

Global L2VPN Overlay over Layer 3 Transport

Design, Addressing and Operations Reference

Document control

Field	Value
Document ID	NET-DES-0142
Title	Global L2VPN Overlay over Layer 3 Transport
Revision	2.1
Date	2026-09-24
Status	Approved (sanitized sample)
Classification	Internal; sanitized for external distribution
Owner	Network Engineering [REDACTED]
Sites in scope	SFO, IAD, FRA, AMS, SIN

Sanitization notice. Hostnames, addresses, provider names, certificate identifiers, software product names and personnel have been replaced with representative values. All IP addressing is drawn from 10.0.0.0/8 (RFC 1918) and does not describe any production network. Configuration excerpts are illustrative and must be validated before use.

Contents

1 Purpose and Scope	2
2 Design Summary	2
3 Topology	3
4 Encapsulation and Transport	4
5 IP Addressing Plan	5
6 Tunnel Endpoint Build	7
7 Loop Prevention and BUM Control	8
8 Redundancy and Failover	9
9 Validation Results	10
10 Operations Runbook	10
11 Security Controls	11
12 Known Limitations and Roadmap	12
Appendix A Revision History	12
Appendix B Glossary	12

Conventions

Notation	Meaning
[REDACTED]	Value removed during sanitization
<site>-l2tep-NN	Tunnel endpoint that only operates as a spoke
<site>-l2hub-NN	Tunnel endpoint that can operate as a hub
L-0x / L-1x	Active overlay links / standby overlay links
Site index	SFO = 10, IAD = 20, FRA = 30, AMS = 40, SIN = 50 (third octet of per-site blocks)
RTT	Median round-trip time, milliseconds

1 Purpose and Scope

This document describes the design, addressing, configuration baseline and operating procedures for the global Layer 2 VPN (L2VPN) overlay that extends a single Ethernet broadcast domain across five data center sites: San Francisco (SFO), Ashburn (IAD), Frankfurt (FRA), Amsterdam (AMS) and Singapore (SIN). The overlay carries 802.1Q-tagged Ethernet frames inside encrypted TLS sessions over the existing routed Layer 3 (L3) transport, so no provider Layer 2 service (VPLS, EVPN or dark fiber) is required.

In scope: tunnel endpoints (TEPs), the hub-and-spoke link topology, stretched VLANs 100, 200 and 999, underlay reachability requirements, failover and day-2 operations. **Out of scope:** underlay WAN design, routed inter-VLAN policy beyond the default gateway, and application-level clustering.

Business drivers:

- Workload mobility between sites without re-addressing (live migration and DR cut-over).
- Database cluster heartbeat and replication that requires Layer 2 adjacency (VLAN 200).
- Transport independence: the overlay runs over any IP path that permits TCP/443 between endpoints.

2 Design Summary

Attribute	Value
Topology	Hub-and-spoke. Primary hub IAD, standby hub FRA
Sites	5: SFO, IAD, FRA, AMS, SIN
Overlay service	Single switch instance GLOBAL-L2 carrying an 802.1Q trunk (VLANs 100, 200, 999)
Transport	TLS 1.3 over TCP/443 across the routed L3 underlay
Encryption / authentication	AES-256-GCM; mutual X.509 certificate authentication
Sessions per link	4 parallel TCP sessions, frames hashed per flow
Inner MTU	1500 bytes (1522-byte tagged frames)
Loop prevention	Strict hub-and-spoke, BPDU filtering at the overlay edge, storm control, MAC limits
Failover	Manual, runbook-driven hub move to FRA. RTO target 15 min (drill: 3 min 40 s)
Capacity	Hub sized for 4 Gb/s TLS throughput; observed peak 1.6 Gb/s

Each site runs one tunnel endpoint (TEP): a hardened Linux virtual machine with an inside interface on a VLAN trunk from the site distribution switch and an outside interface on the routed underlay. A VLAN-aware Linux bridge joins the inside trunk to a TAP interface owned by the overlay agent, which forwards frames to the hub over TLS. The hub behaves as a learning switch: it learns MAC addresses per link, forwards known unicast only to the owning spoke and floods broadcast, unknown-unicast and multicast (BUM) frames to every link.

