cisco The bridge to possible

VXLAN or GENEVE

What is a better choice for your Data Center

Lukas Krattiger, Distinguished Engineer
@CCIE21921

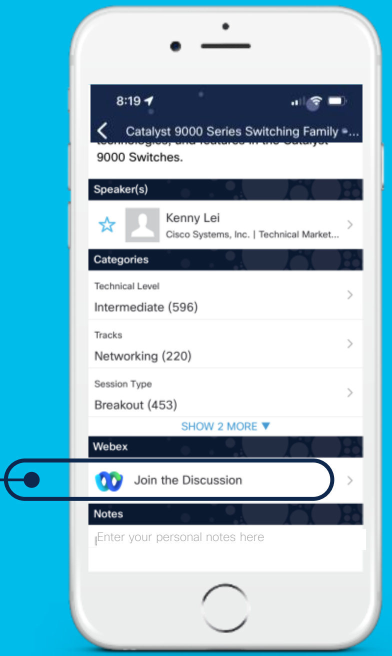
Cisco Webex App

Questions?

Use Cisco Webex App to chat
with the speaker after the session

How

1. Find this session in the Cisco Live Mobile App
2. Click “Join the Discussion”
3. Install the Webex App or go directly to the Webex space
4. Enter messages/questions in the Webex space



8:19

Catalyst 9000 Series Switching Family •...

9000 Switches.

Speaker(s)

Kenny Lei
Cisco Systems, Inc. | Technical Market...

Categories

Technical Level
Intermediate (596)

Tracks
Networking (220)

Session Type
Breakout (453)

SHOW 2 MORE ▼

Webex

Join the Discussion

Notes

Enter your personal notes here

Webex spaces will be moderated
until February 24, 2023.

Abstract

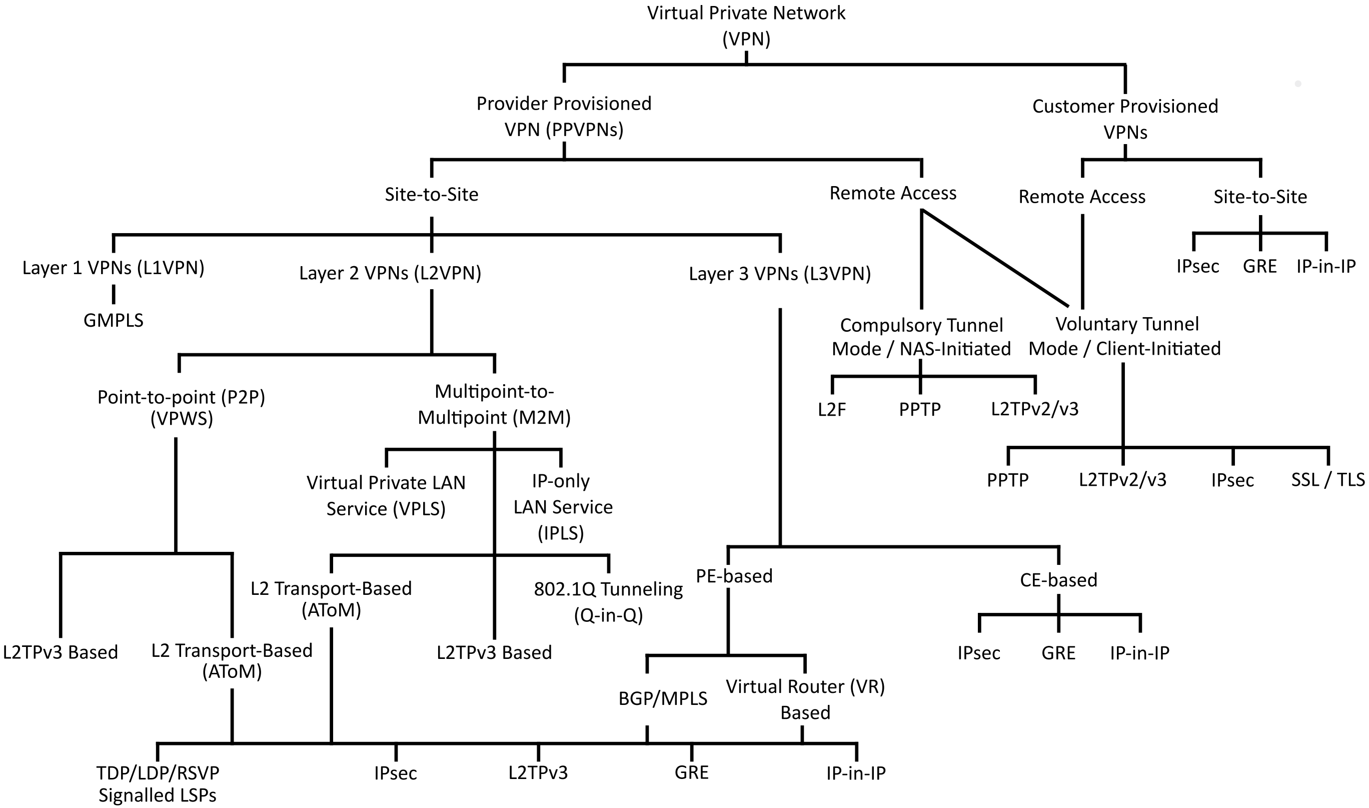
Many network engineers and architects want to know what the right overlay is for their data center fabrics. In this session we will discuss the advantages and capabilities of both VXLAN and GENEVE based overlays for production data centers. We will cover what the real world use cases are. We will begin with a network virtualization primer, and discuss what is covered in the RFCs for both encapsulations, and what is supported in switching hardware and software. We will also touch on security and service insertion implications of both encaps. At the end of this session the attendee will be able to determine which encap is right for their implementation.

Agenda

- Network Virtualization Primer
- Encapsulation Overview
  - GENEVE
  - VXLAN
- Use-Cases
- Score Board
- Conclusion

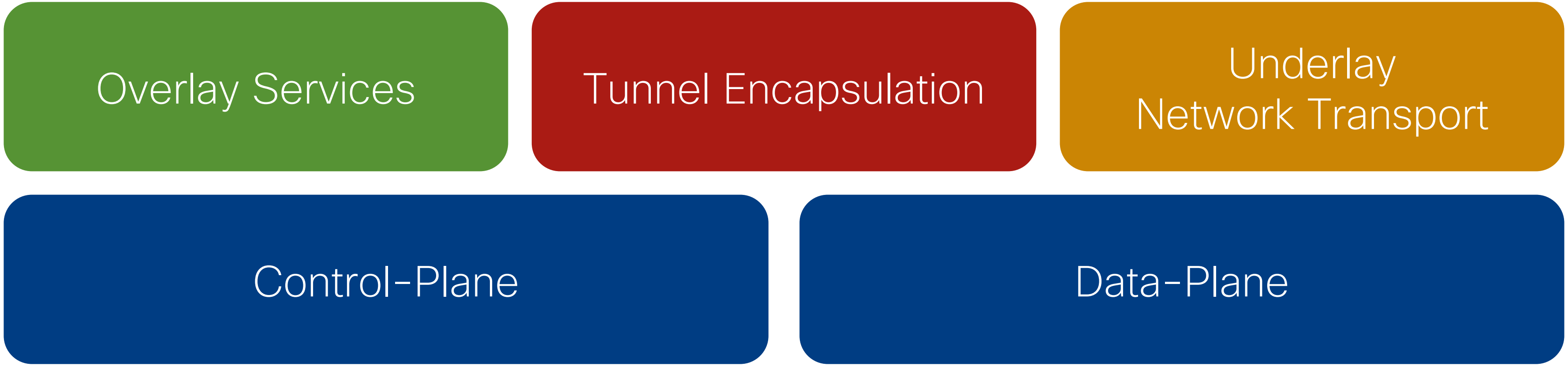
Network Virtualization Primer

Virtual Private Networks


Virtual Private Network (VPN)
Provider Provisioned VPN (PPVPNs)
Customer Provisioned VPNs
Site-to-Site
Remote Access
Remote Access
Site-to-Site
IPsec
GRE
IP-in-IP
Layer 1 VPNs (L1VPN)
Layer 2 VPNs (L2VPN)
Layer 3 VPNs (L3VPN)
GMPLS
Compulsory Tunnel Mode / NAS-Initiated
Voluntary Tunnel Mode / Client-Initiated
L2F
PPTP
L2TPv2/v3
Point-to-point (P2P) (VPWS)
Multipoint-to-Multipoint (M2M)
PPTP
L2TPv2/v3
IPsec
SSL / TLS
Virtual Private LAN Service (VPLS)
IP-only LAN Service (IPLS)
L2 Transport-Based (AToM)
802.1Q Tunneling (Q-in-Q)
PE-based
CE-based
L2TPv3 Based
L2 Transport-Based (AToM)
L2TPv3 Based
IPsec
GRE
IP-in-IP
BGP/MPLS
Virtual Router (VR) Based
TDP/LDP/RSVP Signalled LSPs
IPsec
L2TPv3
GRE
IP-in-IP

