NV03 OAM DT Status report

Reported by Erik Nordmark

- Started by the WG chairs at IETF94
- Invited members (in alphabetical order):
  - Ignas Bagdonas
  - Matthew Bocci
  - Philips Chenhao
  - Saumya Dikshit
  - Diego Garcia del Rio
  - Anoop Ghanwani
  - Pradeep Jain
  - Deepak Kumar (lead)
  - Erik Nordmark
  - Reshad Rahman
  - Benson Schliesser
  - Ravi Shekhar
  - Tina Tsou
  - Li Yizhou
- Scheduled weekly phone calls

NVO3 DT Activity

- Deployment Scenario and Use Cases
- Data Packet Encapsulation Scenarios
- Different OAM Proposal(s) Header Format
- Use case of Reserved Extra OAM bit (Marking Bit)
- Questions for Working Group (Overlay Data plane Encapsulation)

Deployment Scenarios and Use cases

Data Center (Including Call Flows and Packet Encapsulation)	Service Provider
Leaf-&-Spine (3-tier fat tree) – Massively Scalable Data Center (MSDC): Leafs are L2/L3 NVE Edge with SVI configuration for routing.	NVO3 tunnel as a service chaining building component
Two Tier Leaf-&-Spine Data Center Architecture	Overlay tunnel as a service handoff mechanism
Leaf-&-Spine (3-tier fat tree) – Massively Scalable Data Center (MSDC): Leafs are L2 NVE Edge without any SVI configuration for routing. Super Spines are Default Gateway	Overlay tunnel with same functionality as L3VPN
Intra-POD connect (L2) with and without Redundancy Flows	Redundancy active-active (all links are used, all links are available) as compared to active-standby
Inter-POD connect (L2) with and without Redundancy flows	Interworking: • NVO3 – IP/MPLS – TRILL (Define Interworking) – It goes over existing transport • NVO3 – Pseudowire – MPLS – NVO3 (Interworking) • NVO3 – LISP – NVO3
Intra-DC L3 connect (same DC-core network)	
Inter-DC L3 connect over DCI with NVO3 encapsulation supported	
L2 or L3 connects over DCI (Wan gateway) with non-NVO3 encapsulation supported in remote DC-core	
SDN based DC deployment topologies	

