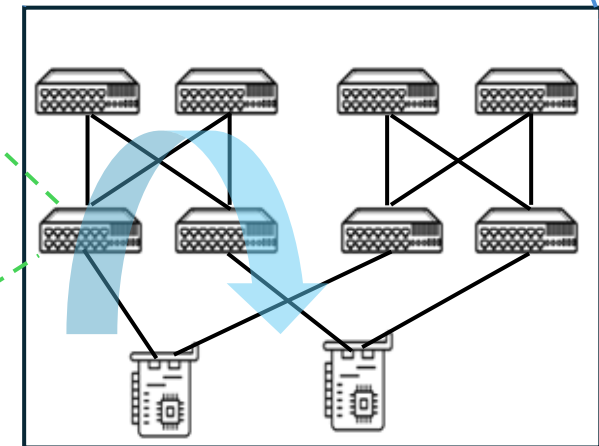
config DB

bgpcfgd

SRV6_MY_SIDS
SRV6_MY_LOCATORS

HW Switch

SDN
Controller



What happened when a T1 stopped forwarding

75K

GPUs in the production job

~25%

of QPs affected

~580K

packets dropped



[Resilient AI Supercomputer Networking using MRC and SRv6](#)



OCP
GLOBAL
SUMMIT

2026

Scaling Innovation for the AI Era.

Cisco Confidential

Scale Up



OCP
GLOBAL
SUMMIT

2026

Scaling Innovation for the AI Era.

Cisco Confidential

Scale Up Requirements

- Ultra-low Overhead
 - Remain highly efficient for 32, 64, 128-byte packet sizes (Flit transactions)
- Simplified packet processing
 - Single Header Architecture
 - Skip unnecessary building-blocks in HW (e.g., LB)
- Native strict source routing support
- Two-tier fabric support
 - Scale Up and combined Scale Up/Scale Out fabrics
- High-radix support (>576)

Call to Action

- For more information about SRv6 <https://www.segment-routing.net/>
- Inviting contributions to SRv6 SONiC
 - SONiC Community meeting: <https://sonic-net.github.io/SONiC/Calendar.html>
 - SONiC Routing Workgroup: sonic-wg-routing@lists.sonicfoundation.dev | [Home](#)



OCP
GLOBAL
SUMMIT

2026

Scaling Innovation for the AI Era.

Cisco Confidential

Thank You!



OCP
GLOBAL
SUMMIT

2026