Spokes never connect to each other. This keeps the Layer 2 topology loop-free by construction and removes any need to run spanning tree across the WAN.

2.1 Key design decisions

Decision	Rationale	Alternative considered
Encapsulate in TLS over TCP/443	Crosses every transit firewall and proxy without per-site exceptions; mutual certificate authentication is built in.	VXLAN over IPsec: better throughput, but needs UDP 4789 plus IKE/ESP opened across three providers.
Hub-and-spoke, not full mesh	No redundant L2 path, so no loop and no WAN spanning tree.	Full mesh with split-horizon: lower latency, much higher loop risk and operational complexity.
Manual hub failover	Automatic failover between two L2 hubs can make both active during a partial underlay outage (split brain), which is a loop.	Automatic failover with a witness: deferred until tooling is proven (Section 12).

3 Topology

Overlay GLOBAL-L2 (one broadcast domain): VLAN 100 10.100.0.0/21, VLAN 200 10.200.0.0/24, VLAN 999 10.199.0.0/24
Underlay: routed L3 transport, 10.0.0.0/16

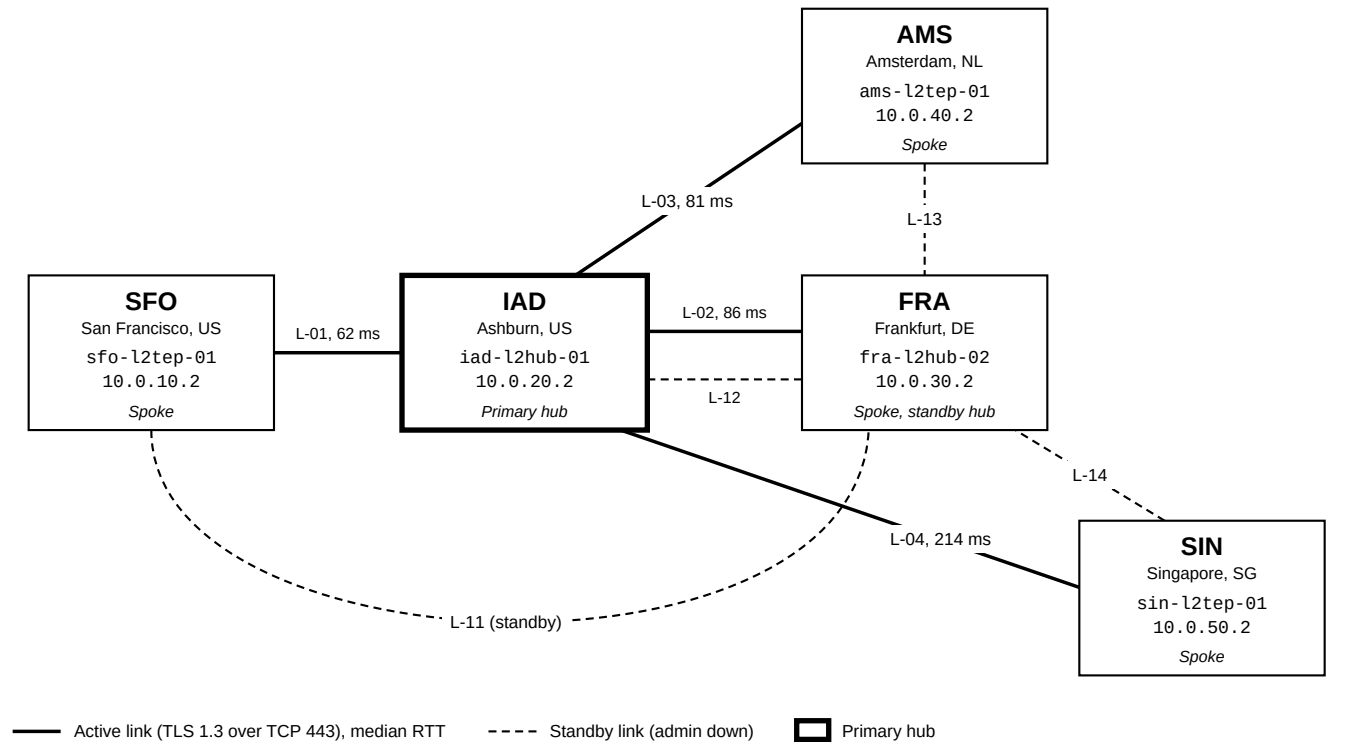


Figure 1. Logical topology. Solid lines are active overlay links with median underlay RTT; dashed lines are pre-provisioned standby links that stay administratively down until a hub failover (Section 8).

3.1 Site inventory

Site	Location	Role	TEP hostname	Underlay (eth0)	Management
SFO	San Francisco, US	Spoke	sfo-l2tep-01	10.0.10.2/29	10.10.10.10/24
IAD	Ashburn, US	Primary hub	iad-l2hub-01	10.0.20.2/29	10.10.20.10/24
FRA	Frankfurt, DE	Spoke + standby hub	fra-l2hub-02	10.0.30.2/29	10.10.30.10/24
AMS	Amsterdam, NL	Spoke	ams-l2tep-01	10.0.40.2/29	10.10.40.10/24
SIN	Singapore, SG	Spoke	sin-l2tep-01	10.0.50.2/29	10.10.50.10/24

3.2 Overlay link matrix

Link	From (spoke)	To (hub, TCP/443)	State	RTT	Commit	Sessions