https://en.wikipedia.org/wiki/Virtual_private_network

Overlay Taxonomy



Overlay Services

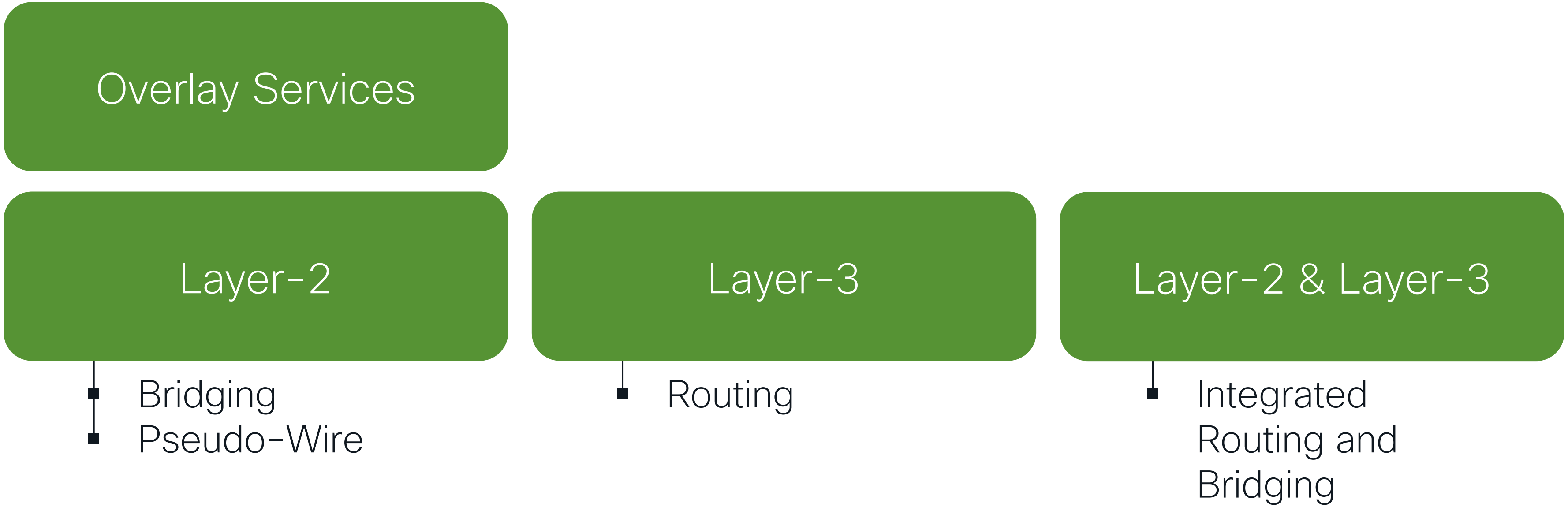
Tunnel Encapsulation

Underlay
Network Transport

Control-Plane

Data-Plane

Overlay Services



Overlay Services

Layer-2

- Bridging
- Pseudo-Wire

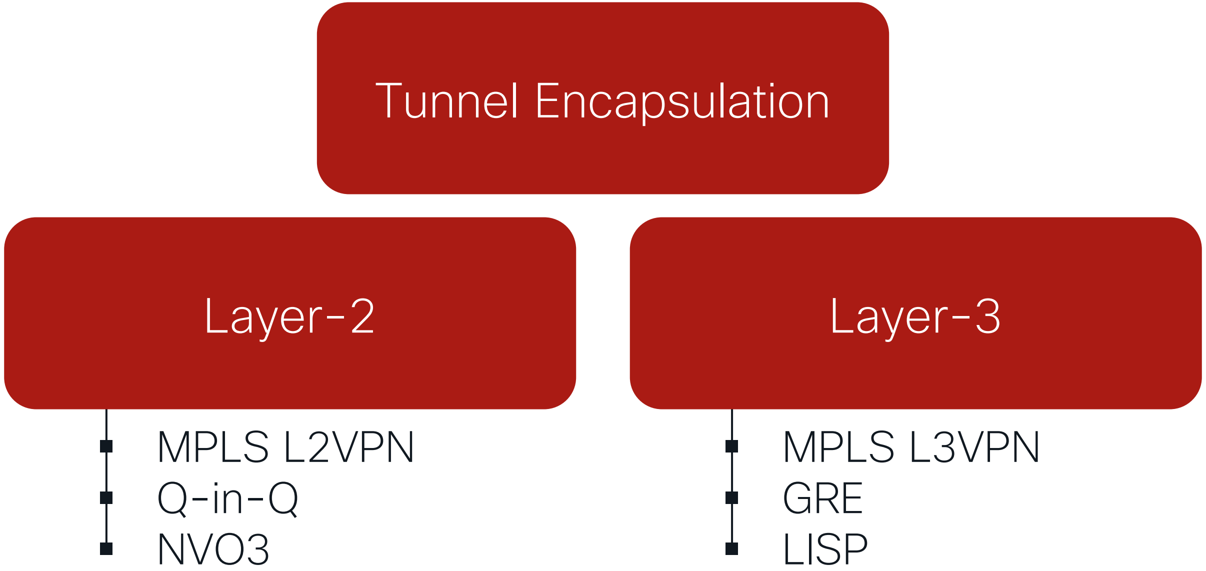
Layer-3

- Routing

Layer-2 & Layer-3

- Integrated Routing and Bridging

Tunnel Encapsulation



Tunnel Encapsulation

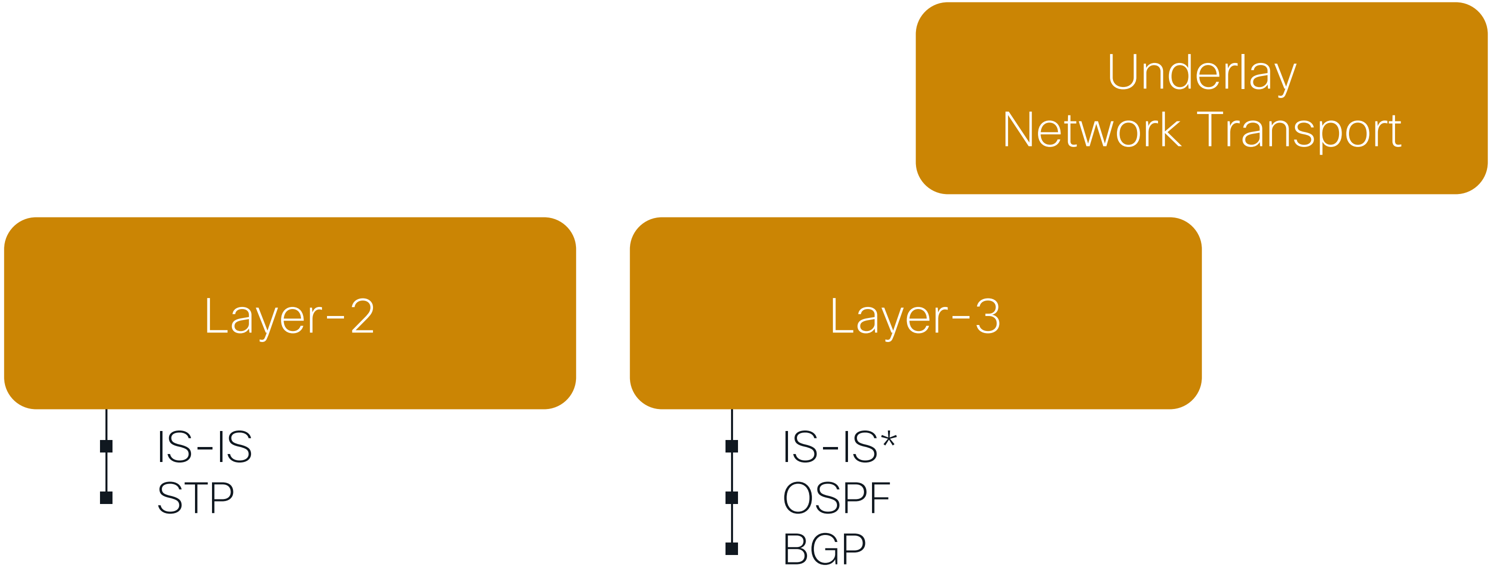
Layer-2

- MPLS L2VPN
- Q-in-Q
- NVO3

Layer-3

- MPLS L3VPN
- GRE
- LISP

Underlay Network Transport



