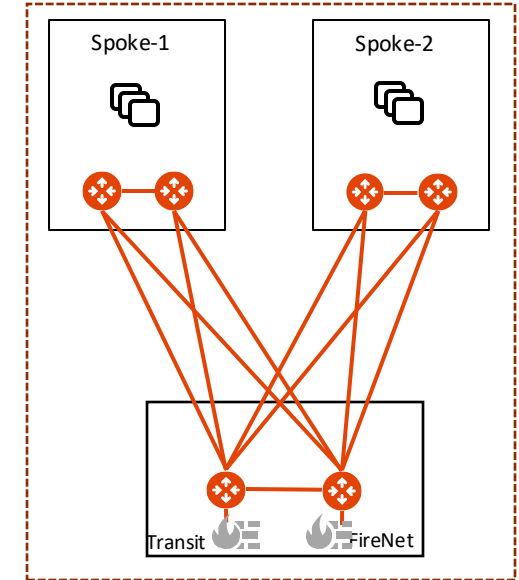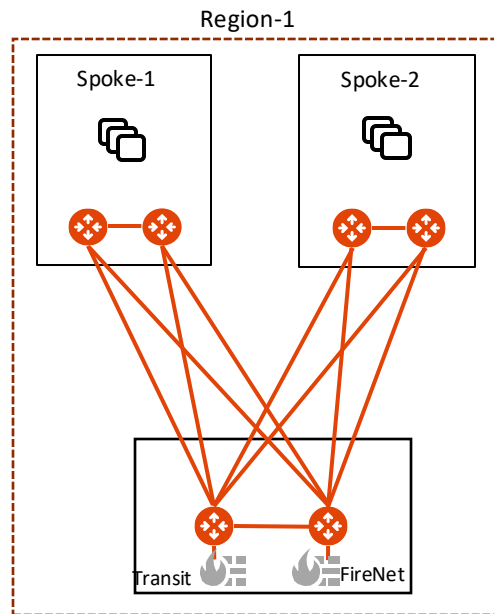# ActiveMesh

ACE Team

Overview

What is it ?

- Provides network **resiliency**, improved convergence time and high performance

- Two Aviatrix gateways in a VPC/VNet/VCN form a cluster

- Both gateways forward traffic simultaneously via ECMP

- Each gateway in a Spoke VPC/VNet/VCN builds IPsec tunnels to **both** Transit gateways

- Number of Transit and Spoke gateways as well as their **instance sizes** are independent of each other:

  - ➤ Maximum **2x** Transit Gateways can be deployed per Transit VPC/VNet/VCN

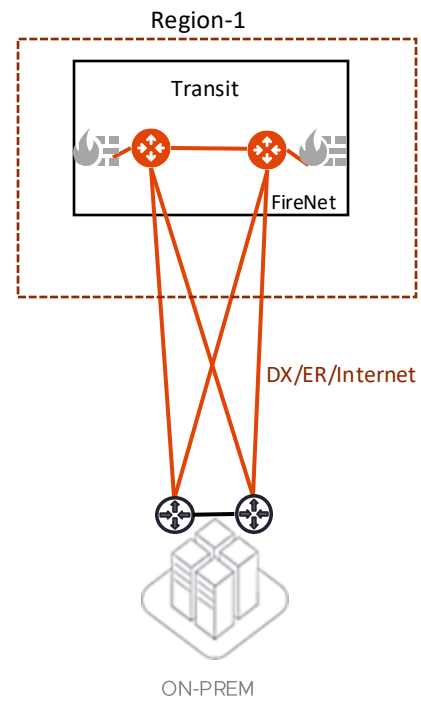 - ➤ Maximum **15x** Spoke Gateways can be deployed per Spoke VPC/VNet/VCN
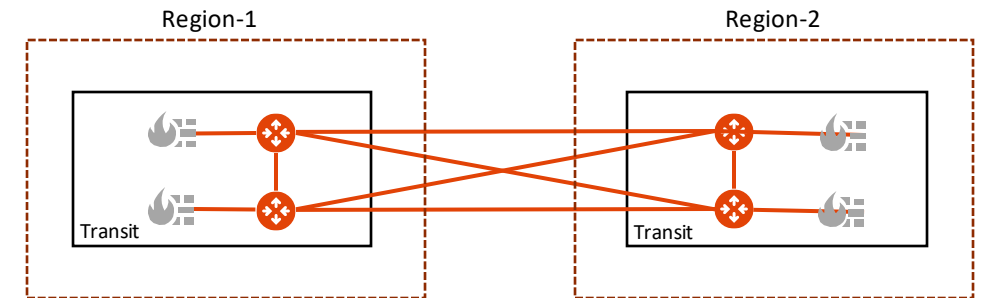
Use Cases