L-01	sfo-l2tep-01	iad-l2hub-01	Active	62 ms	500 Mb/s	4
L-02	fra-l2hub-02	iad-l2hub-01	Active	86 ms	1 Gb/s	4
L-03	ams-l2tep-01	iad-l2hub-01	Active	81 ms	1 Gb/s	4
L-04	sin-l2tep-01	iad-l2hub-01	Active	214 ms	300 Mb/s	4
L-11	sfo-l2tep-01	fra-l2hub-02	Standby (admin down)	146 ms	500 Mb/s	4
L-12	iad-l2hub-01	fra-l2hub-02	Standby (admin down)	86 ms	1 Gb/s	4
L-13	ams-l2tep-01	fra-l2hub-02	Standby (admin down)	7 ms	1 Gb/s	4
L-14	sin-l2tep-01	fra-l2hub-02	Standby (admin down)	158 ms	300 Mb/s	4

3.3 Site-to-site latency over the overlay

RTT (ms)	SFO	IAD	FRA	AMS	SIN
SFO	-	62	150	145	278
IAD	62	-	86	81	214
FRA	150	86	-	169	302
AMS	145	81	169	-	297
SIN	278	214	302	297	-

Median RTT on VLAN 999 probe addresses (ICMP, 1,000 samples per pair, 2026-09-18). Spoke-to-spoke traffic transits the IAD hub. Direct underlay latency is compared in Section 12.

4 Encapsulation and Transport

Frames are carried unmodified, including the 802.1Q tag, inside a TLS 1.3 byte stream. Because the outer transport is TCP, the underlay never fragments overlay traffic: a full 1518-byte inner frame is split across TCP segments by the sender and reassembled by the receiver. Hosts therefore keep a standard 1500-byte MTU, and no MSS clamping is needed on the site LANs.

Original frame on the site LAN (VLAN 100)

Dst MAC	Src MAC	802.1Q	Type	IPv4 10.100.1.25 -> 10.100.5.40 + payload	FCS
6 B	6 B	VID 100	0x0800	46-1500 B	4 B

On the L3 underlay (link L-01)

Outer IPv4	TCP	TLS hdr	Frame hdr	Encapsulated Ethernet frame	AEAD tag
10.0.10.2 -> 10.0.20.2	:49812 -> :443	5 B	len + flags	up to 1518 B, FCS stripped	16 B

Cleartext, routed by the underlay

Encrypted and authenticated (TLS 1.3, AES-256-GCM)

Figure 2. Encapsulation. The site frame (top) becomes the payload of a TLS record carried in TCP/443 between tunnel endpoint addresses (bottom).

