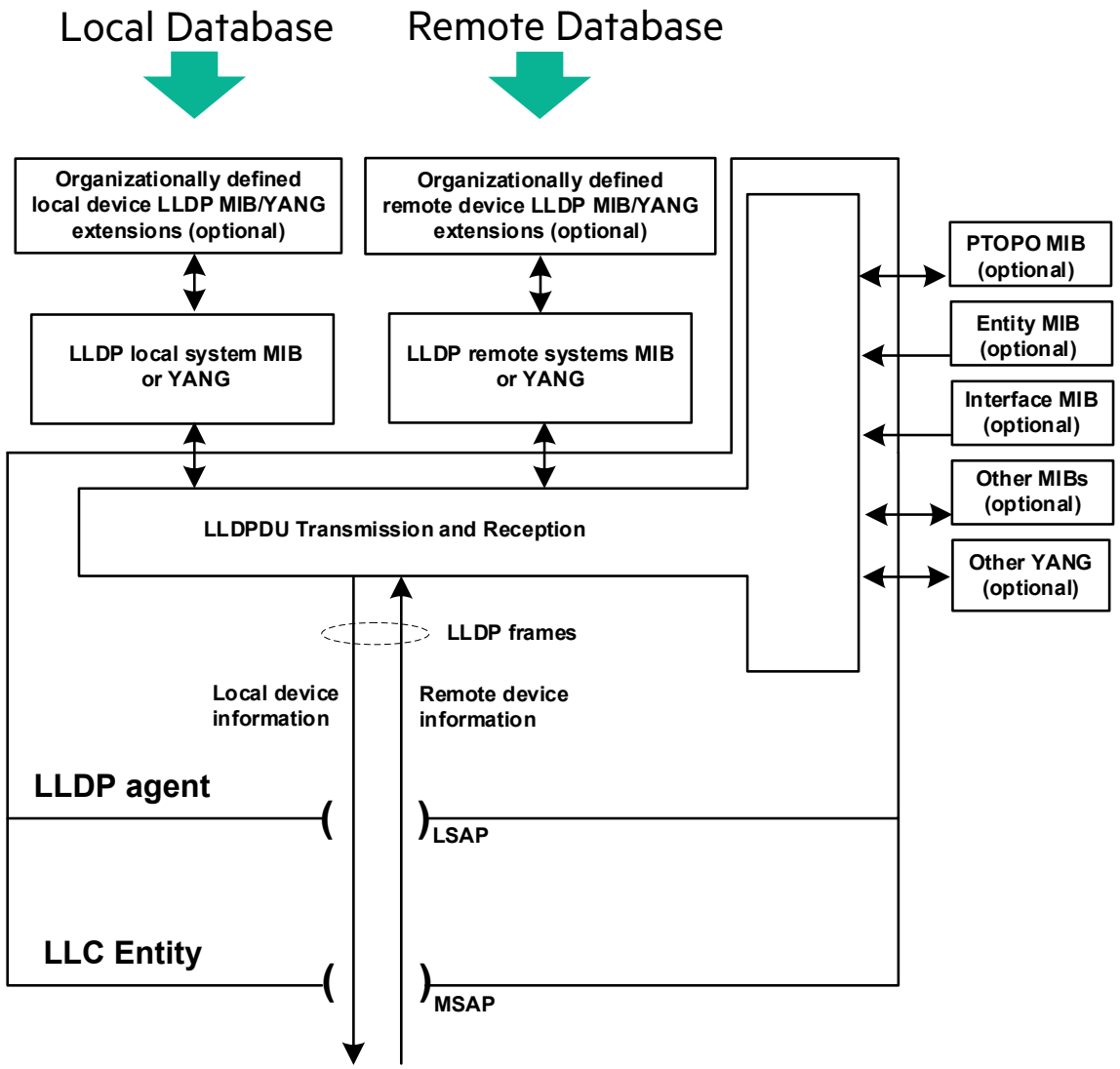
Hewlett Packard
Enterprise
HPE Aruba Networking