Intra-Region Spoke-Spoke

Region-1

Spoke-1   Spoke-2

Transit   FireNet

Cloud to On-Prem

Region-1

Transit

FireNet

DX/ER/Internet

ON-PREM

Inter-Region / Multi-Cloud

Region-1   Region-2

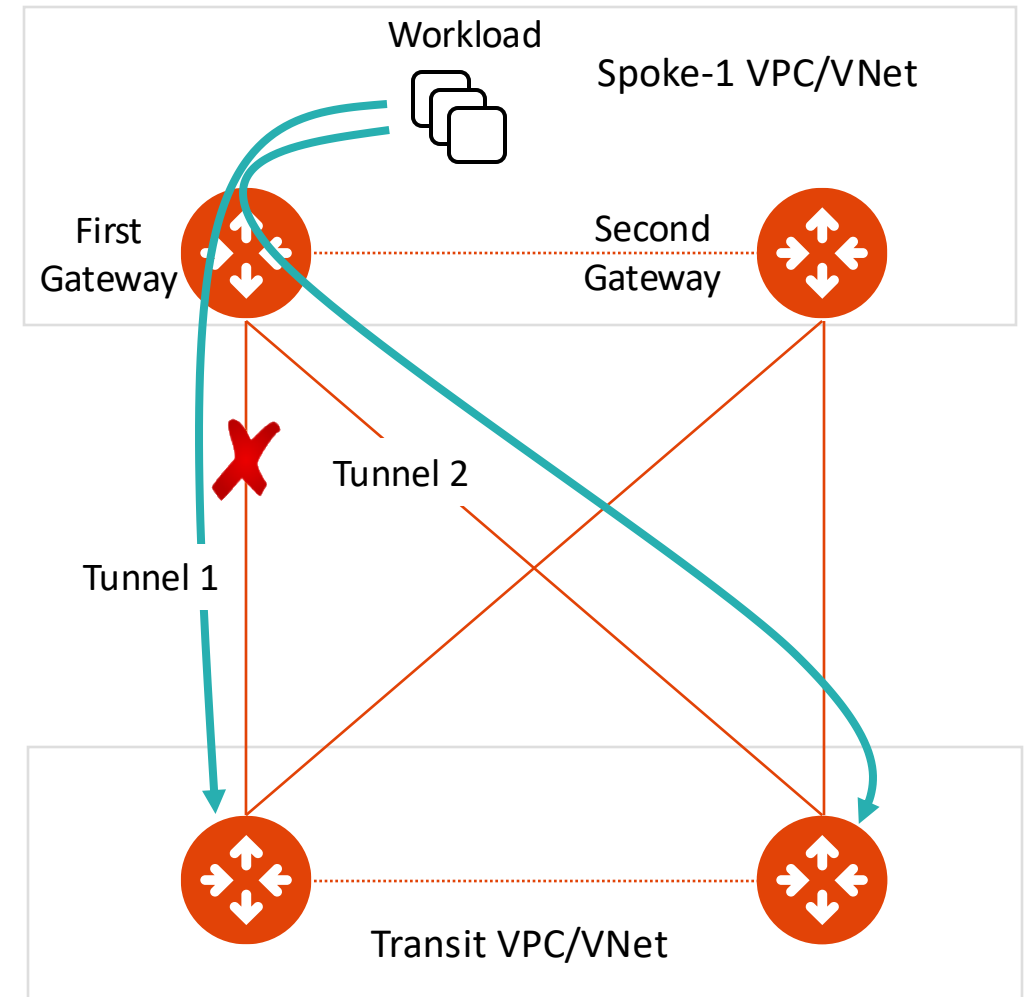Transit   Transit

AVIATRIX

# Resiliency

Failover Scenario 1

- Workload in Spoke-1 VPC/VNet traverses Primary gateway, Tunnel 1, onto Transit to Spoke-2 VPC/VNet (not shown)

- If Tunnel 1 at the First Spoke Gateway fails,

  - Then, the traffic uses Tunnel 2, connected to the SecondTransit Gateway

  - This tunnel was already active and was forwarding half of the traffic (same metric 100)

- **No re-convergence** of the routes in the VPC/VNet route table

- Gateway handles the change on its own

- Controller is aware of the tunnel going down event, but **it is not involved** in making the change

Failover Scenario 2

- Workload in Spoke-1 VPC/VNet traverses Primary gateway, Tunnel 2, onto Transit to Spoke-2 VPC/VNet (not shown)

- If both Spoke ←→ Transit tunnels fail on Primary Spoke gateway:
  - The traffic gets forwarded from the Primary Spoke gateway through the interconnected link to the Secondary Spoke Gateway
  - Secondary Spoke Gateway forwards the traffic to any of the Transit Gateways via ECMP (usual behavior – metric 100 on both downstream links)

- No re-convergence of the routes in the VPC/VNet route table

- Gateway handles the change on its own

- Controller is aware of the tunnel going down event, but it is not involved in making the change