Underlay
Network Transport

Layer-2

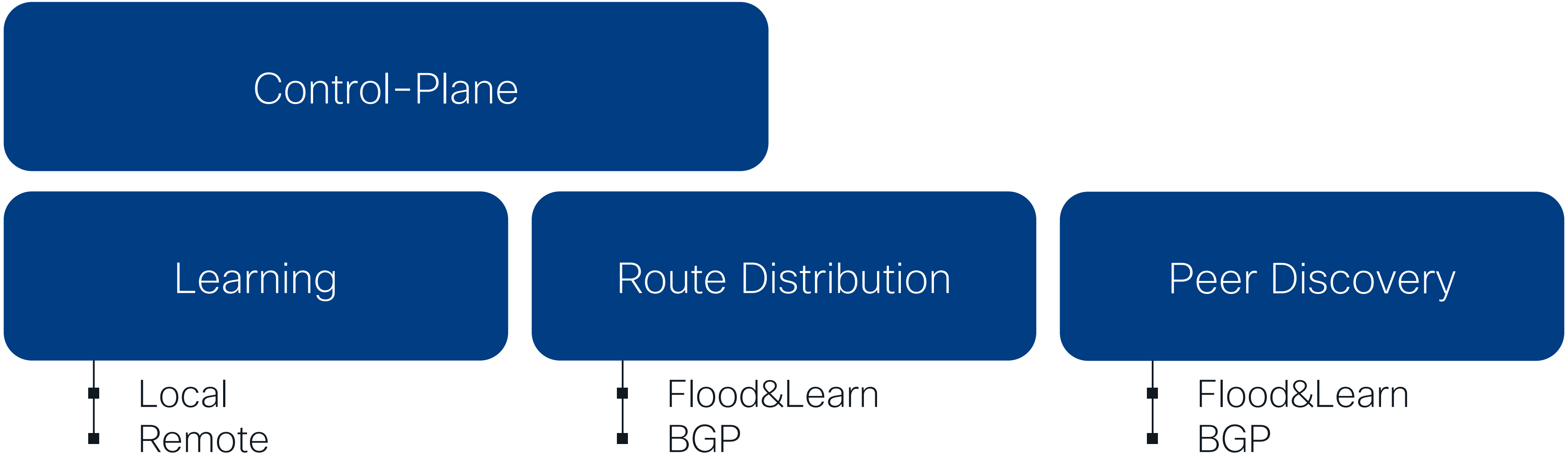
- IS-IS
- STP

Layer-3

- IS-IS*
- OSPF
- BGP

*IS-IS operates at L2 but transports L2 & L3

Control-Plane



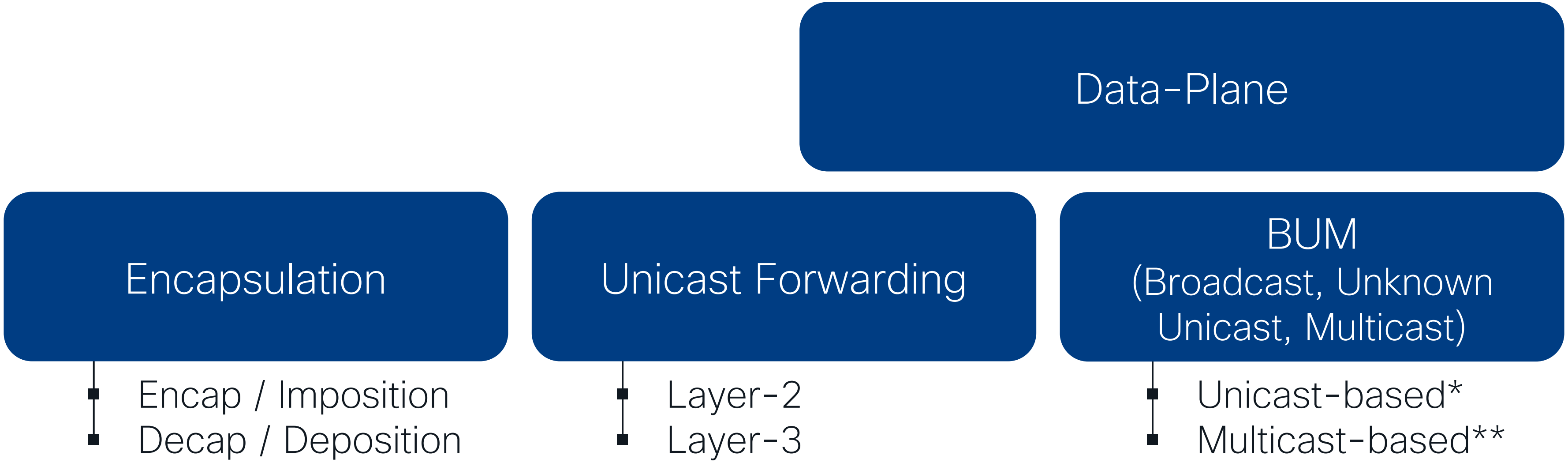
Control-Plane

Learning
- Local
- Remote

Route Distribution
- Flood&Learn
- BGP

Peer Discovery
- Flood&Learn
- BGP

Data-Plane



Data-Plane

Encapsulation
- Encap / Imposition
- Decap / Deposition

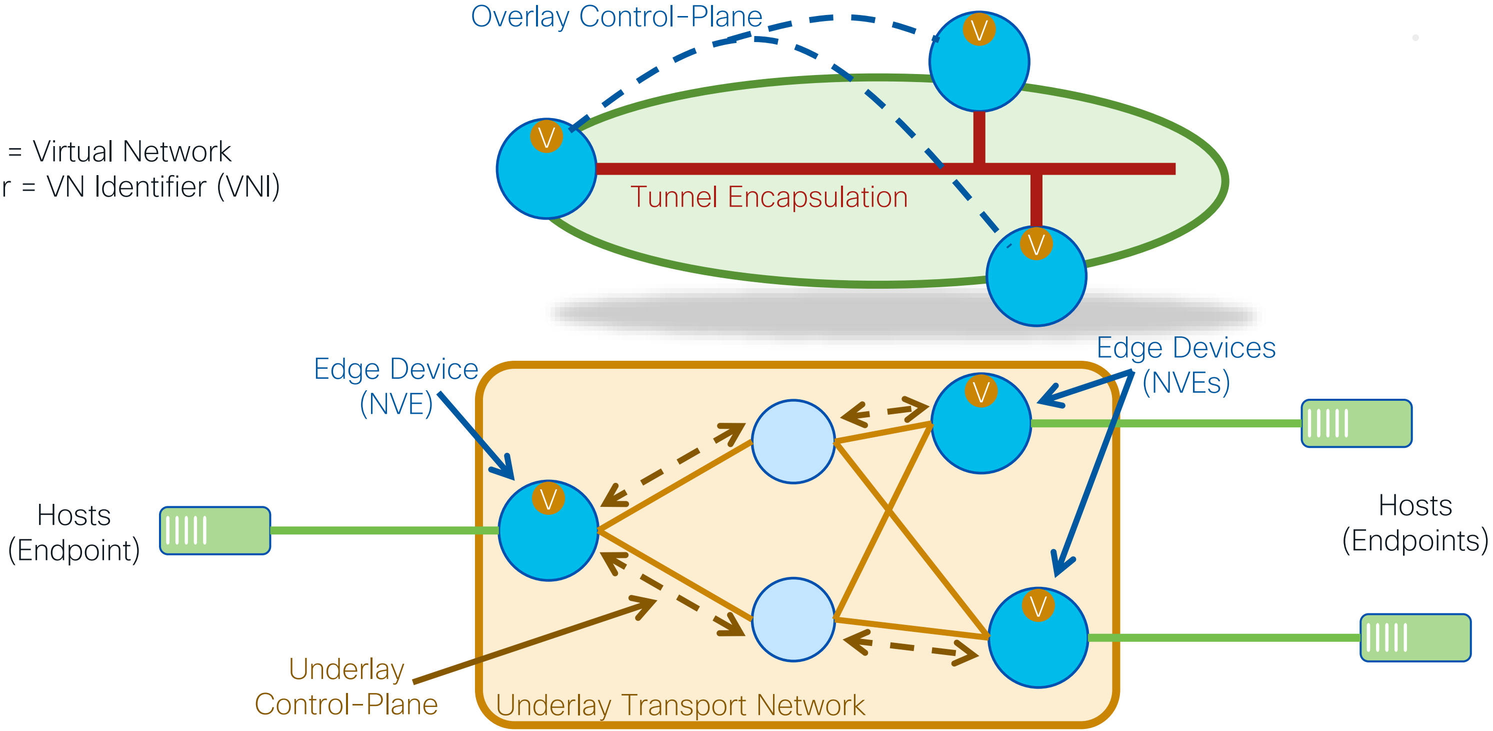
Unicast Forwarding
- Layer-2
- Layer-3

BUM (Broadcast, Unknown Unicast, Multicast)
- Unicast-based*