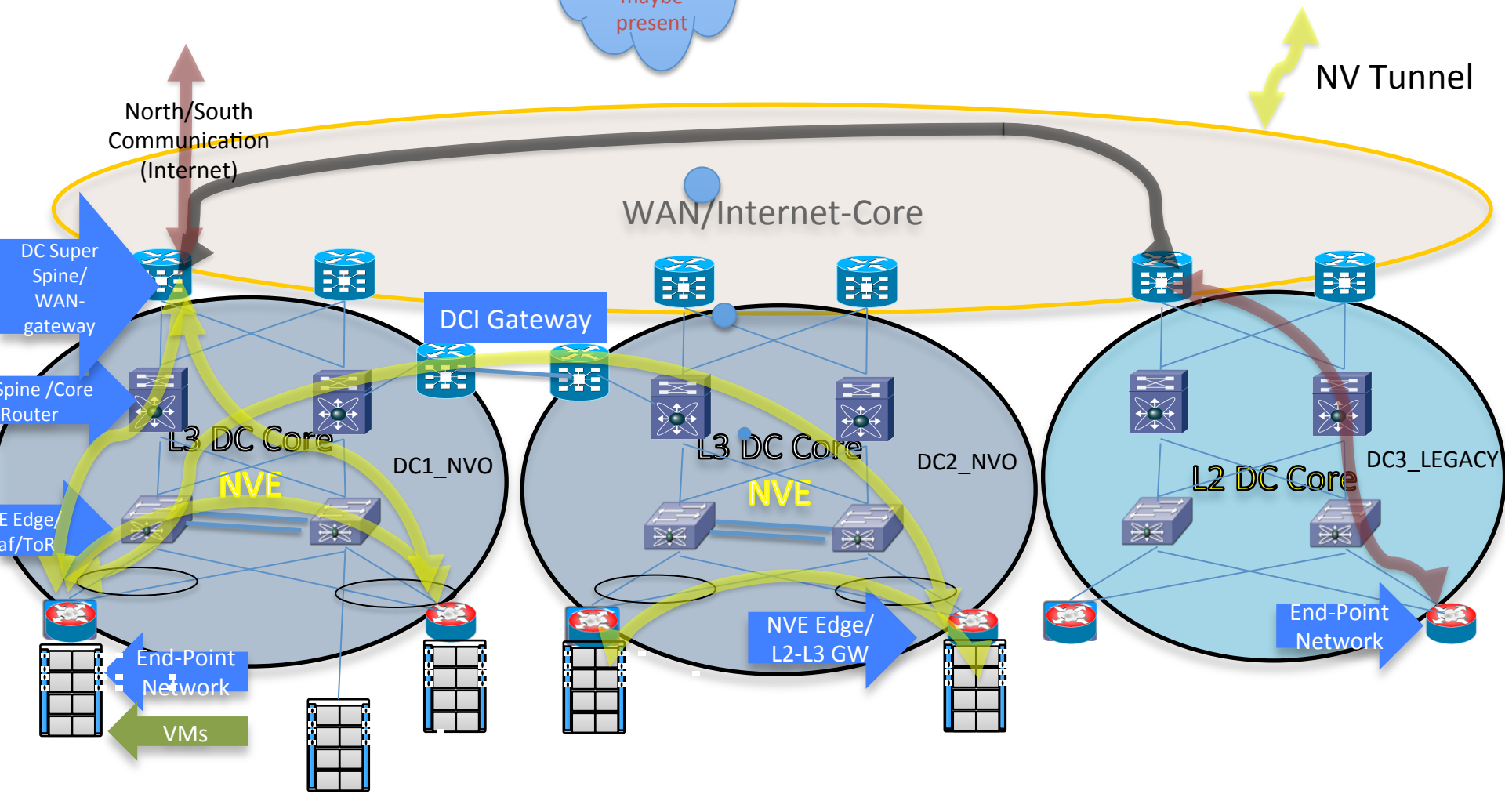
Datacenter Layout

Encapsulation end-points (NVE)

Intra-DC : ToR/Leaf – ToR/Leaf
ToR/Leaf– WAN-gateway
ToR/Leaf- DCI gateway

DC – EVPN/ Flood & Learn/ Ingress Replication Controller maybe present

Inter-DC: ToR – ToR
ToR/vPE – DC Edge Router(non vxlan)