Workload

Spoke-1 VPC/VNet

First Gateway

Second Gateway

Tunnel 2

Tunnel 1

Transit VPC/VNet
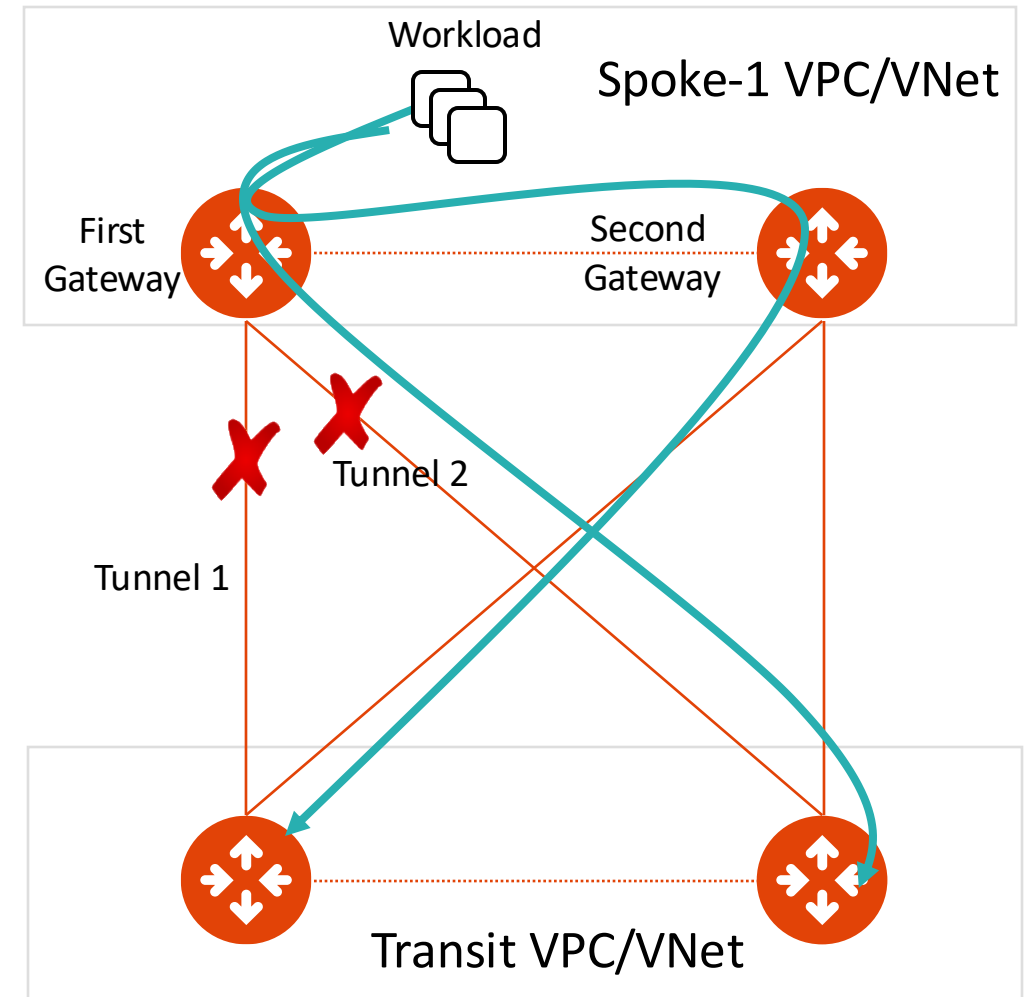
Failover Scenario 3

- The workload in Spoke-1 VPC/VNet needs to reach Spoke-2 VPC/VNet (not shown), but the Gateway is down.

- If the First Gateways fails, the Controller will detect this event through the periodic keepalive messages exchanged between itself and the gateways.

- In this scenario, the Controller will **reprogram the routing table** in the AZ-A, updating the next hop of the three RFC1918 routes with the <u>ENI of the Second Spoke Gateway</u>, in AZ-B

© Aviatrix Certified Engineer
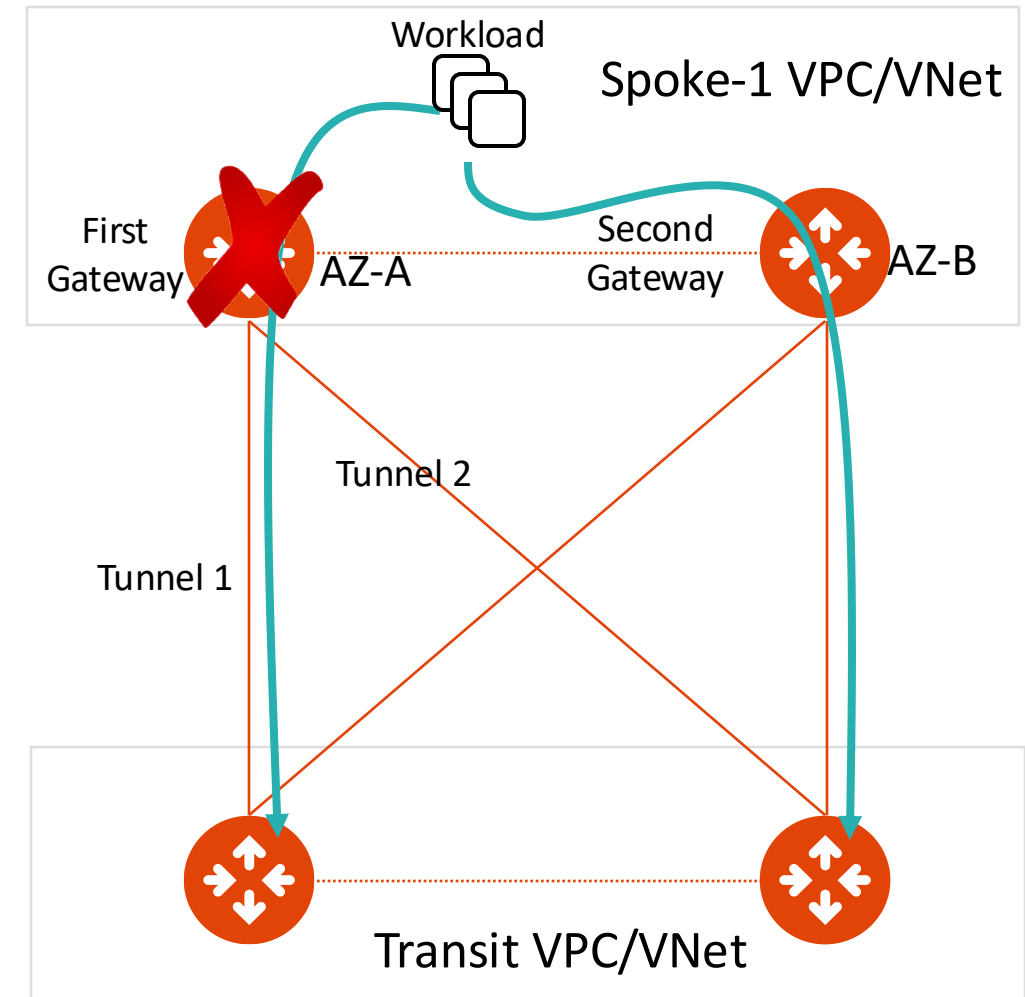
Active-Standby Mode (introduced in Controller version 6.6)