- Multicast-based**

*Ingress/Head-End Replication
**PIM

Overlay Taxonomy

Service = Virtual Network
Identifier = VN Identifier (VNI)



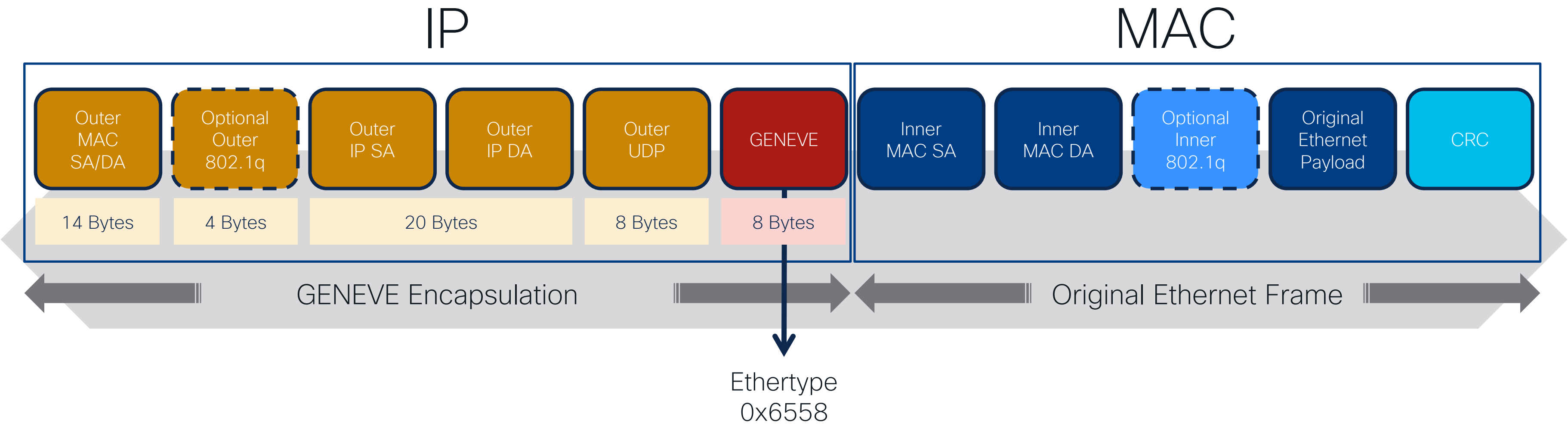
Encapsulation Overview

GENEVE – Generic Network Virtualization Encapsulation

What is GENEVE?

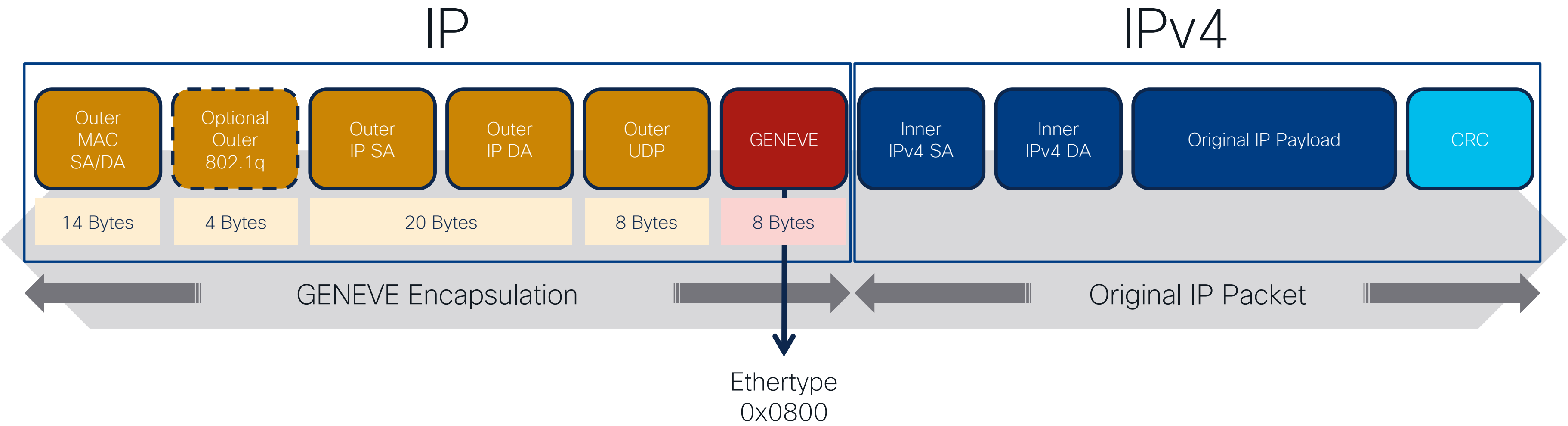
- Standards based Encapsulation
  - RFC 8926
  - MAC(and more)-in-IP
    - Dynamic Inner-Header
- Transport Independent
  - Layer-3 Transport (Underlay)
- Uses UDP-Encapsulation
  - Multipath Capable
  - Uses Per-Flow Entropy
- Flexible Namespace
  - Allows Segmentations

