



Beyond Traditional BGP: Decoupling Route Distribution for Large- Scale AI Clusters at Alibaba

Eddie Ruan

Dennis Cai

Alibaba Cloud

Recap what Microsoft presented last year

BGP



Two layer CLOS Network Capacity = $C(\text{switch})^2 / 2^B(\text{links})$

	51.2T	102.4T
800Gb/s	1,638.4T	13,107.2T
100Gb/s	13,107.2T	104,857.6T

Skinny links are desirable to get more bandwidth in the fabric !

BUT

As ports increase, BGP sessions 64 => 512 !



SONIC OPTIMIZATIONS TO SCALE BGP

- FRR upgrade 10.0.1 + ~20 patches
- Scale up FRR to support 514 sessions and nexthop
- 512 BGP session w/ 1K route 512 nexthop:
- Single port flapping <0.1s data plane down time and ~5s route convergence time.
- 100-150s port/bgp startup time (from +Inf)

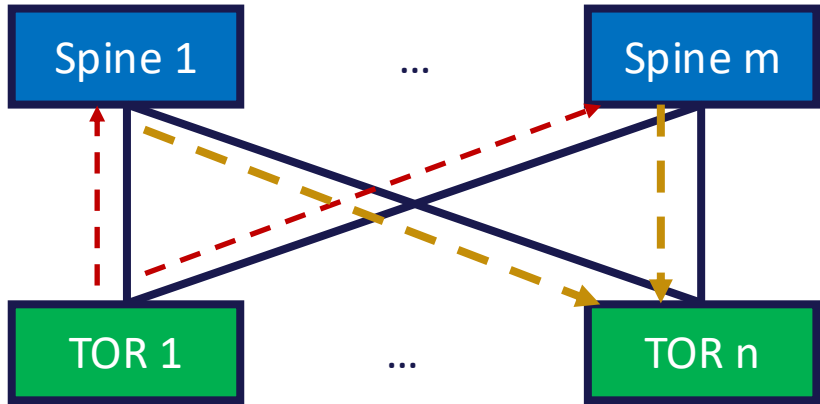


OCP
GLOBAL
SUMMIT

2026

Scaling Innovation for the AI Era.

Issues with traditional BGP approach



Assume we have

- $m = 200$ spine switches
- $n = 500$ TOR switches
- Each TOR supports 200 service routes

The number of infra routes : $m + n = 700$

The number of services routes : $200 * 500 = 100K$

Each TOR needs to learn

$100K * 200 = 20 \text{ M}$ service routes' information from all spine switches.

Microsoft case:

$1000 * 512 = 512K$

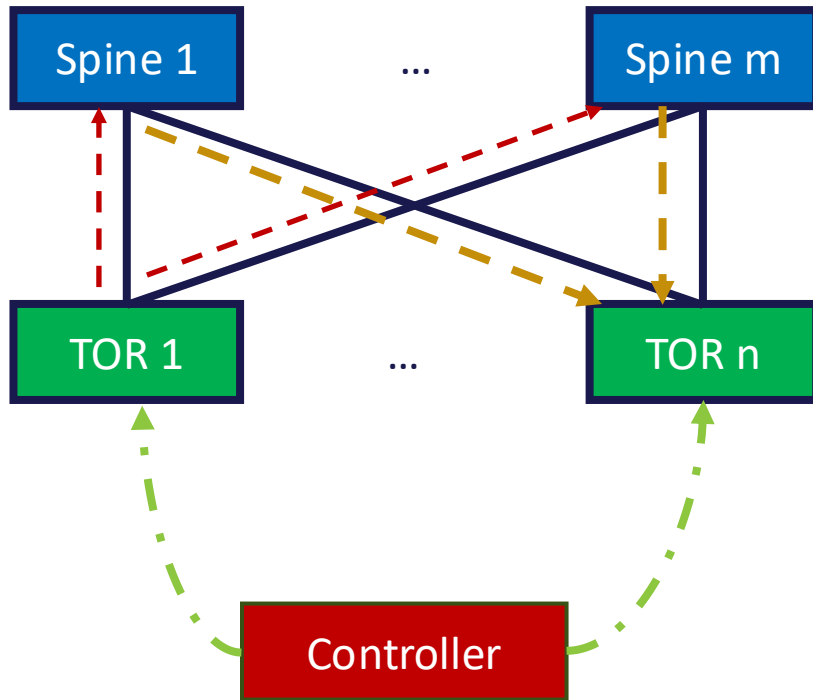


OCP
GLOBAL
SUMMIT

2026

Scaling Innovation for the AI Era.