4.1 Tunnel parameters

Parameter	Value	Notes
Transport	TCP, destination port 443	Traverses stateful firewalls and proxies without exceptions
TLS version	1.3 only	TLS 1.2 and below refused at handshake
Cipher suite	TLS_AES_256_GCM_SHA384	Uses AES-NI offload on all TEPs
Authentication	Mutual X.509, ECDSA P-384	Issued by internal CA EXAMPLE-NET-CA-G2 [REDACTED]
Parallel sessions	4 per link	Per-flow hashing preserves frame order within a flow
Keepalive	10 s interval, dead after 3 missed	Link declared down after 30 s
Reconnect backoff	5 s initial, doubling, 60 s max	±20% jitter avoids synchronized reconnects
Compression	Disabled	No gain on already-encrypted application payloads
Inner MTU	1500 (1522-byte tagged frame)	Jumbo frames are not supported across the overlay
Overhead	About 3-5% at full-size frames	TLS record plus TCP/IP headers, amortized over the stream

5 IP Addressing Plan

All addressing is drawn from 10.0.0.0/8. Blocks are allocated by function in the second octet and by site in the third octet using the site index (SFO = 10, IAD = 20, FRA = 30, AMS = 40, SIN = 50).

5.1 Allocation summary

Block	Function	Notes
10.0.0.0/16	Underlay transport	TEP outside interfaces, one /29 per site
10.10.0.0/16	Out-of-band management	One /24 per site; NOC tooling in 10.10.20.0/24
10.100.0.0/21	VLAN 100 L2-APP (stretched)	One subnet across all five sites
10.199.0.0/24	VLAN 999 L2-PROBE (stretched)	TEP monitoring addresses
10.200.0.0/24	VLAN 200 L2-DBREPL (stretched)	Not routed
Remainder of 10.0.0.0/8	Reserved	Not used by the overlay

5.2 Underlay (transport)

Site	Subnet	Gateway (provider edge)	TEP eth0	Reserved
SFO	10.0.10.0/29	10.0.10.1	10.0.10.2	10.0.10.3 - .6
IAD	10.0.20.0/29	10.0.20.1	10.0.20.2	10.0.20.3 - .6
FRA	10.0.30.0/29	10.0.30.1	10.0.30.2	10.0.30.3 - .6
AMS	10.0.40.0/29	10.0.40.1	10.0.40.2	10.0.40.3 - .6
SIN	10.0.50.0/29	10.0.50.1	10.0.50.2	10.0.50.3 - .6

Each /29 is advertised into the provider transport. The only reachability the overlay needs is TCP/443 from every TEP to the hub and standby hub addresses (Section 11).

5.3 Stretched overlay segments

VLAN	Name	Subnet	Default gateway	Routed	Purpose
100	L2-APP	10.100.0.0/21	10.100.0.1 (VRRP VIP)	Yes, at IAD (FRA backup)	Application tier, workload mobility
200	L2-DBREPL	10.200.0.0/24	None	No	Cluster heartbeat and replication
999	L2-PROBE	10.199.0.0/24	None	No	TEP reachability probes

5.4 VLAN 100 allocation by site

Range	Assigned to	Notes
10.100.0.0/24	Shared infrastructure	.1 VRRP VIP; .2 iad-gw01; .3 iad-gw02; .4 fra-gw01; .5 fra-gw02; .10 DHCP server (IAD)
10.100.1.0/24	SFO workloads	DHCP pool .50 - .250
10.100.2.0/24	IAD workloads	DHCP pool .50 - .250
10.100.3.0/24	FRA workloads	DHCP pool .50 - .250
10.100.4.0/24	AMS workloads	DHCP pool .50 - .250
10.100.5.0/24	SIN workloads	DHCP pool .50 - .250
10.100.6.0/23	Reserved	Growth

Note: Per-site ranges are administrative only. The /21 is a single broadcast domain; a workload keeps its address when it migrates, so the range shows where it was provisioned, not where it runs today.