MAC(and more)-in-IP Encapsulation
GENEVE



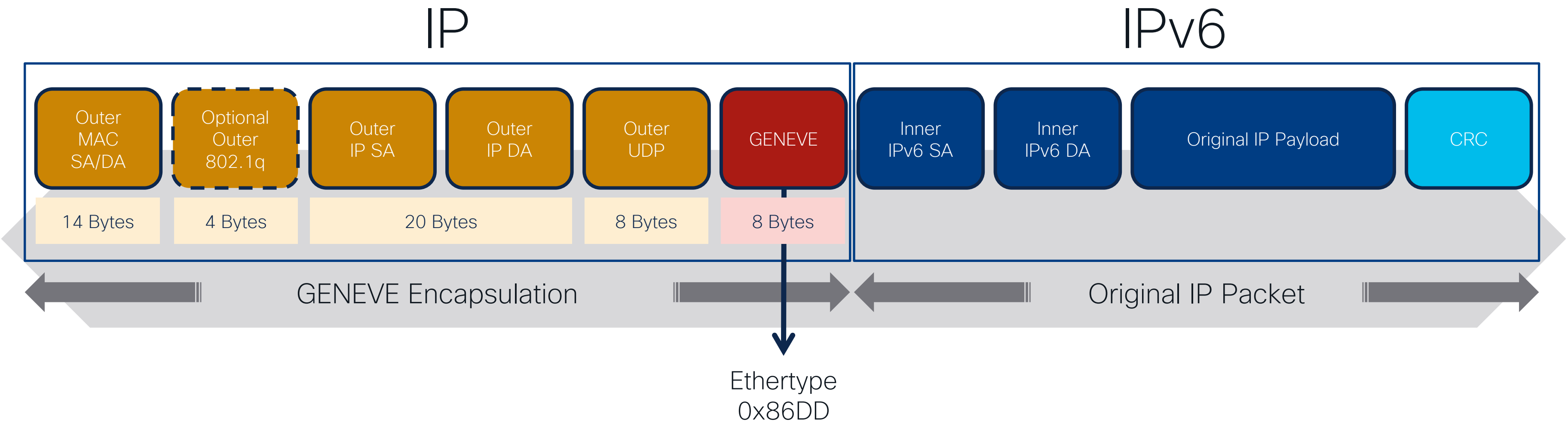
- Default Protocol Type in GENEVE is Ethernet (0x6558)
- IEEE 802 Numbers at IANA (*https://www.iana.org/assignments/ieee-802-numbers/ieee-802-numbers.xhtml)

MAC(and more)-in-IP Encapsulation
GENEVE



- Optional Protocol Type for IPv4 (0x0800)
- IEEE 802 Numbers at IANA (*https://www.iana.org/assignments/ieee-802-numbers/ieee-802-numbers.xhtml)

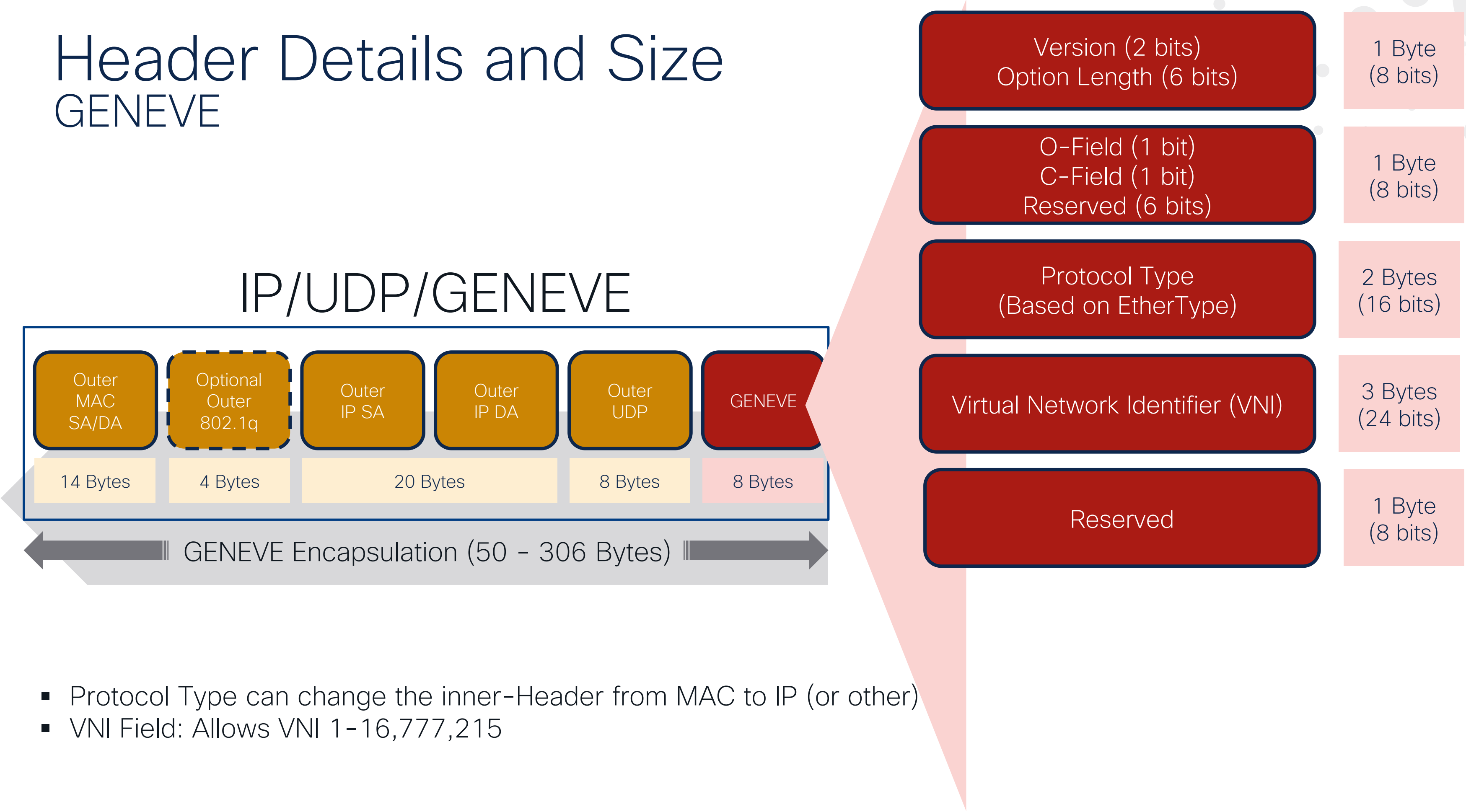
MAC(and more)-in-IP Encapsulation
GENEVE



- Optional Protocol Type for IPv6 (0x86DD)
- IEEE 802 Numbers at IANA (*https://www.iana.org/assignments/ieee-802-numbers/ieee-802-numbers.xhtml)

Header Details and Size

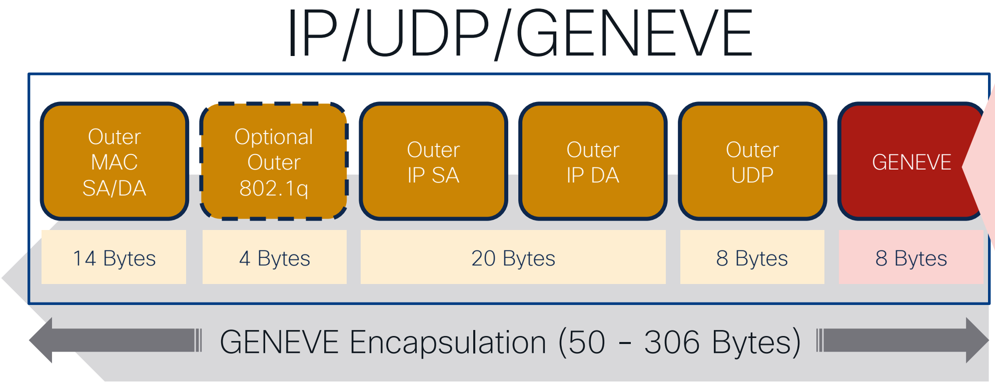
GENEVE



- Protocol Type can change the inner-Header from MAC to IP (or other)
- VNI Field: Allows VNI 1-16,777,215