Google's suggestion



Assume we have

- $m = 200$ spine switches
- $n = 500$ TOR switches
- Each TOR supports 200 service routes

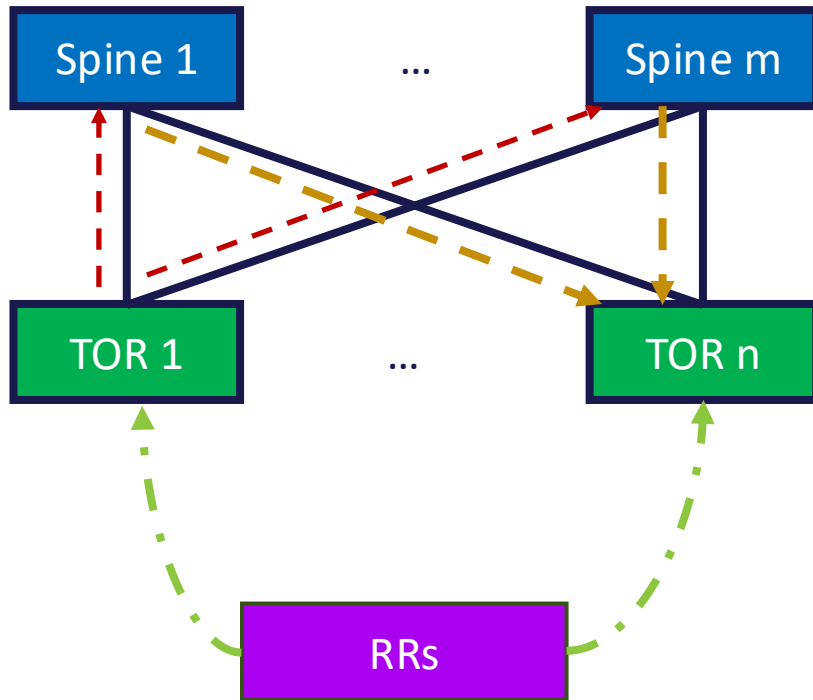
The number of infra routes : $m + n = 700$

The number of services routes : $200 * 500 = 100K$

Each TOR learns

- Ony infra routes learnt via BGP from all spine switches
 $(m+n) * m = 700 * 200 = 140K$
- Service routes learnt from controller
each service route via TOR_i

Cisco's suggestion



Assume we have

- $m = 200$ spine switches
- $n = 500$ TOR switches
- Each TOR supports 200 service routes

The number of infra routes : $m + n = 700$

The number of services routes : $200 * 500 = 100K$

Each TOR learns

- Ony infra routes learnt via BGP from all spine switches
 $(m+n) * m = 700 * 200 = 140K$
- Service routes learnt via RRs
each service route via TOR_i