LLDP Between LSVR Routers

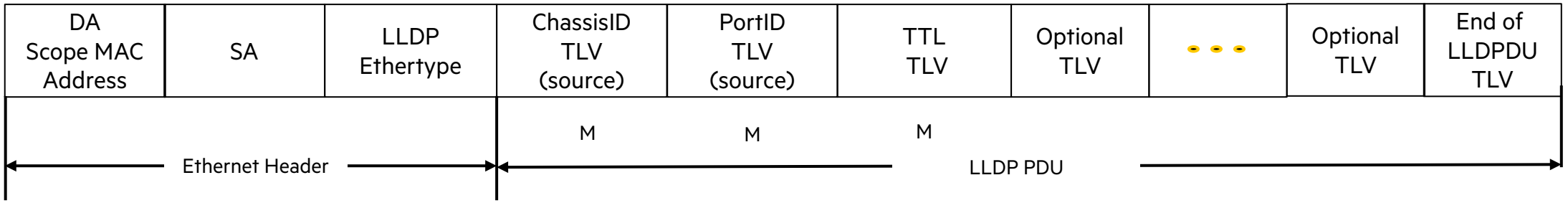
Paul Bottorff
04/22/2025

LLDP is an extremely simple database exchange protocol



- LLDP starts with a local database which it advertises in a single frame LLDPDU using a Scope MAC Address as the L2 destination.
- The Scope MAC Address determines the extent of the LLDPDU propagation.
- LLDP advertisements are received and stored at each LLDP port listening to the corresponding Scope MAC Address.
- The LLDP receivers build a remote database containing the advertised information from each LLDP transmitter.
- LLDP transmitters are identified by a Chassis ID and Port ID encoded in each LLDPDU
- LLDP transmitters advertise their local database periodically to maintain liveliness

Classic LLDPDU single frame format (IEEE Std 802.1AB-2016)



LLDP TLV Format

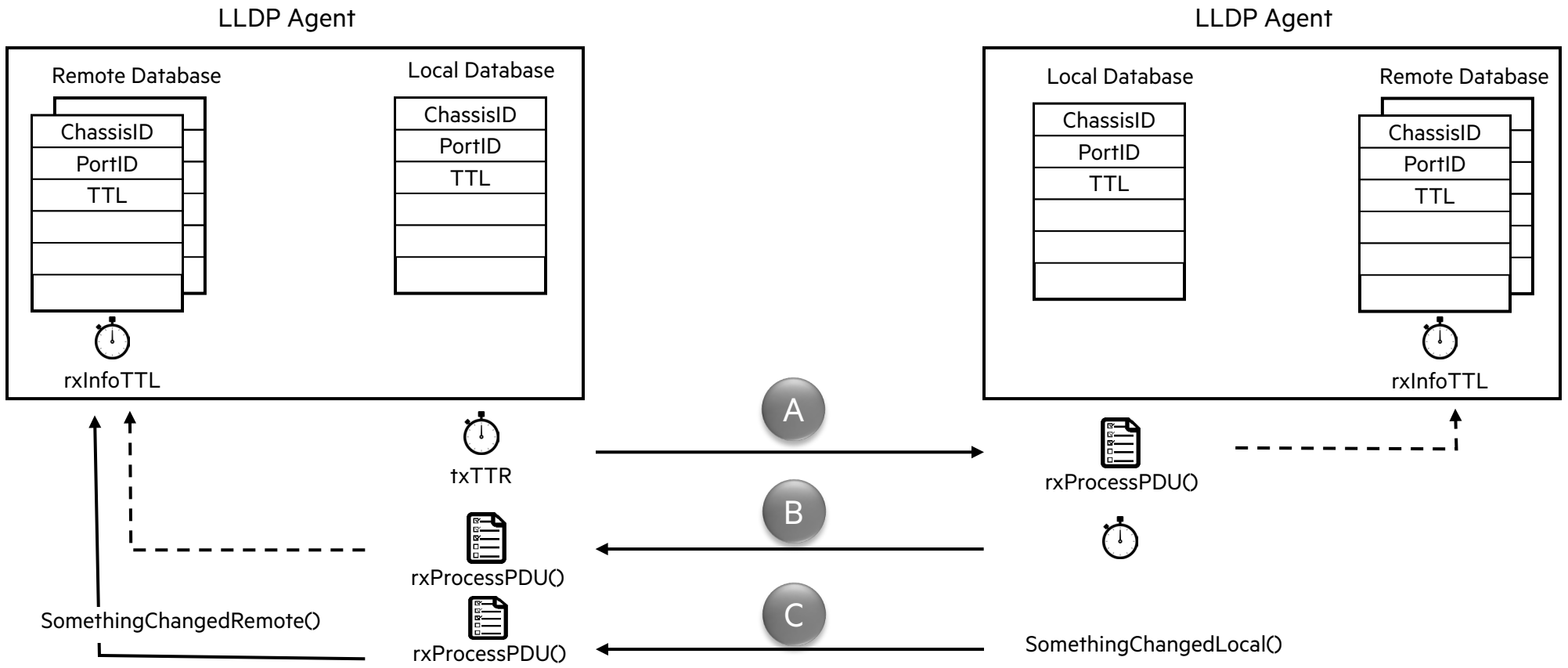
TLV type (7 bits)	TLV information String length (9 bits)	TLV information