5.5 VLAN 200 and probe assignments

Site	VLAN 200 host range	VLAN 999 probe (TEP br-l2.999)
SFO	10.200.0.10 - 10.200.0.29	10.199.0.11
IAD	10.200.0.30 - 10.200.0.49	10.199.0.21
FRA	10.200.0.50 - 10.200.0.69	10.199.0.31
AMS	10.200.0.70 - 10.200.0.89	10.199.0.41
SIN	10.200.0.90 - 10.200.0.109	10.199.0.51

6 Tunnel Endpoint Build

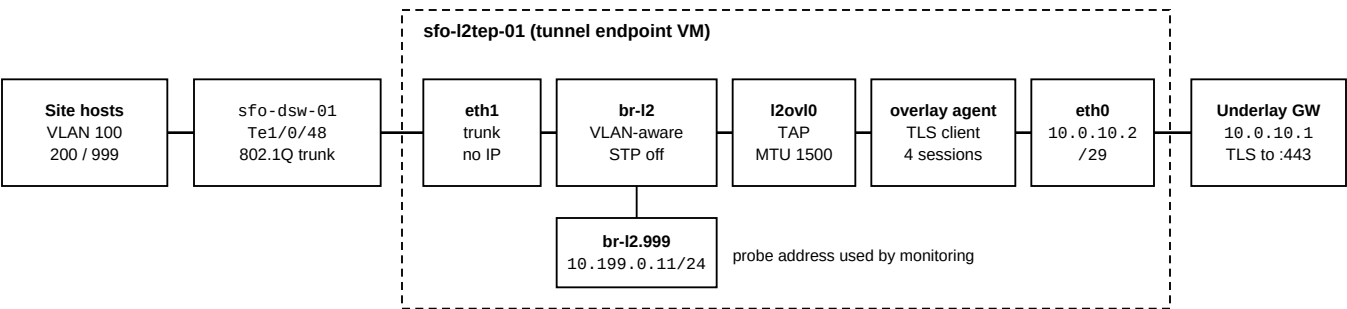


Figure 3. Tunnel endpoint internals (SFO shown). All sites are identical apart from addressing.

6.1 Interfaces

Interface	Type	Address	Purpose
eth0	vNIC, underlay	10.0.<site>.2/29	TLS sessions to the hub(s)
eth1	vNIC, trunk	none	802.1Q trunk from the distribution switch, VLANs 100 / 200 / 999
br-l2	VLAN-aware bridge	none	Joins eth1 and l2ovl0; STP disabled
l2ovl0	TAP	none	Frame interface of the overlay agent
br-l2.999	VLAN sub-interface	10.199.0.<n>/24	Probe address for monitoring
mgmt0	vNIC, out-of-band	10.10.<site>.10/24	SSH, SNMPv3, syslog

Hypervisor requirement: the port group behind eth1 must allow promiscuous mode, MAC address changes and forged transmits. Without these, the hypervisor drops frames whose source MAC belongs to a remote site and the overlay appears up while passing no traffic.

6.2 Configuration baseline

/etc/l2ovl/bridge-up.sh (sfo-l2tep-01, sanitized)

```
ip link add br-l2 type bridge vlan_filtering 1 stp_state 0 ageing_time 30000 fdb_max_learned 4096
ip link set eth1 master br-l2
ip link set l2ovl0 master br-l2
for p in eth1 l2ovl0; do
  bridge vlan del dev $p vid 1
  for v in 100 200 999; do bridge vlan add dev $p vid $v; done
done
bridge vlan add dev br-l2 vid 999 self
ip link add link br-l2 name br-l2.999 type vlan id 999
ip addr add 10.199.0.11/24 dev br-l2.999
for i in eth1 l2ovl0 br-l2 br-l2.999; do ip link set dev $i mtu 1500 up; done
```