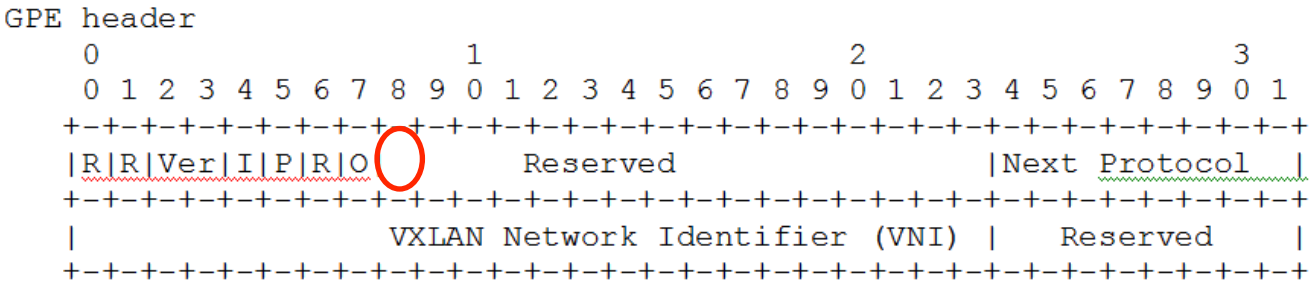


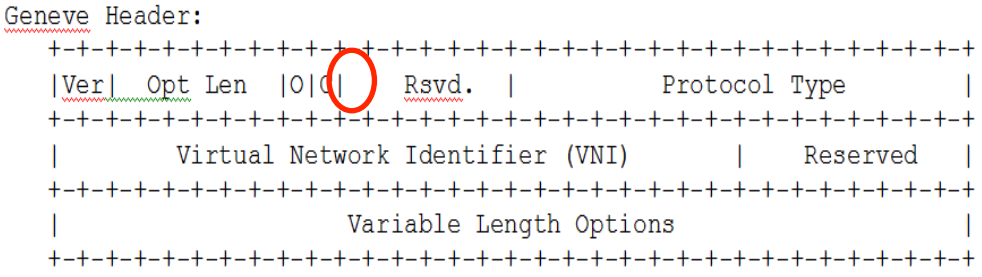
Generic Queries Across All Deployments

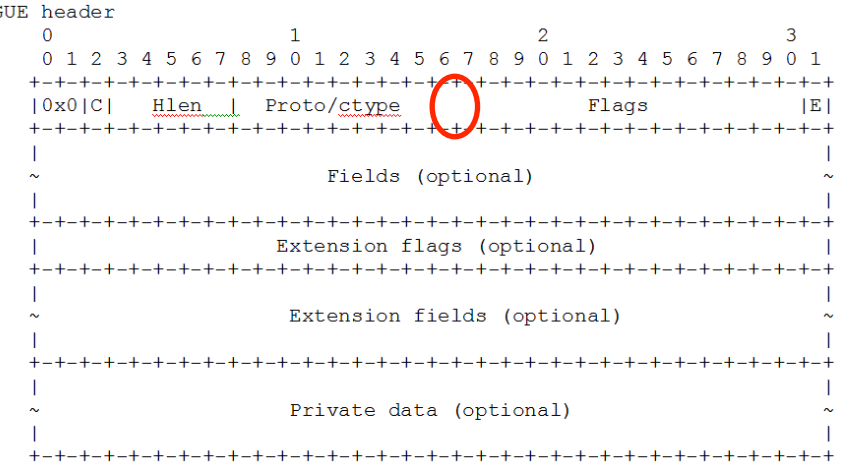
Generic Open Items and Queries
Knowledge about Underlay Capabilities or agnostic to same? Open item for inter-op with other layer OAM (application level and underlay)
Tunnel Depth in the core. Or should it be generic and unaffected
Scope of Administrative Domains?
Congruent support across all NVO3 encapsulation to carry and identify OAM Common Header Encapsulation?
Operational Mechanics shouldn't change based on overlay Mechanism • Generic Requirements and verify whether they can be covered by all encapsulation • pro-active BFD use case from Customer Edge Point • Scenario to detect high loss link due to optics issues
SFC OAM <-> Overlay OAM Interworking

M bit – M for marking

- O bit – OAM control message
- M bit - Marking bit is normally set in the real user traffic. Forwarding should be as usual. Affects counters.







OAM Functionality

- Continuity Check
- Continuity Verification/Reachability Verification
- Trace/Path Verification
  - Underlay visibility is important (Can't just rely on TTL expiry for all scenarios)
  - Telemetry with Trace/Path
- ECMP Verification
- Path MTU
- Delay/Loss
  - Delay and Loss using markers
- Logging/Tracing – (Which Events that need to be logged, @ NVE and NVA)
- Control Plane Verification (EVPN Scenario)
- Multicast Tracing
- Interworking
- Telemetry

OAM Headers

- Need OAM identification bits and few extra reserved bit for Marking and Performance in Overlay Encapsulation
- Need OAM Common Header format if possible
- Re-use Existing functionality if possible
  - Optional "O" bit inference for underlay
    - Fragmentation
    - Provide architecture reference
    - Difficult to define protocol procedures

Question For WG

- Do we want consistent definitions/semantics across the different encapsulations?
  - Encoding can differ to fit the different encapsulation headers

- Need for OAM payload type for out-of-band OAM? Or just OAM bit?
- Does C-bit in GUE mean the same as O-bit in other encapsulations?
- Can intermediate (underlay) nodes look at OAM packets and participate? (Geneve doesn't allow this)
- Underlay error reports? (All encapsulations are silent on this)
- Able to set OAM bit on normal payload for in-band OAM? (GUE makes it exclusive)
- Are packets with OAM bit set always dropped by destination NVE?
- No impact on ECMP/LAG hashing due to OAM bit (Explicit in Geneve)
- Define a common M-bit?
  - Don't want different definition of marking bits in different encapsulations.
- Need the same definitions of the above for all the encapsulations?