- **Use case:** Deployment scenario where an on-prem device such as a firewall does not support asymmetric routing on two tunnels
- Upon failure, the Second gateway takes over from the First gateway
- The first does not become active unless there is a manual switchover or the Second fails
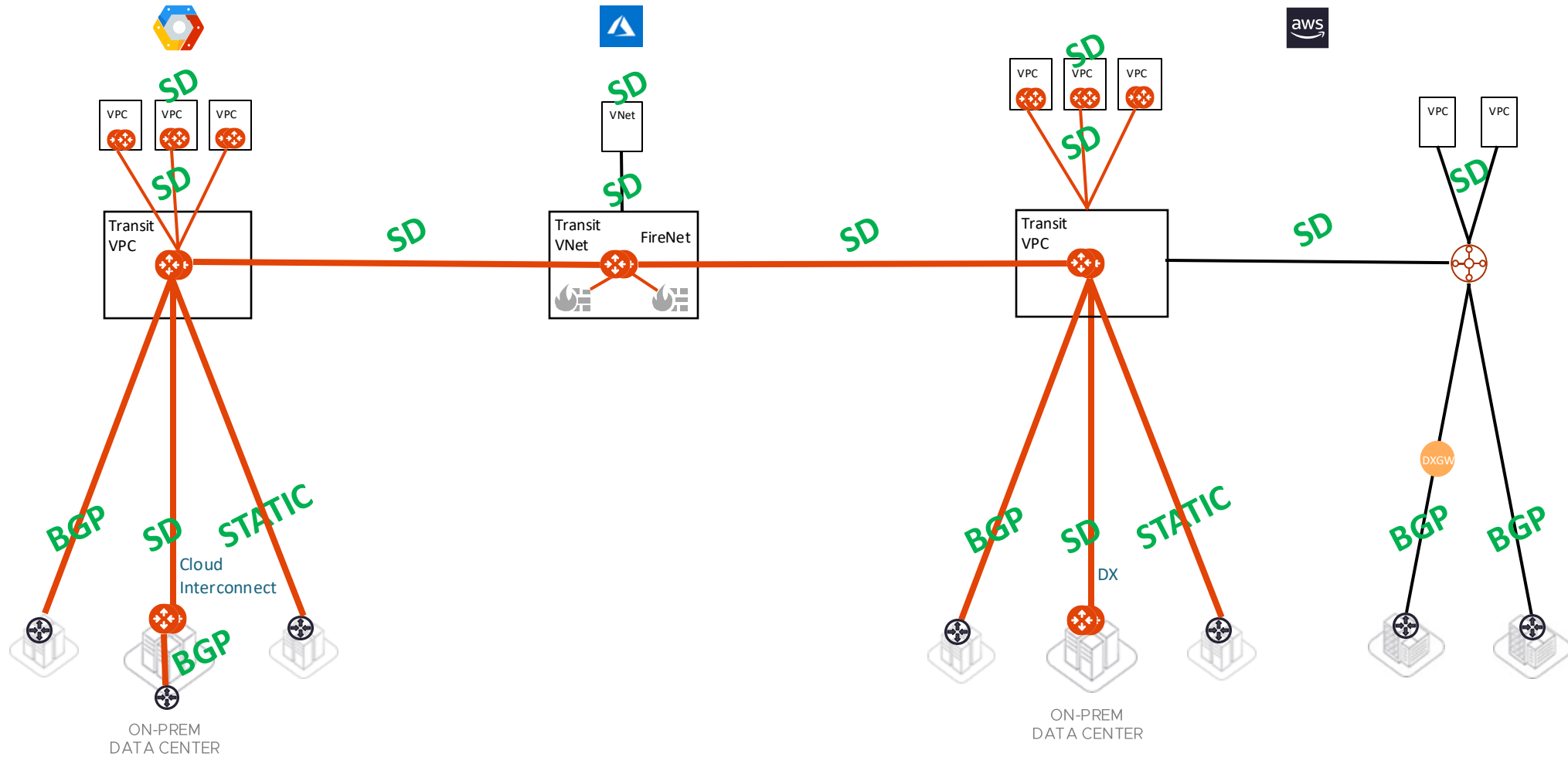- UI provides an option for customers to choose **Preemptive** or **Non-Preemptive** behavior.
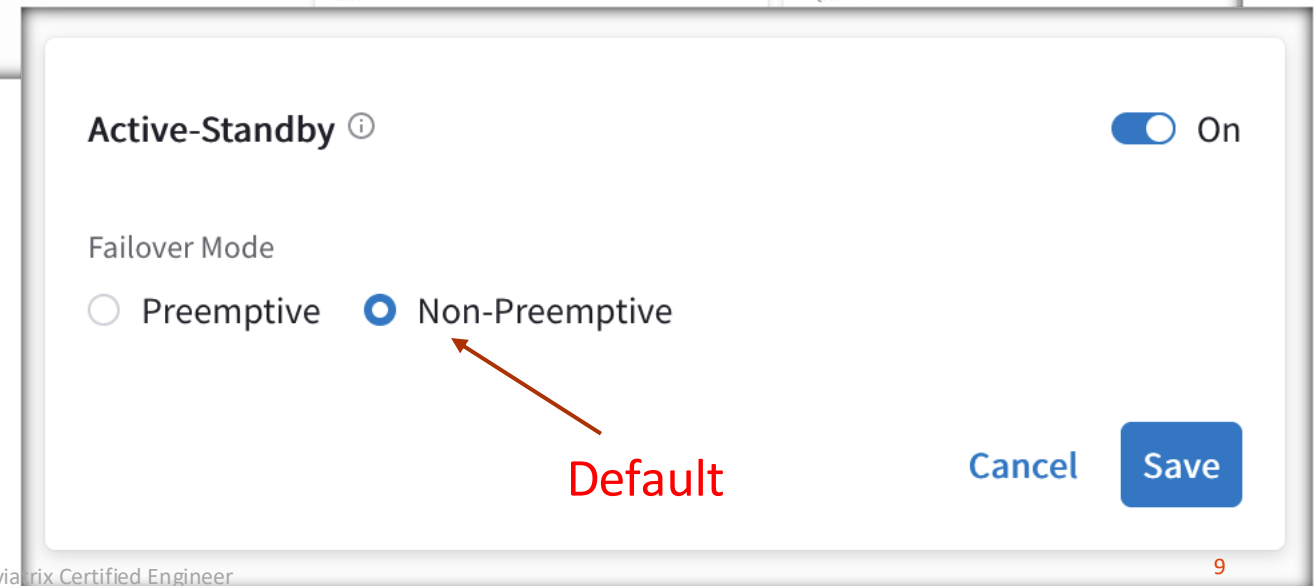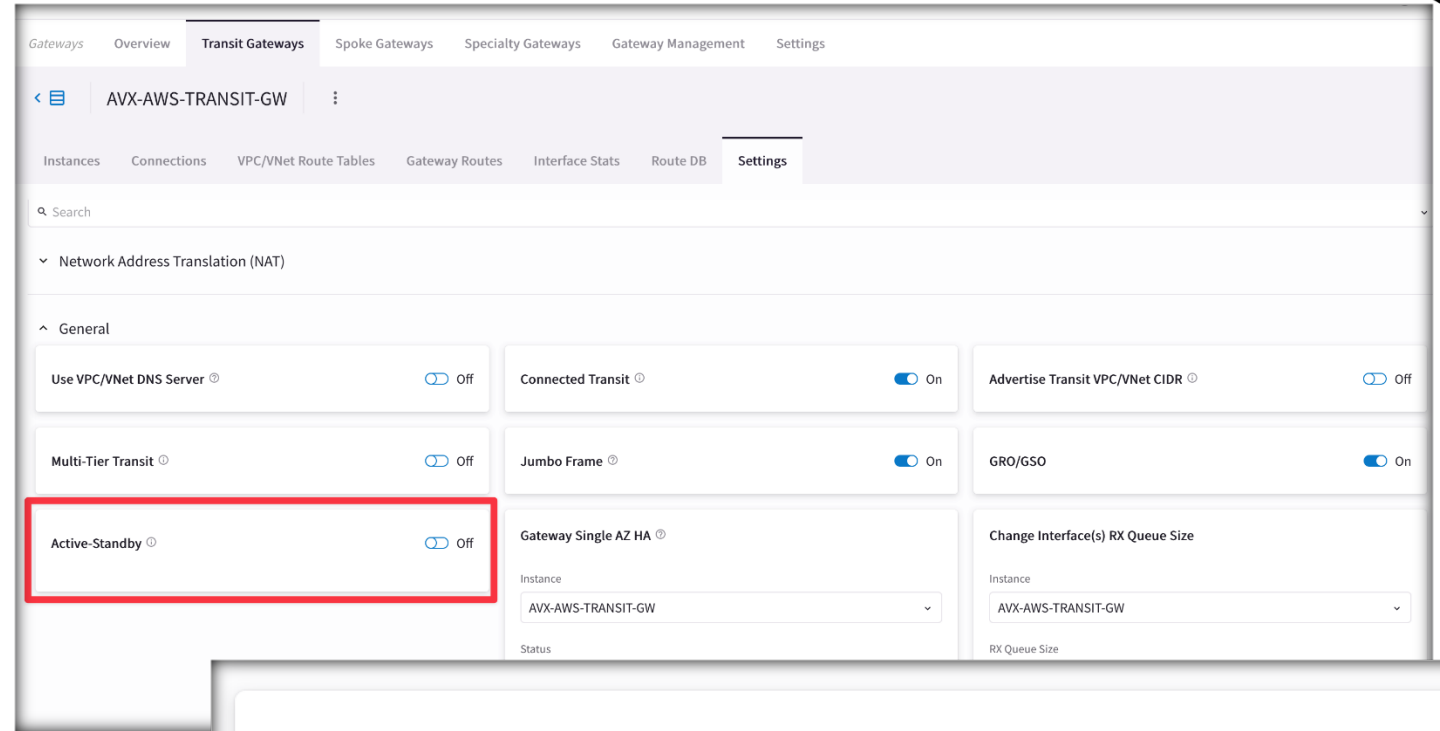
Aviatrix Control Plane
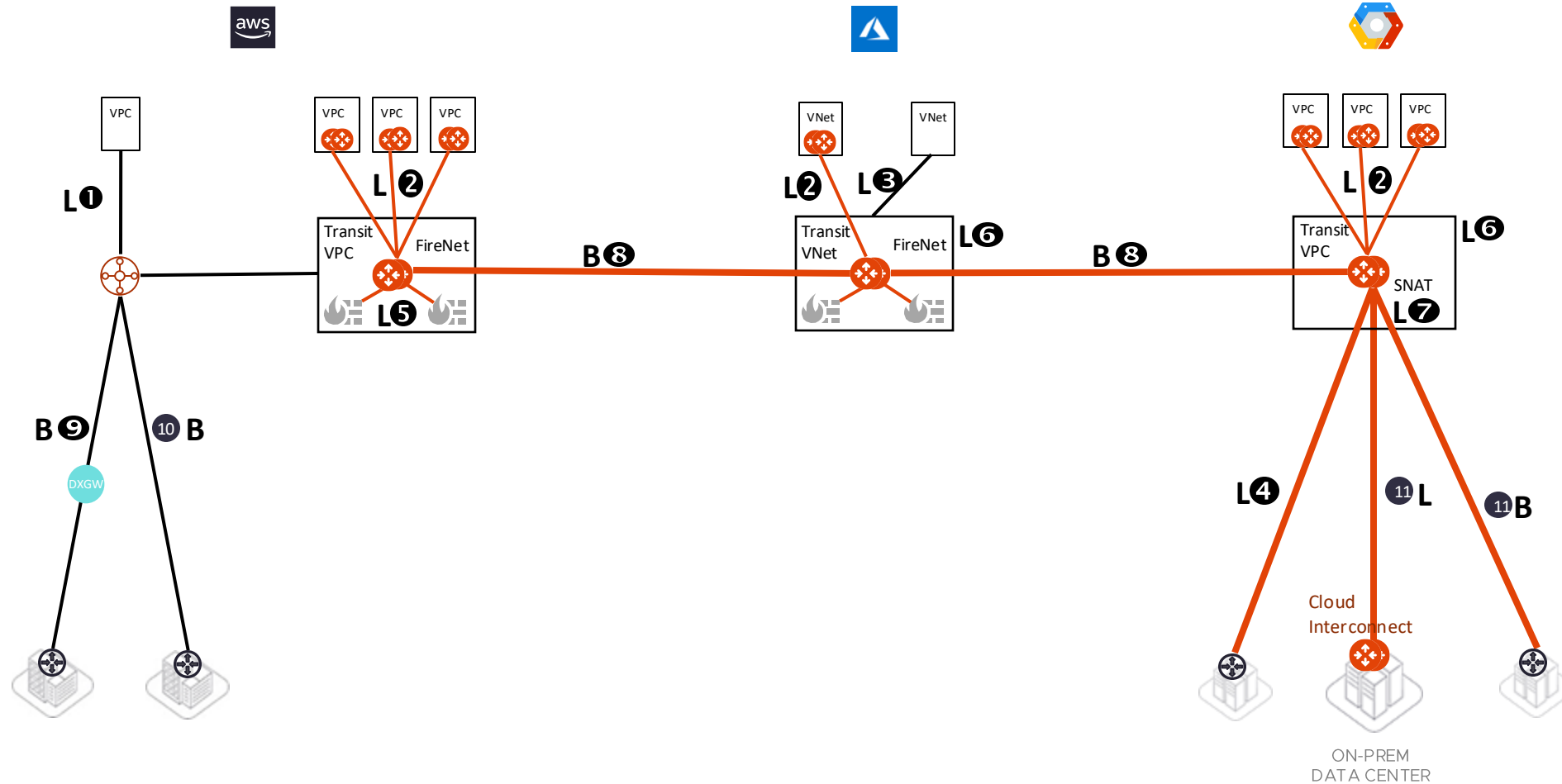
Deterministic Routing