/etc/l2ovl/links.conf (sfo-l2tep-01; agent name and syntax sanitized)

```
[link L-01]
role          = spoke
switch_instance = GLOBAL-L2
remote        = 10.0.20.2:443
local_bind    = 10.0.10.2
tap_interface  = l2ovl0
tls_min_version = 1.3
cipher        = TLS_AES_256_GCM_SHA384
cert          = /etc/l2ovl/pki/sfo-l2tep-01.crt
key           = /etc/l2ovl/pki/sfo-l2tep-01.key           ; [REDACTED]
ca_bundle     = /etc/l2ovl/pki/example-net-ca-g2.pem
sessions      = 4
keepalive     = 10s
dead_after    = 3
admin_state   = up

[link L-11]
role          = spoke
switch_instance = GLOBAL-L2
remote        = 10.0.30.2:443
local_bind    = 10.0.10.2
tap_interface  = l2ovl0
sessions      = 4
admin_state   = down                                ; standby, see Section 8
```

sfo-dsw-01, port facing the tunnel endpoint

```
interface TenGigabitEthernet1/0/48
description L2OVL sfo-l2tep-01 eth1 (trunk)
switchport mode trunk
switchport trunk native vlan 4000
switchport trunk allowed vlan 100,200,999
spanning-tree portfast trunk
spanning-tree bpdupfilter enable
storm-control broadcast level 1.00
storm-control multicast level 2.00
storm-control unicast level 2.00
storm-control action trap
```

7 Loop Prevention and BUM Control

A stretched broadcast domain has no TTL: one loop anywhere replicates frames at line rate to every site. The design relies on five layered controls instead of running spanning tree across the WAN.

Control	Where	Setting	Effect
Hub-and-spoke only	Overlay agents	Spokes accept no inbound links; standby links admin down	No redundant L2 path exists in normal operation
STP isolation	Site switch port, TEP bridge	bpdufilter enable; stp_state 0	Each site stays its own STP domain; WAN events cannot trigger site re-convergence
Storm control	Site switch port	Broadcast 1%, multicast 2%, unknown unicast 2%	Caps BUM entering the overlay at 100-200 Mb/s per site on a 10G port
MAC limit	TEP bridge	fdb_max_learned 4096, ageing 300 s	Contains MAC floods and accidental bridging
VLAN pruning	Site switch trunk	Allowed 100, 200, 999; native VLAN 4000 unused	Site-local VLANs are never extended by accident

Warning: never raise a standby link while a primary link is up. Enabling any of L-11 to L-14 while L-01 to L-04 are up gives two hubs accepting spokes at the same time, which is a Layer 2 loop across the WAN. Section 8 gives the only approved order of operations.

8 Redundancy and Failover

The overlay has one active hub. Hub failover is deliberate and manual (see Section 2.1). FRA was chosen as standby hub for its 7 ms path to AMS and acceptable latency to SFO (146 ms) and SIN (158 ms).

Triggers:

- iad-l2hub-01 unreachable on all of L-01 to L-04 for more than 5 minutes, confirmed from at least two spokes; or
- planned IAD maintenance under an approved change record.

8.1 Failover procedure (IAD to FRA)

Step	Action	Where	Verify
1	Confirm hub loss from two spokes and open an incident	NOC	l2ovl-cli link show: L-0x DOWN at two or more sites
2	Set L-02 admin down so FRA stops being a spoke	fra-l2hub-02	L-02 DOWN
3	Enable the hub listener for GLOBAL-L2	fra-l2hub-02	Listening on 10.0.30.2:443
4	One spoke at a time (SFO, AMS, SIN): set L-0x admin down, then L-1x admin up	Spokes	L-1x ESTABLISHED, sessions 4/4
5	If the IAD site is up but its hub node was repaired: bring up L-12	iad-l2hub-01	L-12 ESTABLISHED
6	Confirm the gateway moved: fra-gw01 is VRRP master for 10.100.0.1	fra-gw01	VRRP state Master
7	Run the probe matrix (test T1, Section 9)	NOC	0% loss on all probe pairs

Failback reverses the order in a maintenance window: drop each spoke's L-1x before raising its L-0x, and disable the FRA hub listener last. The 2026-09-18 drill measured 3 min 40 s from step 2 to step 7.