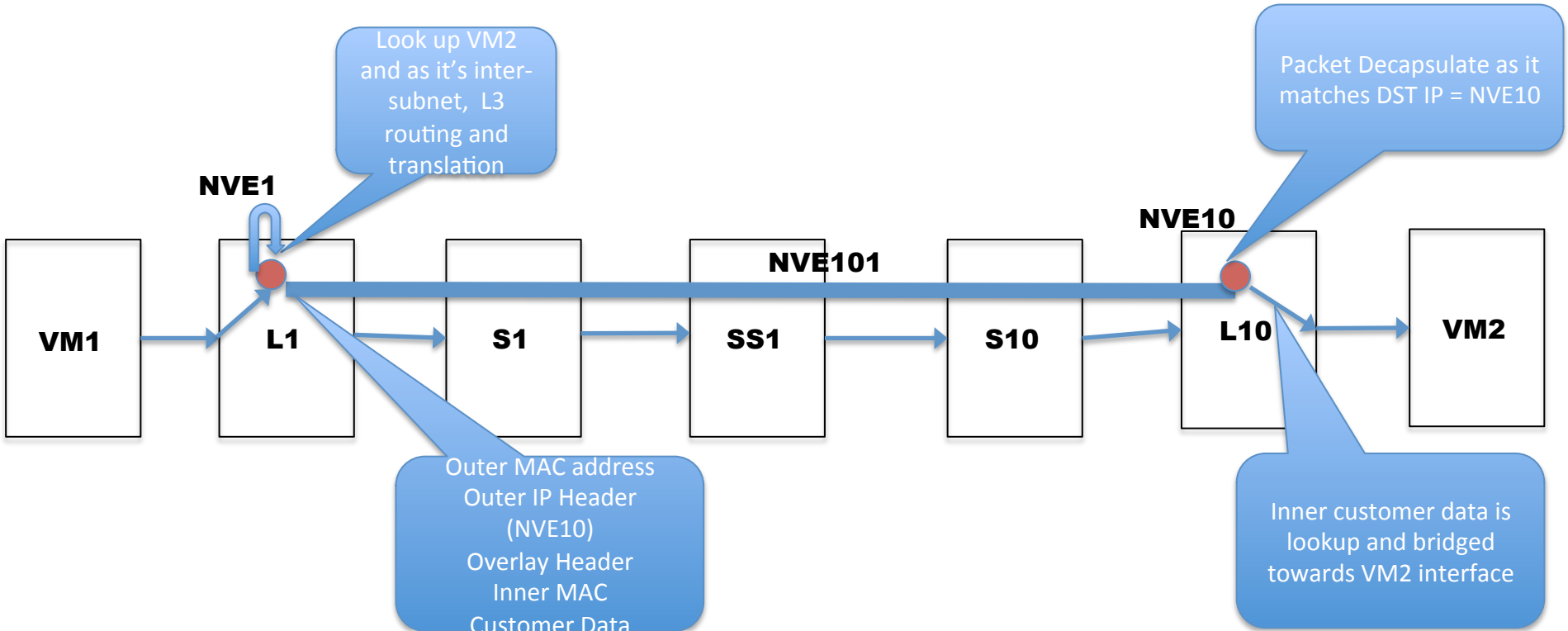
Next Steps

- Working more closely with RTG area Overlay OAM team
- Look more at BFD
- Document for OAM Functionality, Encapsulation, Procedure for Geneve, GUE, and GPE functionality.