Route Programming: Software-Defined (SD) / Static / Dynamic

Route Classification

1. AWS TGW Attachment [L]

2. Aviatrix Spokes (VPC/VNet) [L]

3. Azure Native Spokes [L]

4. Aviatrix Transit GW – on-prem (*static*) [L]

5. Firewall Egress 0/0 [L]

6. Transit VPC/VNet associated prefixes [L]

7. Transit GW SNAT IP [L]

8. Remote Transit GW (Transit Peering) [B]

9. TGW DXGW [B]

10. TGW VPN [B]

11. Site2Cloud BGP on Transit GW (including Edge routes) and Site2Cloud BGP on Spoke GW [B]
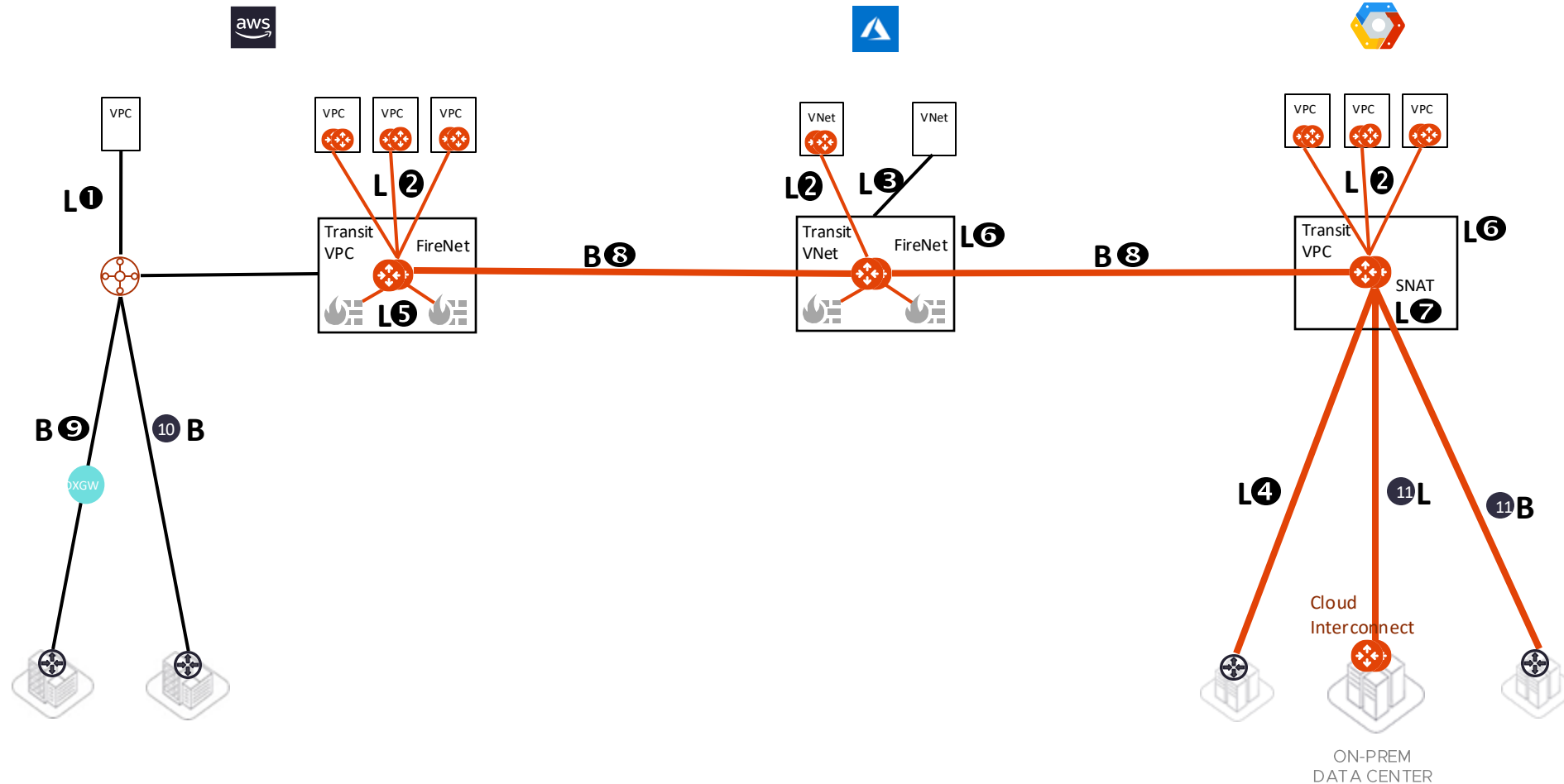