8.2 Default gateway redundancy (VLAN 100)

Router	Site	Address	VRRP priority	Normal role
iad-gw01	IAD	10.100.0.2	120	Master for 10.100.0.1
iad-gw02	IAD	10.100.0.3	110	Backup
fra-gw01	FRA	10.100.0.4	100	Backup; master after a hub move
fra-gw02	FRA	10.100.0.5	90	Backup

Preemption is enabled with a 300 s delay so that a recovering IAD gateway does not take the VIP back before the overlay has settled.

9 Validation Results

Acceptance tests run on 2026-09-18 after the SIN onboarding and TLS 1.3 migration.

ID	Test	Path	Criterion	Result
T1	ICMP full mesh on VLAN 999, 1,000 packets per pair	All 10 site pairs	0% loss; RTT within 10% of Table 3.3	PASS: 0% loss; max 302 ms (FRA-SIN)
T2	ARP resolution across the overlay	10.100.4.21 (AMS) to 10.100.1.25 (SFO)	Resolves in under 1 s	PASS: 148 ms
T3	Full-size frame, DF set (ping -M do -s 1472)	FRA to SIN probe	No fragmentation or loss	PASS
T4	Throughput, iperf3 -P 8, 60 s	AMS to IAD	800 Mb/s or more	PASS: 912 Mb/s
T5	Throughput, iperf3 -P 8, 60 s	SIN to IAD	250 Mb/s or more	PASS: 268 Mb/s
T6	Segment isolation	VLAN 200 host to VLAN 100 host	No reachability	PASS
T7	Live migration keeps 10.100.1.25	SFO to IAD	Fewer than 3 pings lost	PASS: 1 lost (gratuitous ARP)
T8	Planned hub failover to FRA	All spokes	Under 15 min	PASS: 3 min 40 s
T9	Loop injection, lab replica only	L-11 raised with L-01 up	Storm-control trap within 5 s	PASS: trap at 1.2 s; BUM capped

10 Operations Runbook

10.1 Health checks

Check	Command (on the TEP)	Expected
Agent running	systemctl is-active l2ovl-agent	active
Links established	l2ovl-cli link show	Active links ESTABLISHED, sessions 4/4
Bridge ports forwarding	bridge link show master br-l2	eth1 and l2ovl0 forwarding
MAC table size	bridge fdb show br br-l2 dynamic wc -l	Under 3,000 (limit 4,096)
Probe reachability	ping -c 20 -I br-l2.999 10.199.0.21	0% loss
Path MTU	ping -M do -s 1472 -c 3 10.199.0.51	Replies; no "message too long"
Underlay to hub	nc -zv -w 3 10.0.20.2 443	Connection succeeded
Certificate expiry	openssl x509 -enddate -noout -in /etc/l2ovl/pki/<host>.crt	More than 30 days remaining

The NMS (10.10.20.50) runs these checks every 60 s. Any link down for more than 2 minutes raises a P2; the hub unreachable from two or more spokes raises a P1 and pages the on-call network engineer.

10.2 Troubleshooting

Symptom	Likely cause	Action
Link drops and re-establishes every few minutes	Firewall or NAT idle timeout on the underlay; underlay packet loss	Confirm TCP idle timeout of 3,600 s or more on transit firewalls; run mtr to the hub
Throughput collapses during congestion	TCP-over-TCP retransmission when underlay loss exceeds about 0.5%	Confirm underlay loss; verify sessions 4/4; open a provider ticket
Small packets pass, large transfers stall	MTU below 1500 on the site trunk or hypervisor port group	Run the path MTU check; confirm 1500 (1522 tagged) on switch trunk, port group and eth1
Remote hosts in ARP table but unreachable	Port group rejecting frames with foreign source MACs	Allow promiscuous mode, MAC changes and forged transmits on the trunk port group
Storm-control traps at several sites	Loop: a standby link raised while a primary is up	Set all L-1x links admin down immediately; review recent changes
New hosts cannot communicate, existing hosts fine	MAC limit of 4,096 reached on a TEP bridge	Use bridge fdb show to find the port with excess MACs; remove the rogue bridge
Duplicate-IP warnings after a migration	Stale ARP on the gateway; VM did not send gratuitous ARP	arping -U from the VM; clear the ARP entry on iad-gw01/02