Backup Slides

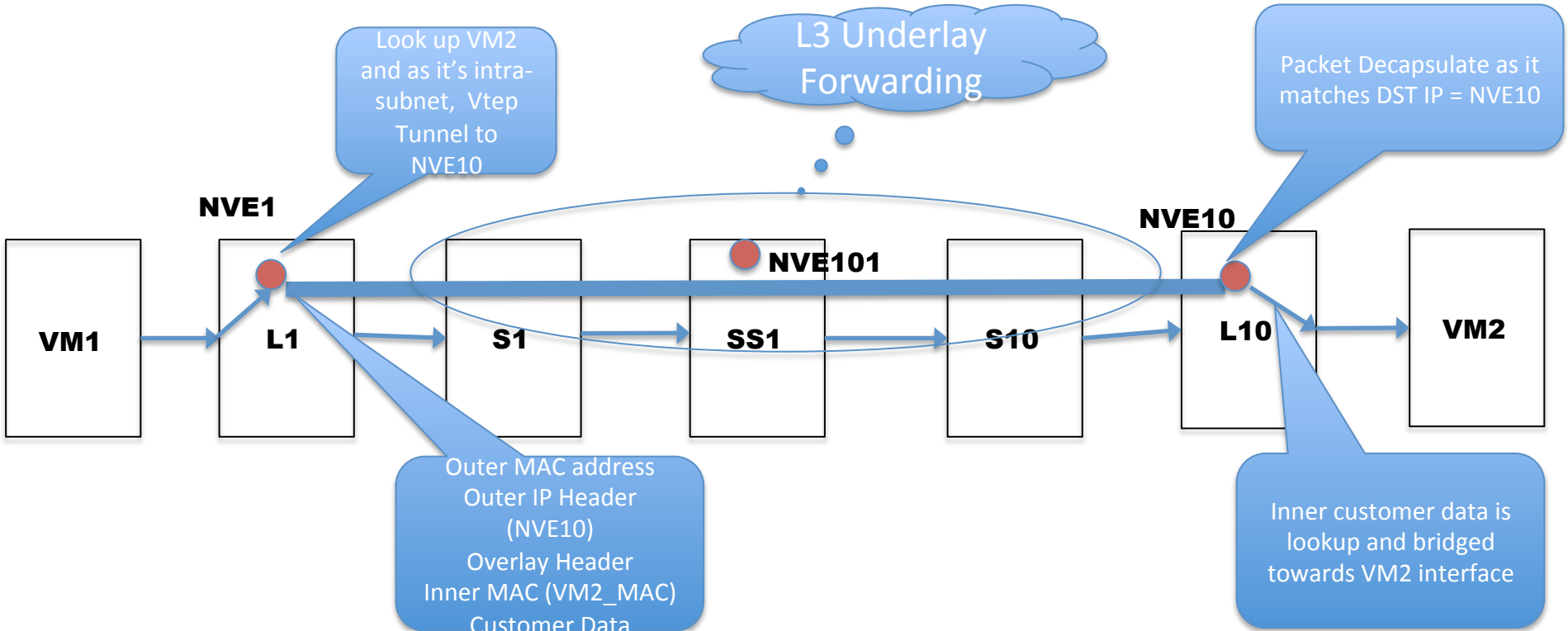
Packet Encapsulation Flow – inter subnet

- Leaf has both L2/L3 routes
- Packet Destined to VM2 on L1, goes through L3 translation to find right overlay VNI for multi –tenacy
- Overlay tunnel generated towards NVE10, packet gets encapsulated with Outer IP, udp header, overlay header, and customer data.