L = Routes considered local by Controller
B= BGP learned routes

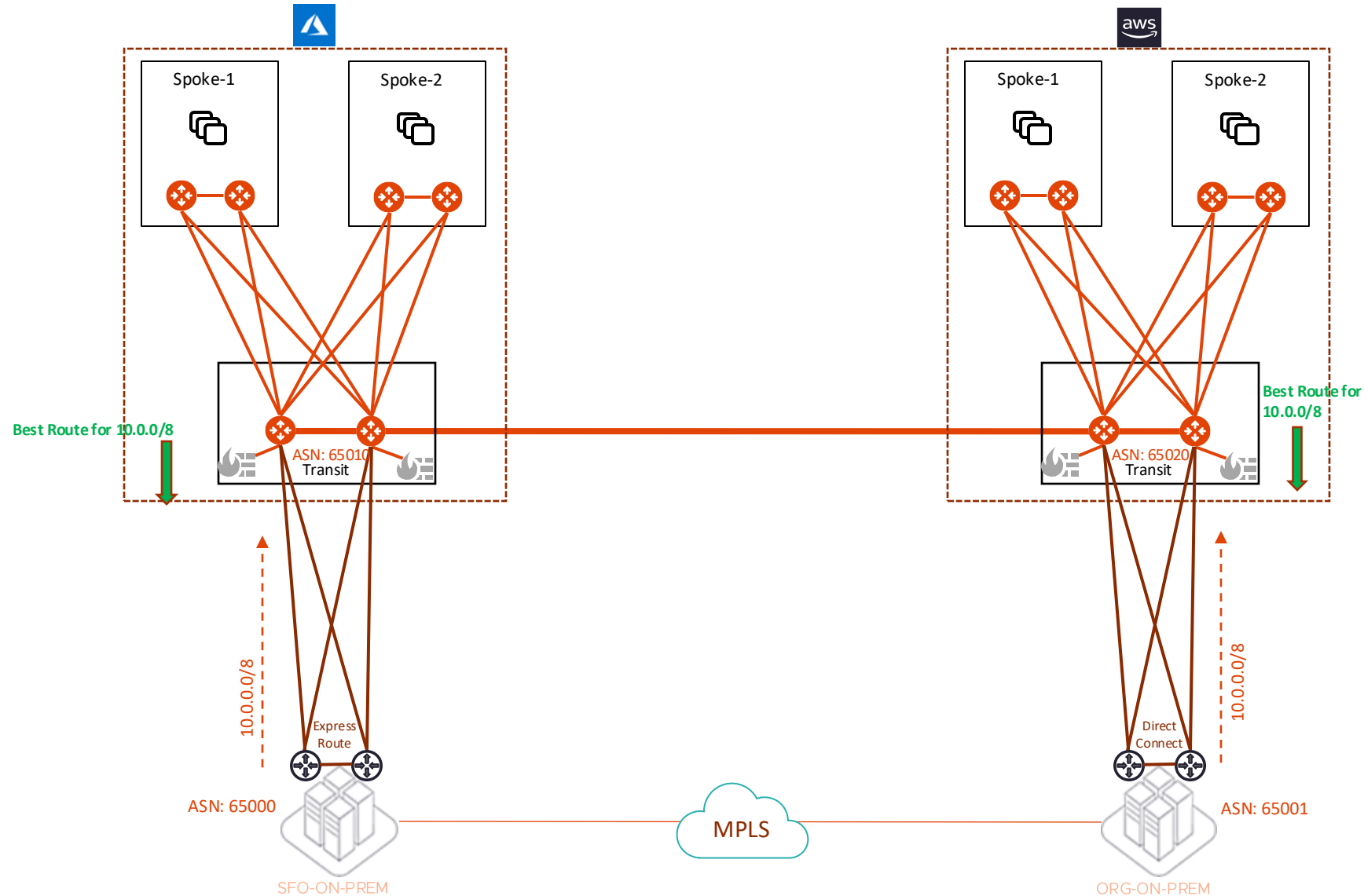# Path Selection Algorithm for Deterministic Next-Hop Selection

Route Selection Algorithm

1. Longest prefix match

2. If equal length, then local route is chosen

3. If routes are of the same type, then shortest AS-path length is chosen

4. If AS-path length is the same, then lowest metric is chosen

5. If metric is the same, then

   • If ECMP is enabled, then traffic is distributed to available routes

   • If ECMP disabled, then the route first programmed in the table is chosen

      • If programmed at the same time, then lower integer IP next hop is chosen

L = Routes considered local by Controller
B= BGP learned routes


© Aviatrix Certified Engineer

13

# Example of using Transit as an alternate path 1/3

- At steady state

  - Each transit is learning 10/8 locally from on-prem

  - For each transit, Controller DB will have 10/8 via local and peer transit

  - Route via peer will have as-path-len 2

  - Each transit and its spokes will get to on-prem via local private path (DX/ER) as best path

© Aviatrix Certified Engineer

Example of using Transit as an alternate path 2/3
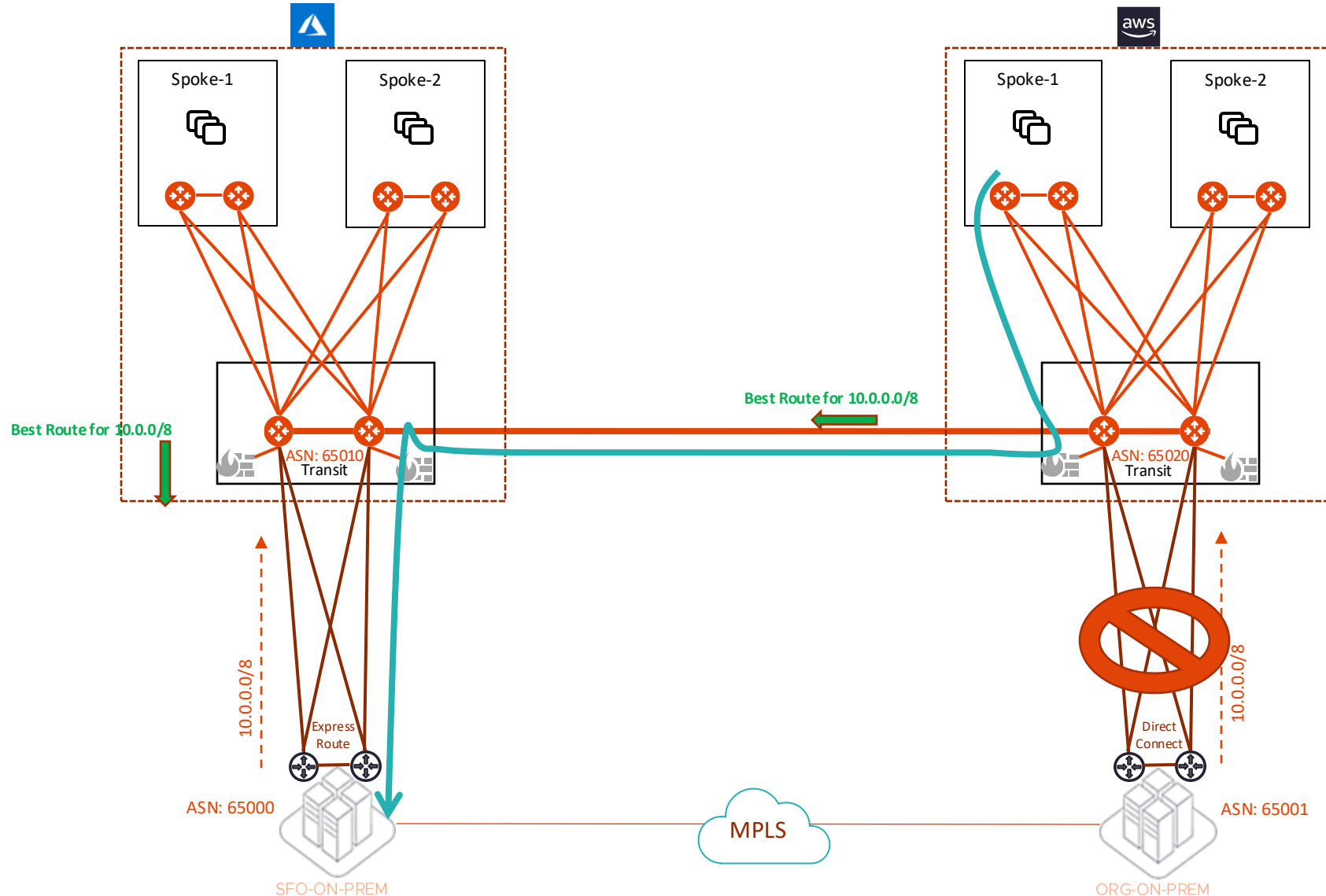
- When on-prem connection goes down