11 Security Controls

- Link authentication is certificate-only; no shared secrets or passwords exist on any link. Certificates are ECDSA P-384, valid 365 days and rotated at day 300 under a change record. The issuing CA is offline.
- Private keys are root-owned, mode 0600, never leave the TEP and are [REDACTED] in all exported configurations.
- Only the hub and standby hub listen on TCP/443. Spokes open no inbound ports on eth0.
- Only internal VLANs are bridged. Internet-facing, DMZ and cardholder-data segments must never be added to the trunk; any VLAN change needs design-authority approval.
- TEP management is reachable only on mgmt0: SSH keys, SNMPv3 authPriv, syslog to 10.10.20.50.
- Link up/down events, handshake failures and certificate errors are forwarded to the SIEM [REDACTED].

11.1 Underlay access control

Rule	Source	Destination	Service	Action
10	10.0.10.2, 10.0.30.2, 10.0.40.2, 10.0.50.2	10.0.20.2 (hub)	TCP/443	Permit
20	10.0.10.2, 10.0.20.2, 10.0.40.2, 10.0.50.2	10.0.30.2 (standby hub)	TCP/443	Permit
30	10.10.20.0/24 (NOC)	10.10.0.0/16 (TEP mgmt0)	TCP/22, UDP/161	Permit
40	10.10.0.0/16 (TEP mgmt0)	10.10.20.50 (syslog, NMS)	UDP/514, UDP/162	Permit
99	any	any TEP address	any	Deny, log

12 Known Limitations and Roadmap

Limitation	Impact	Mitigation / roadmap
Spoke-to-spoke hairpin through IAD	FRA-AMS is 169 ms over the overlay versus 7 ms direct; SFO-SIN is 278 ms versus 171 ms direct	Q1 2027: assess an EMEA regional hub, subject to a loop-safety design review
Single gateway location	Routed traffic from SIN trombones through IAD (+214 ms)	Local routed egress for north-south traffic; keep the overlay for east-west L2 needs
TCP-over-TCP	Above about 0.5% underlay loss, inner and outer TCP both retransmit and throughput collapses	Four sessions per link; underlay QoS marking AF31; loss alarm at 0.2%
One broadcast domain across five sites	BUM from any site reaches every site; ARP volume grows with host count	/21 cap, storm control; design review at 1,500 hosts (currently 612)
Manual failover	RTO depends on NOC response time	Q4 2026: scripted failover with pre-checks
Hub capacity	4 Gb/s TLS ceiling on iad-l2hub-01	Peak 1.6 Gb/s (40%); scale-up trigger at 70% sustained

Appendix A Revision History

Rev	Date	Author	Change
1.0	2025-11-03	[REDACTED]	Initial design: SFO, IAD, FRA
1.1	2026-02-14	[REDACTED]	Added AMS; FRA designated standby hub; standby links L-11 to L-13 provisioned
2.0	2026-06-30	[REDACTED]	Added SIN (L-04, standby L-14); TLS 1.3 only; sessions per link raised from 2 to 4
2.1	2026-09-24	[REDACTED]	Failover drill and acceptance results; document sanitized for external sharing

Appendix B Glossary

Term	Definition
BUM	Broadcast, unknown-unicast and multicast traffic; flooded to every link
FDB	Forwarding database: the bridge MAC address table
Hairpin	Traffic between two spokes that must transit the hub
Hub	Tunnel endpoint that accepts spoke links and switches frames between them
L2VPN	A VPN that carries Ethernet frames, making remote hosts appear on the same LAN
MTU / MSS	Maximum transmission unit / maximum TCP segment size
Overlay	The stretched Layer 2 network carried inside the tunnels
Spoke	Tunnel endpoint that keeps exactly one active link, to the hub
TEP	Tunnel endpoint: the per-site VM that bridges site VLANs into the overlay
Underlay	The routed Layer 3 network that carries the tunnels
VRRP	Virtual Router Redundancy Protocol; provides the 10.100.0.1 gateway address