Header Details and Size

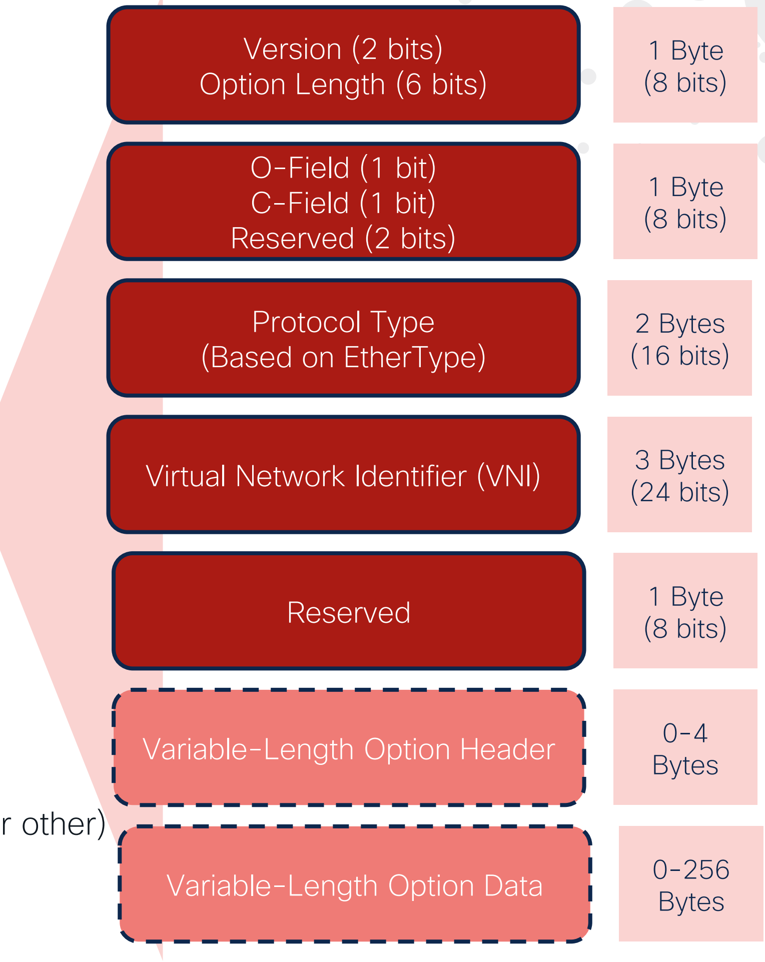
GENEVE



IP/UDP/GENEVE

Outer MAC SA/DA	Optional Outer 802.1q	Outer IP SA	Outer IP DA	Outer UDP	GENEVE
14 Bytes	4 Bytes	20 Bytes		8 Bytes	8 Bytes

GENEVE Encapsulation (50 - 306 Bytes)



Version (2 bits) Option Length (6 bits)	1 Byte (8 bits)
O-Field (1 bit) C-Field (1 bit) Reserved (2 bits)	1 Byte (8 bits)
Protocol Type (Based on EtherType)	2 Bytes (16 bits)
Virtual Network Identifier (VNI)	3 Bytes (24 bits)
Reserved	1 Byte (8 bits)
Variable-Length Option Header	0-4 Bytes
Variable-Length Option Data	0-256 Bytes

- Protocol Type can change the inner-Header from MAC to IP (or other)
- VNI Field: Allows VNI 1-16,777,215
- Variable-Length Option can be Zero

Header Extension
GENEVE

- Extensible Headers for Adding Use-Cases
  - Use-Cases for Variable Length Options
  - Some Existing Proposals
    - GBP– Group Based Policy
      - https://datatracker.ietf.org/doc/html/draft-lemon-geneve-gbp
    - INT – In-Band Network Telemetry
      - https://datatracker.ietf.org/doc/html/draft-brockners-ippm-ioam-geneve

Header Extension
GENEVE

- Extensible Headers for Adding Use-Cases
  - Variable Length Options Ranges
  - Details of Options Ranges
    - Options and Vendor options registered with IANA
    - Total of 65k of possible Option Registration (First Come, First Serve)
    - https://www.iana.org/assignments/nvo3/nvo3.xhtml#geneve-option-class

Registration Procedure	Range
IETF Review	0x0000-0x00FF
First Come First Served	0x0100-0xFEFF
Experimental Use	0xFF00-0xFFFF

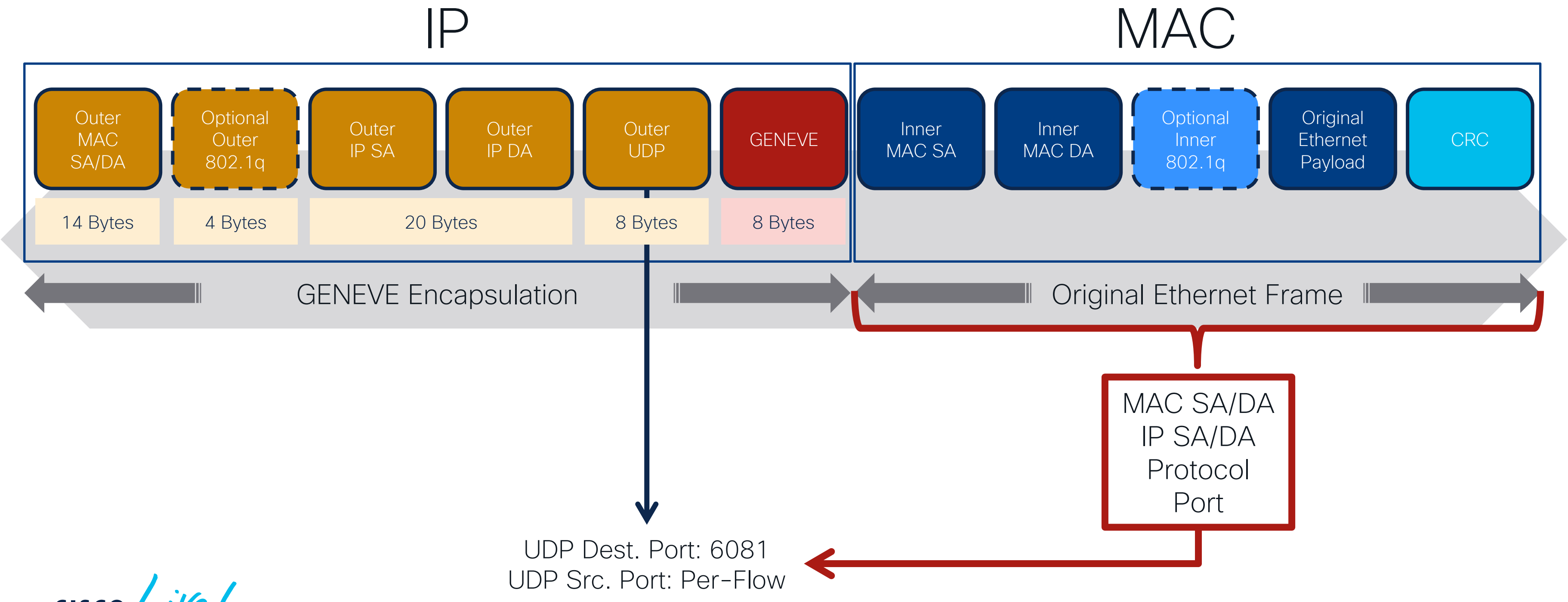
*https://www.iana.org/assignments/nvo3/nvo3.xhtml#geneve-option-class

Header Extension
GENEVE

- Extensible Headers for Adding Use-Cases
  - Details of Vendor Options are PARTIALLY documented

Description	Option Class	Description	Option Class
Linux	0x0100	Ericsson	0x0119-0x0128
Open vSwitch (OVS)	0x0101	Oxide Computer Company	0x0129
Open Virtual Networking (OVN)	0x0102	Google	0x0132-0x0135
In-Band Network Telemetry (INT)	0x0103	InfoQuick Global Connection Tech Ltd.	0x0136
VMware, Inc.	0x0104	Alibaba, inc	0x0137-0x0140
Amazon.com, Inc.	0x0105, 0x0108-0x0110	Palo Alto Networks	0x0141-0x0144
Cisco Systems, Inc.	0x0106, 0x0130-0x0131	Huawei Technologies Co., Ltd	0x0145-0x0149
Oracle Corporation	0x0107	EMnify GmbH	0x014A
IBM	0x0111-0x0118	Currently Unassigned (01/2023)	0x014B-0xFEFF

MAC(and more)-in-IP Encapsulation
GENEVE



Control-Plane
GENEVE

Flood&Learn (RFC8926)

- Ethernet over IP
  - No Spanning-Tree (terminates at NVE)
  - Endpoint Learning is based on Flood and Learn (it's in the name)
  - Requires Extra Work for Routing
    - FHRP for Default Gateway
    - Over-the-Top VRF-lite for Prefix Routing (or use the Underlay?!)