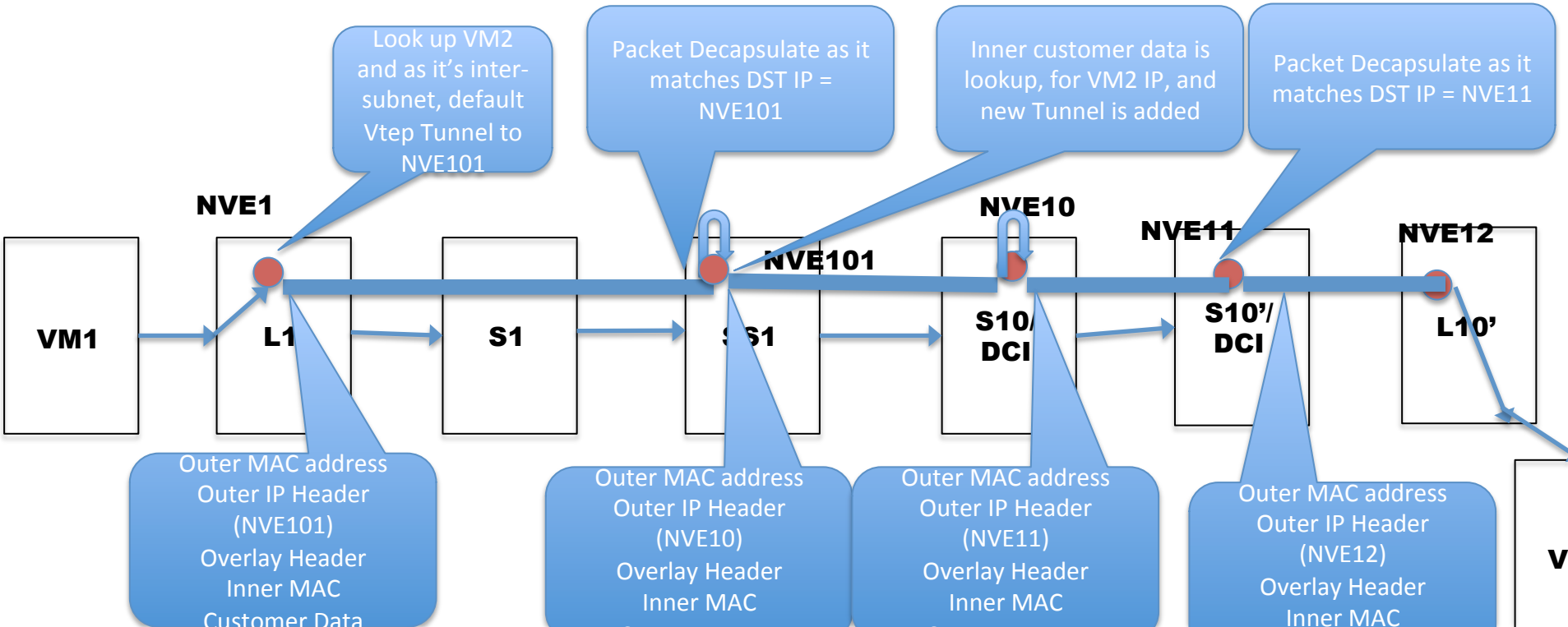
Packet Encapsulation Flow – Simple Tunnel

- Tor is L2 NVE Interface with no SVI
- Spine is Underlay Switch with L3 ECMP
- Super Spine has NVE Interface and also SVI for routing
- Scenario 2 - Intra-subnet Routing between VM across POD (No physical connectivity through spine as it's across POD)