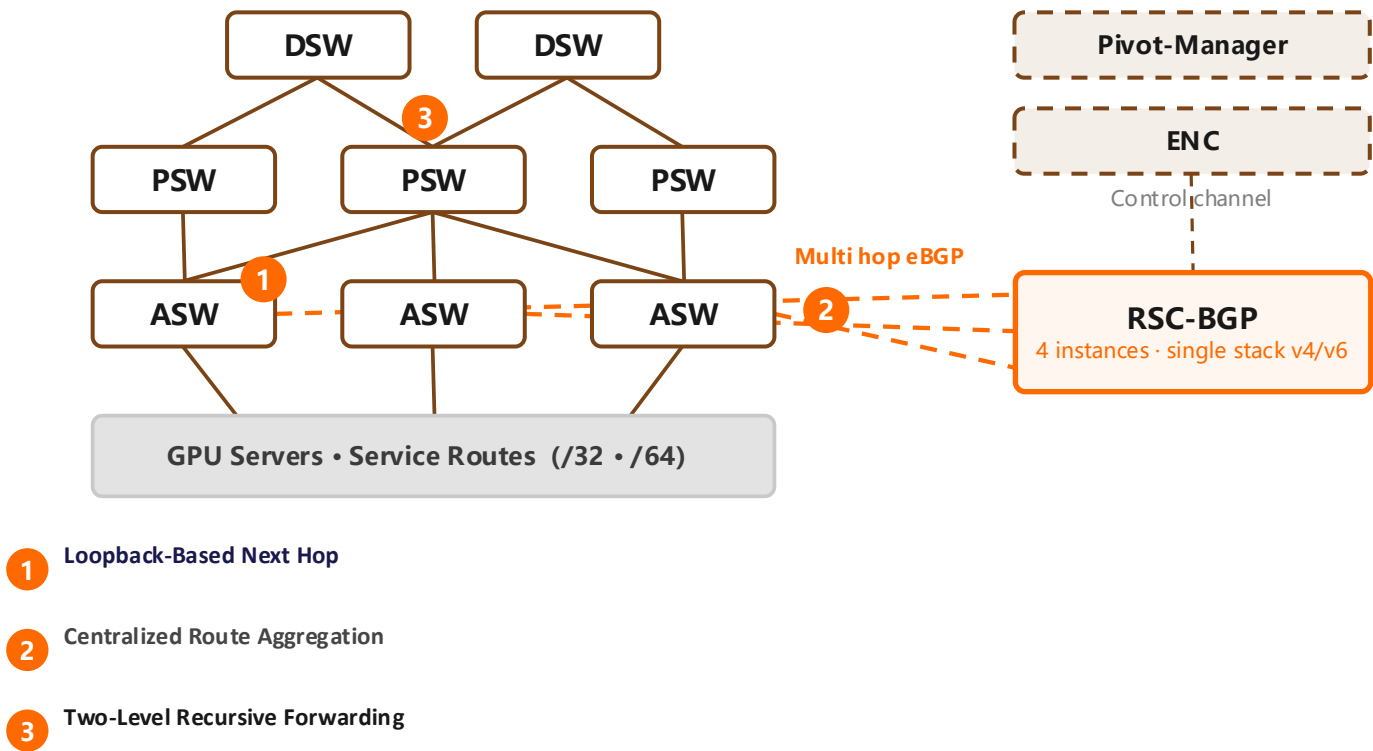
OCN
GLOBAL
SUMMIT

2026

Scaling Innovation for the AI Era.

HPN Pivot: Large-Scale Routing Architecture for Hyperscale AI Clusters

Pivot Routes offload Mechanism • Two Tier network diagram



Challenge

- Exponential Scaling Pressure
- Routing State Explosion
- Degraded Network Resilience

Solution

~~≈over 200~~ → **4**

Massive Peer Reduction
Control Plane Relief
Accelerated Convergence

From Architecture To Deployment

- **Form Factor:** Physical vs. Virtual Appliances
 - **Decision: Virtual RR.** Selected for greater deployment agility, resource efficiency, and seamless integration.
- **Placement:** Centralized Cloud vs. Cluster-Proximate (On-Premises)
 - **Decision: Cluster-Proximate.** Deployed locally to ensure low latency and highly reliable BFD sessions.
- **Operations:** Manual vs. Automated Lifecycle Management
 - **Decision:**
 - **Pivot Manager.** Introduced to automate the provisioning, scaling, and lifecycle management of the RRs
 - No K8



OCF
GLOBAL
SUMMIT

2026

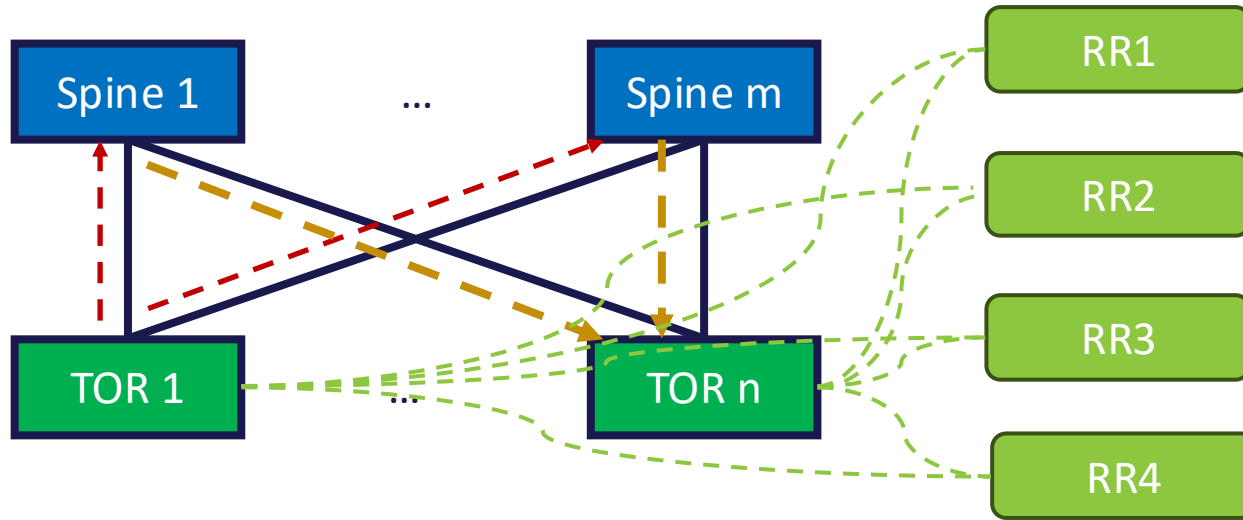
Scaling Innovation for the AI Era.