  - For e.g., DX is down

  - Only route to 10/8 now is via Azure Transit

Example of using Transit as an alternate path 3/3
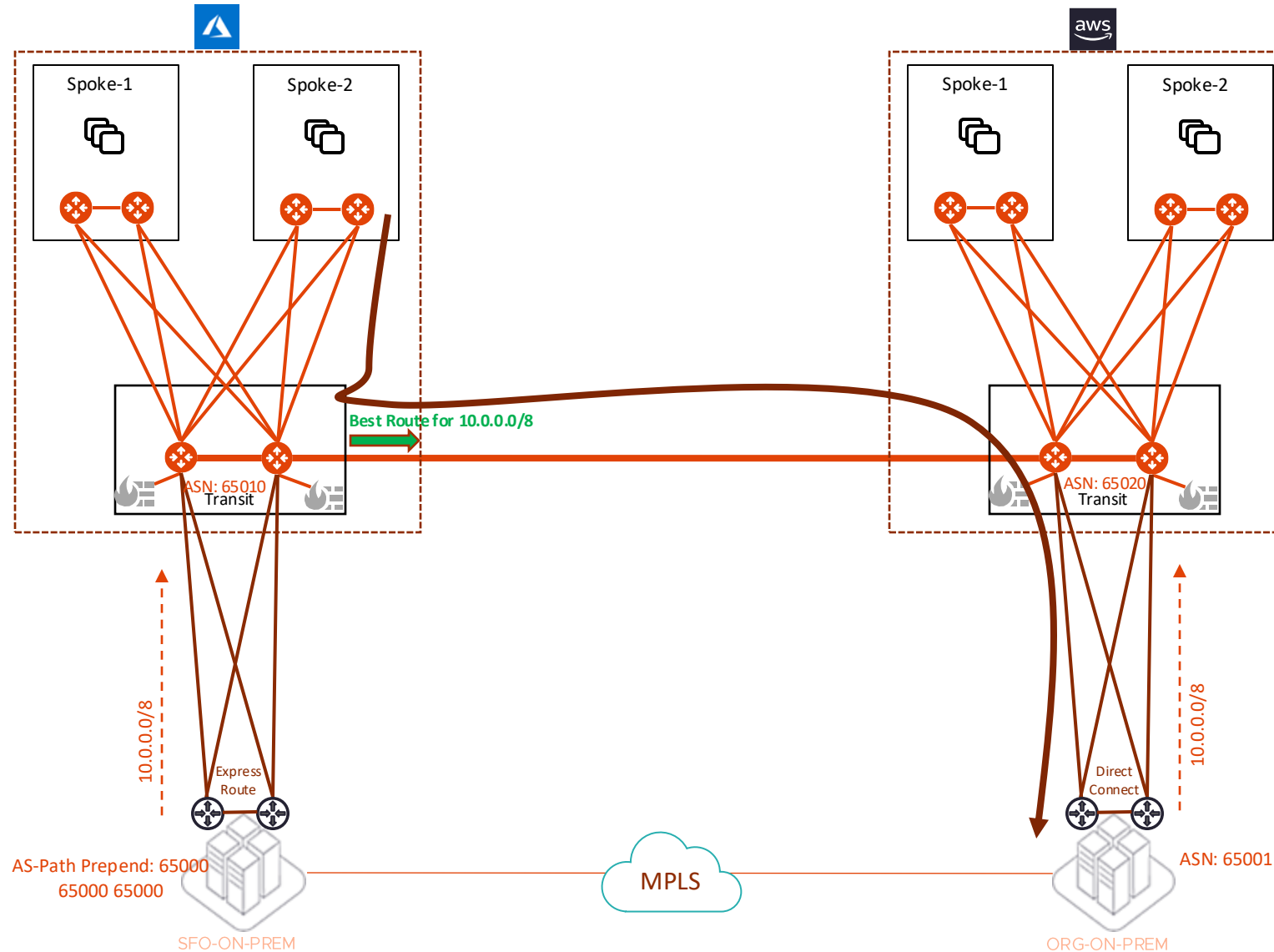
Use AS-PATH Prepend

- E.g, SFO on-prem ER is going under planned maintenance

- You want to avoid sending any traffic through SFO on-prem ER

- You can send AS-paths from SFO on-prem so that AWS Transit becomes the preferred path

Next: Lab 4 – HPE with ActiveMesh