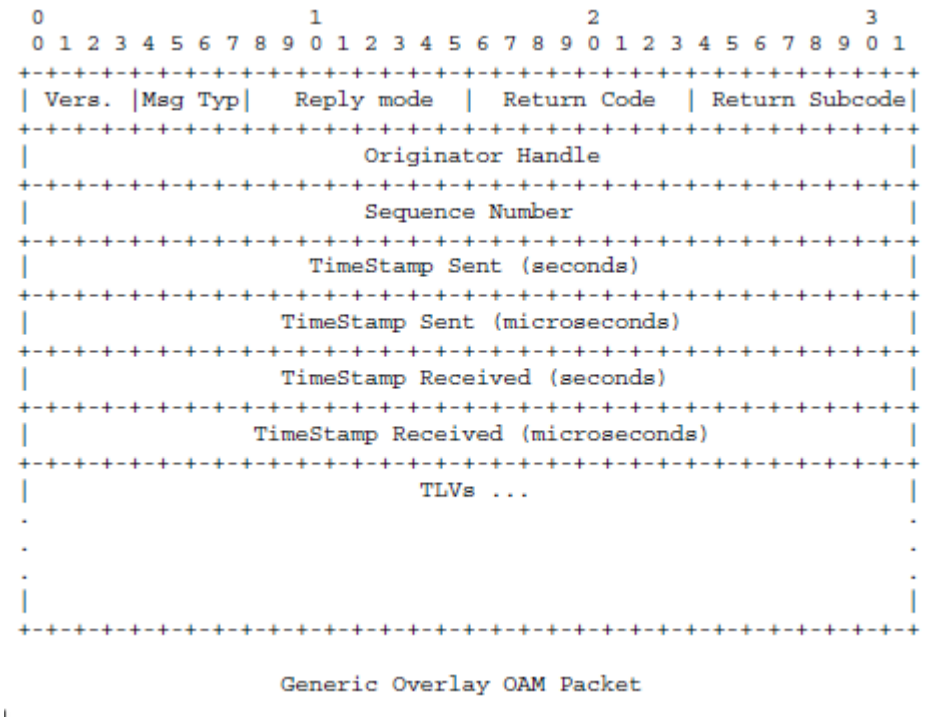


Packet Encapsulation – (Tunnel Stitching Scenario)

- The L3 connect will traverse an L3-gateway for NVO encapsulation to perform the routing. The DC-Super Spine (WAN gateway) also acts as NVO L3 gateway with the same POD. Effectively 4-MEPs in this Path
- ToR Switches: NVO L2-gateway
- Super-Spine: NVO L3-gateway
- DCI-gateway: NVO tunnel End Points Within DC.



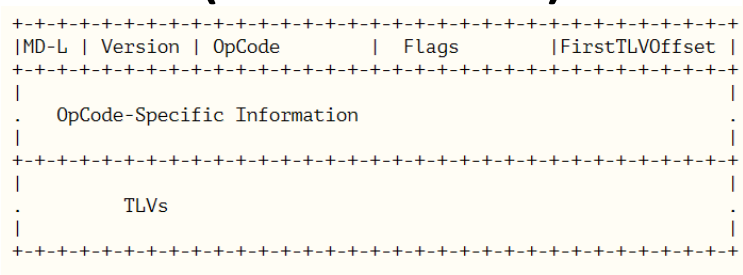
OAM Common Header in draft-jain-nvo3-overlay-oam



Generic Overlay OAM Packet

Other OAM Common Header for reference

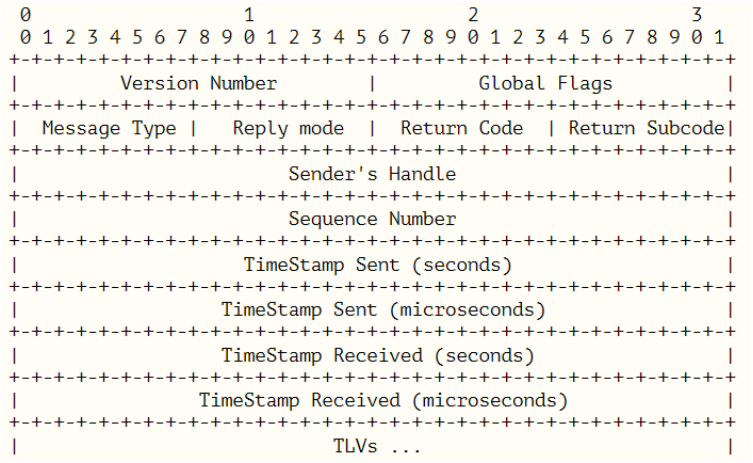
- Trill (RFC 7455)



Type	Name of TLV in [8021Q]
0	End TLV
1	Sender ID TLV
2	Port Status TLV
3	Data TLV
4	Interface Status TLV
5	Reply Ingress TLV
6	Reply Egress TLV
7	LTM Egress Identifier TLV
8	LTR Egress Identifier TLV
9-30	Reserved
31	Organization Specific TLV

Type	TLV Name
64	TRILL OAM Application Identifier TLV
65	Out-of-Band Reply Address TLV
66	Diagnostic Label TLV
67	Original Data Payload TLV
68	RBridge Scope TLV
69	Previous RBridge Nickname TLV
70	Next-Hop RBridge List TLV
71	Multicast Receiver Port Count TLV
72	Flow Identifier TLV
73	Reflector Entropy TLV
74	Authentication TLV

- MPLS (LSP ping RFC 4379) – use UDP



Other OAM Common Header for reference

- draft pang (out of band solution)