EVPN - Ethernet VPN (draft-ietf-bess-evpn-geneve)

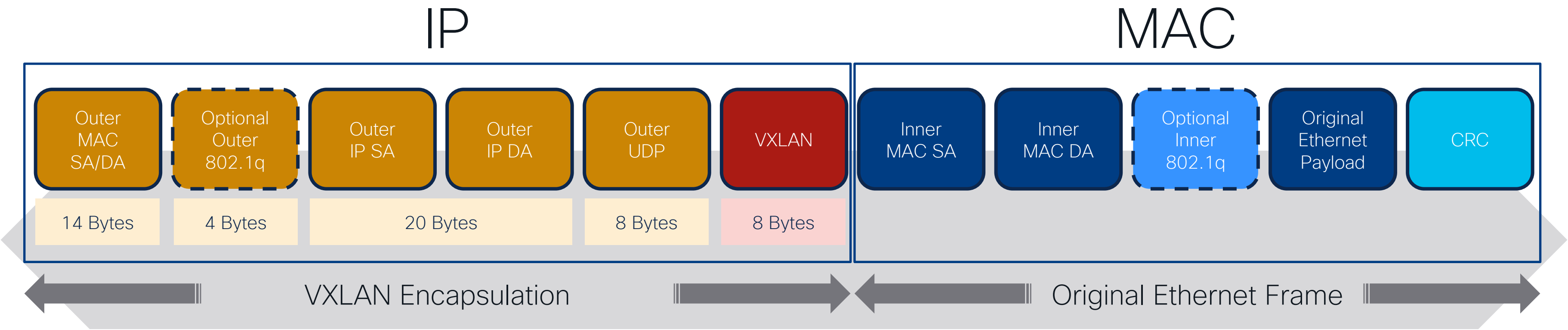
- A Better Ethernet/IP over IP
  - No Spanning-Tree (terminates at NVE)
  - Endpoint Learning is based on BGP exchange (EVPN uses BGP)
  - Provides Integrated Routing & Bridging (IRB)
    - Distributed Anycast Gateway for Default Gateway
    - Uses a Layer-3 VPN approach like MPLS L3VPN
  - And there is much more in EVPN!

VXLAN - Virtual Extensible Local Area Network

What is VXLAN?

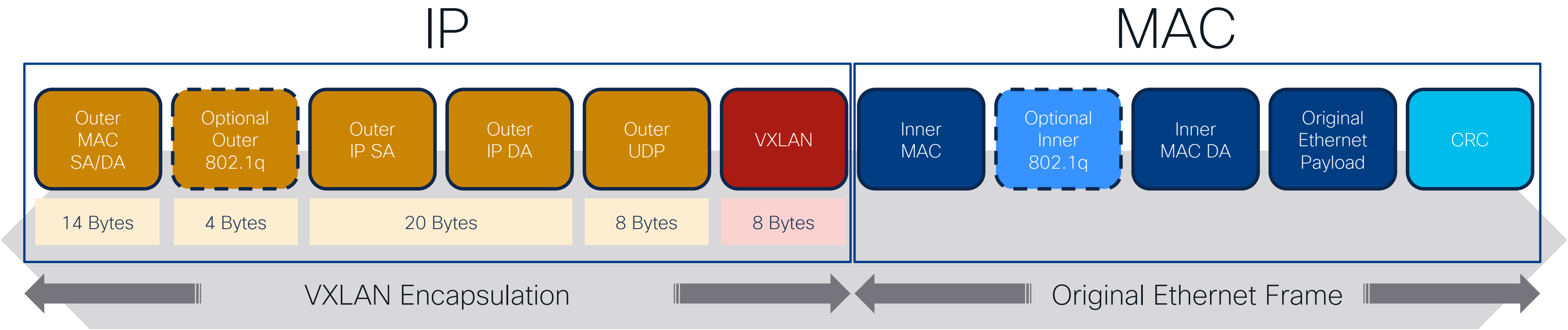
- Standards based Encapsulation
  - RFC 7348
  - MAC-in-IP
    - MAC as inner-Header
- Transport Independent
  - Layer-3 Transport (Underlay)
- Uses UDP-Encapsulation
  - Multipath Capable
  - Uses Per-Flow Entropy
- Flexible Namespace
  - Allows Segmentations

MAC-in-IP Encapsulation
VXLAN



- VXLAN Always has an Inner-MAC Header

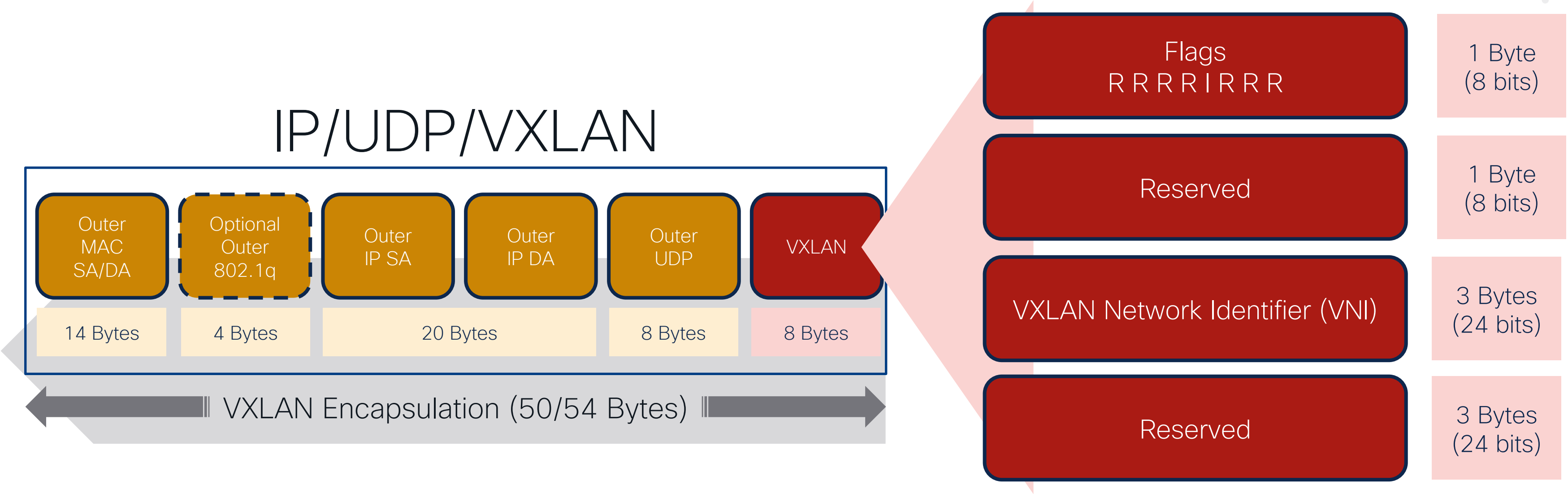
MAC-in-IP Encapsulation
VXLAN



- VXLAN Always has an Inner-MAC Header
- In case of IPv4 or IPv6, Router MAC information are required

Header Details and Size

VXLAN



- Flags Field: I-flag (set to 1) for valid VNI. Other flags remain as R (set to 0)
- VNI Field: Allows VNI 1-16,777,215 (some implementation only 4096-16,777,215)