Reliability is the key

- Bottom line : any RR outage should not impact AI cluster.
- Solution :
 - Use 4 RR instances per plane per protocol
 - Locate 4 RR instances close to AI cluster
 - 4 RR instances are on different physical servers
 - Physical servers are located at different racks, via two different network clusters.
 - Each RR instances runs independently with all TOR switches in the plane
 - The same routes are distributed by 4 RRs.



BGP Group GR



BGP Group GR:

- Only the last peer (in the group) going down triggers GR
- The first upcoming (reconnect or NEW) peer (in the group) resets the active GR

Problem Statements:

What happen if all RRs are gone simultaneously?

Solution:

Introduce BGP Group GR for headless forwarding.

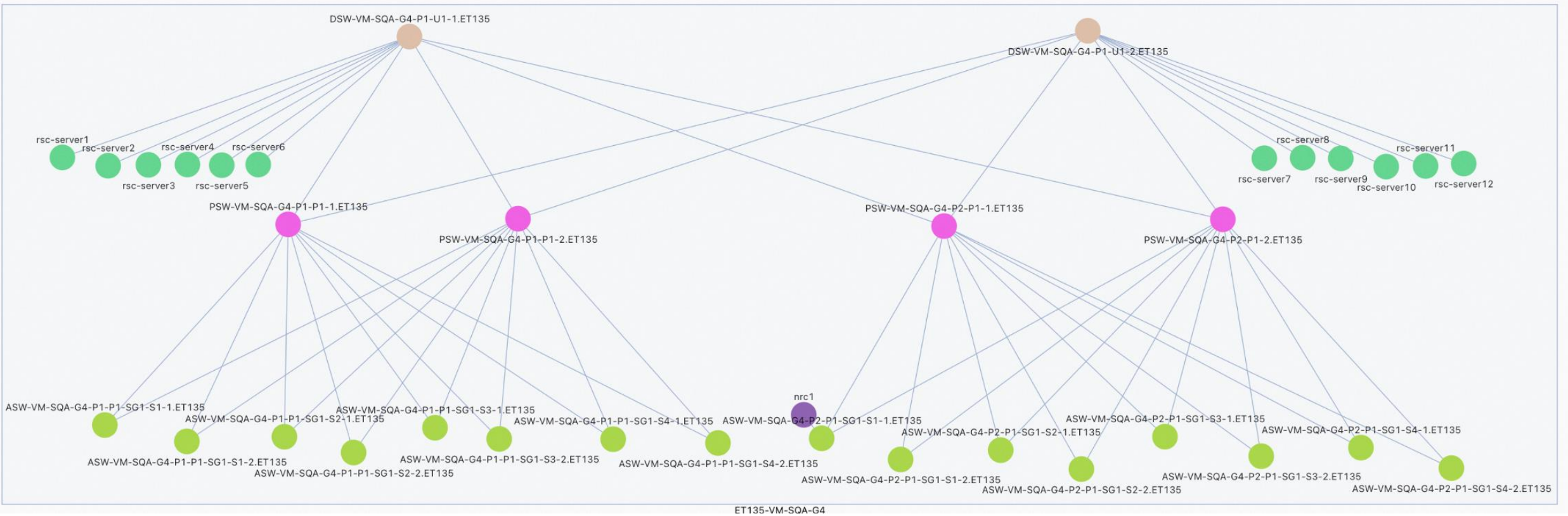


OCPI
GLOBAL
SUMMIT

2026

Scaling Innovation for the AI Era.

Full system Emulation



OCPI
GLOBAL
SUMMIT

2026

Scaling Innovation for the AI Era.

Call to Action

- Problem to solve
 - Continue to enhance BGP's scale / performance along with AI cluster's growth.
 - Enhancement on BGP group GR
 - Heterogeneous RR Software Stack Further Enhances RR Cluster Reliability
- How to get involved
 - <https://lists.sonicfoundation.dev/g/sonic-wg-routing>



OCP
GLOBAL
SUMMIT

2026

Scaling Innovation for the AI Era.

Thank You!



OCP
GLOBAL
SUMMIT

2026