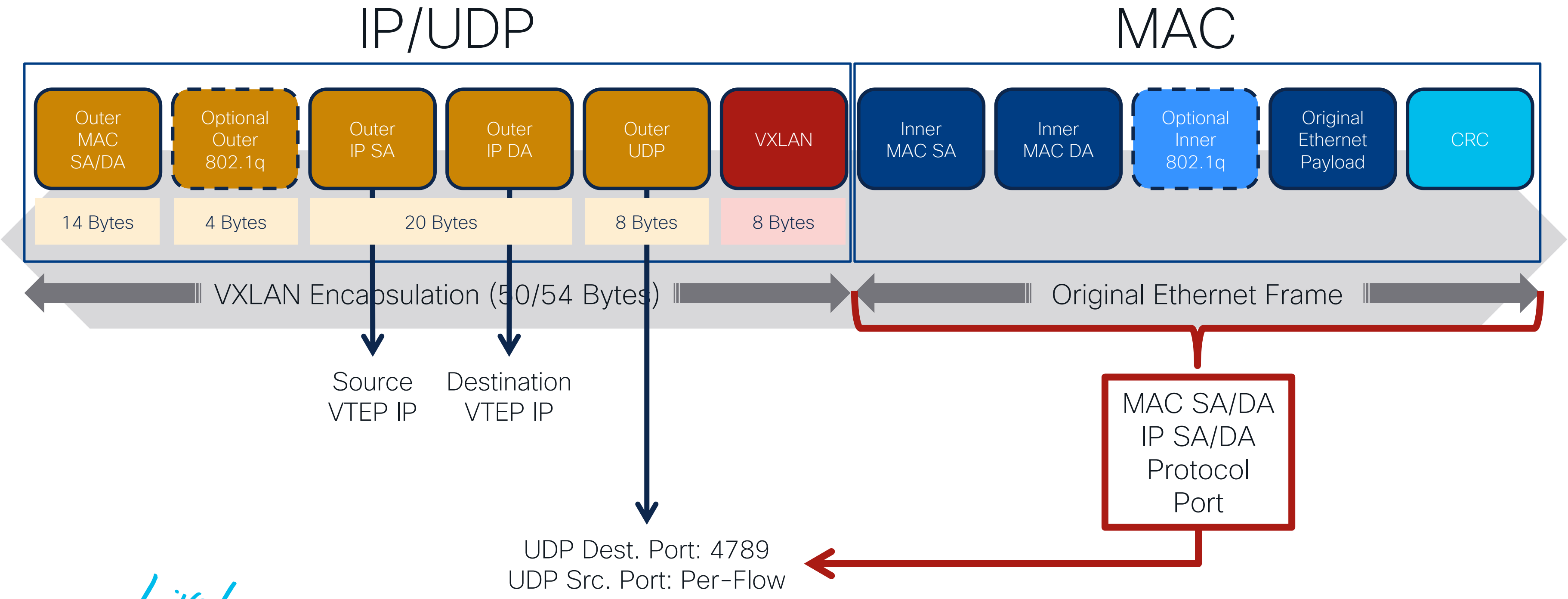
Header Extension
VXLAN

- Extensible Headers don't exist
  - Proposal for Existing Flag and Reserved Field Usage
  - Some Existing Proposals
    - VXLAN Group Policy Option
      - https://datatracker.ietf.org/doc/html/draft-smith-vxlan-group-policy
    - BUM Bit (borrowed from VXLAN-GPE)
      - https://datatracker.ietf.org/doc/html/draft-ietf-nvo3-vxlan-gpe

NOTE: VXLAN and VXLAN-GPE are not interoperable. Still, some VXLAN-GPE concepts could be leveraged by VXLAN (non-Standard)

Multipath Capable

VXLAN



Control-Plane
VXLAN

Flood&Learn (RFC7348)

- Ethernet over IP
  - No Spanning-Tree (terminates at NVE)
  - Endpoint Learning is based on Flood and Learn (it’s in the name)
  - Requires Extra Work for Routing
    - FHRP for Default Gateway
    - Over-the-Top VRF-lite for Prefix Routing (or use the Underlay?!)

EVPN - Ethernet VPN (RFC8365)

- A Better Ethernet over IP
  - No Spanning-Tree (terminates at NVE)
  - Endpoint Learning is based on BGP exchange (EVPN uses BGP)
  - Provides Integrated Routing & Bridging (IRB)
    - Distributed Anycast Gateway for Default Gateway
    - Uses a Layer-3 VPN approach like MPLS L3VPN
  - And there is much more in EVPN!

Use-Cases

Routing and Bridging

GENEVE

- Using FHRP for Flood&Learn
- Integrated Routing & Bridging (IRB) via EVPN
- Missing Usage of Common Control-Plane (F&L, EVPN) or Proprietary Control-Plane Implementation (e.g. NSX)

VXLAN

- Using FHRP for Flood&Learn
- Integrated Routing & Bridging (IRB) via EVPN
- Known and tested interoperability with Flood&Learn and EVPN Control-Plane (Plugfest)

Security and Micro-Segmentation

GENEVE

- Uses Protocol Type and Variable Length Option for Group Based Policy
    - https://datatracker.ietf.org/doc/html/draft-lemon-geneve-gbp
- Implementation seen from vendors; usage of variable length field not documented (e.g. NSX)
- Missing Common implementation of GBP

VXLAN

- Uses Reserved Field for Group Policy Option
    - https://datatracker.ietf.org/doc/html/draft-smith-vxlan-group-policy
- Implementation in ACI, SDA and other vendors seen. Flags and reserved field usage documented
- Common implementation of GPO known but no interoperability tested

Integration of Host and Network Overlay

GENEVE

- Common Data-Plane
- Control-Plane
  - Often Proprietary, Not Documented
- Missing Common implementation
- Interoperability Not Possible with different Control-Plane (e.g. NSX)

VXLAN

- Common Data-Plane
- Control-Plane
  - Flood&Learn or EVPN
- Implementation between Vendors seen
- Interoperability tested with common Control-Plane

‘Use-Case and Interop of Extension Headers is an Open Question’

These days, there is limited or no common implementation for GENEVE amongst vendors

Score Board

Score Board

	GENEVE	Score	VXLAN	Score
IETF	RFC 8926 (Standard)	++	RFC 7348 (Informational)	+
Encapsulation	MAC(and more)-in-IP	++	MAC-in-IP	+
Outer-Header	Fixed, IPv4 or IPv6	++	Fixed, IPv4 or IPv6	++
Entropy	UDP Source Port	++	UDP Source Port	++
Inner-Header	Dynamic (Protocol Type)	++	Fixed (MAC)	+
VNI	24 bits (~16 Million Segments)	++	24 bits (~16 Million Segments)	++
Control-Plane	draft-ietf-bess-evpn-geneve (Pre-Standard)	-	RFC8365 (Standard)	++
Extensibility	Option Class	++	Using Reserved Fields	-
Operations	Integrated	++	External	-
Availability	All Use-Cases can't be implemented, cost-effectively, in Hardware (Number of Gates)	-	Extensive number of Use-Cases are widely available (in Custom & Merchant Silicon)	++
Header-Size	Too large for commonly available parse buffer (50-306 Bytes)*	-	Fits in available Hardware parser across Router/Switch and NIC (50/54 Bytes)	++

*256 bytes in NIC (draft-ietf-nvo3-encap)

Score Board Result

Who gets 22 Points?

GENEVE 13 Points (16-3)	VXLAN 13 Points (15-2)
<u>16 Plus Points</u> Modern Data-Plane Design Extension Header Extensibility of Encapsulation to Use-Cases	<u>15 Plus Points</u> Adoption Implementation Control-Plane Status Ability for Use-Case Execution
<u>3 Minus Points</u> Adoption Implementation Control-Plane Status	<u>2 Minus Point</u> No Extension Header Native Operations Integration

Conclusion

Conclusion

- Neither GENEVE nor VXLAN define a Control-Plane beyond Flood&Learn
- GENEVE and VXLAN both have BGP-based Control-Plane Proposals (EVPN)
  - Same Control-Plane results in Same Use-Cases
- GENEVE uses dynamic Protocol Type while VXLAN has fixed Protocol
    - Protocol Type and Variable Length Options – Flexible But Requires Documentation for interop
- Cisco Hardware Supports VXLAN and GENEVE Encapsulation
  - VXLAN with and without EVPN is widely implemented
  - GENEVE capability is validated on Cisco Nexus 9000 with CloudScale ASIC (not productized)
- Common Implementations are Limited or Absent (Today)
- Score Board Results – What Really Matters?
  - New Bits in the Header or Wide Adoption and Implementation!
  - Why Changing the Data-Plane if it doesn't give you anything New that is usable?

Complete your Session Survey

- Please complete your session survey after each session. Your feedback is important.
- Complete a minimum of 4 session surveys and the Overall Conference survey (open from Thursday) to receive your Cisco Live t-shirt.
- All surveys can be taken in the Cisco Events Mobile App or by logging in to the Session Catalog and clicking the "Attendee Dashboard" at https://www.ciscolive.com/emea/learn/sessions/session-catalog.html

Continue Your Education

- Visit the Cisco Showcase for related demos.
- Book your one-on-one Meet the Engineer meeting.
- Attend any of the related sessions at the DevNet, Capture the Flag, and Walk-in Labs zones.
- Visit the On-Demand Library for more sessions at ciscolive.com/on-demand.

The bridge to possible

Thank you

cisco Live!