LLDP Organizationally Specific TLV Format

TLV type = 127 (7 bits)	TLV information String length (9 bits)	OUI/CID (3 octets)	Subtype (1 octet)	TLV information

- The Scope MAC Address of the single frame Normal LLDPDU determines the limits of propagation of LLDPDUs. It can be the:
  - Nearest bridge group address, nearest customer bridge group address, nearest non-TPMR bridge group address
  - Also, the standard allows any group MAC address or any individual MAC address
- Each LLDPDU includes three mandatory TLVs which are the ChassisID, PortID, and Time To Live TLV
  - ChassisID provides a unique system identifier, i.e. MAC Address, IP Address, other management identifier
  - PortID provides a unique identifier for the interface (or sub-interface) within the system, i.e. IfIndex, Port Number
  - Time To Live provide duration the LLDPDU is valid
- Optional TLVs contain further information
  - An 802.1ABdh Manifest TLV can be used to extend the database size up adding up to 83 additional frames of TLVs

Classic LLDPDU single frame operation (IEEE Std 802.1AB-2016)

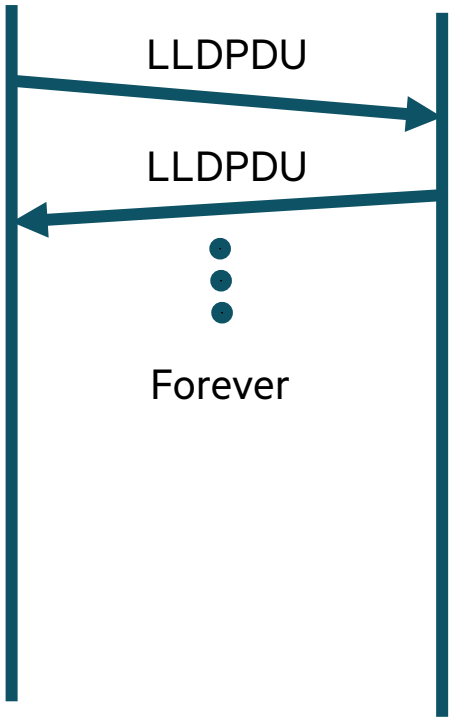


A B Each port periodically advertises its local LLDP database

A B On receipt of an LLDP advertisement, LLDP updates the remote database for the sender

C When something changes in the local database LLDP immediately transmits and update

Classic LLDP (single frame) ladder: (802.1AB-2016)



- Each LLDPDUs a single frame containing the TLVs which encode the LLDP information from the sender
- Each LLDPDU is transmitted to a destination MAC address called the Scope MAC Address
- Every LLDPDU repeats all the sender's current information (in TLVs)
- LLDPDU transmissions repeat at an average rate (default 30 seconds)
- Whenever a new device enters the network or when any change occurs in the content of the LLDPDU a series of fast LLDPDU transmissions occurs (default 4 LLDPDU spaced every second)

LLDP ChassisID + PortID = L3DL Logical Link Endpoint Identifier (LLEI)

Chassis ID TLV Format

TLV type = 1 (7 bits)	TLV information String length (9 bits)	Chassis ID subtype (8 bits)	Chassis ID (1-255 octets)

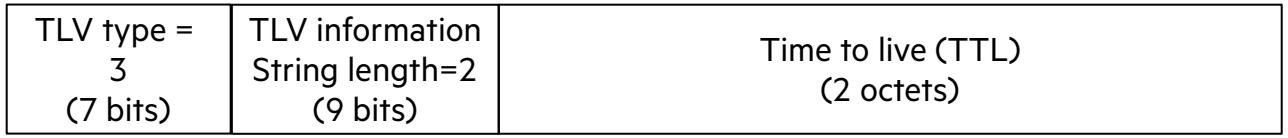
Port ID TLV Format

TLV type = 2 (7 bits)	TLV information String length (9 bits)	Port ID subtype (8 bits)	Port ID (1-255 octets)

- The Chassis ID and Port ID are the first two TLVs in all LLDPDUs (frames). They identify the data base element this LLDPDU is associated with.
- The ChassisID provides a unique system identifier which can be a MAC address, any IANA network address (such as an NSAP, IP, etc), a management identifier, or any locally defined string identifier.
- The PortID provides a unique identifier for the interface within the system which can be an IfIndex, a port number, or any locally defined string identifier.
- Together they can uniquely identify each router interface

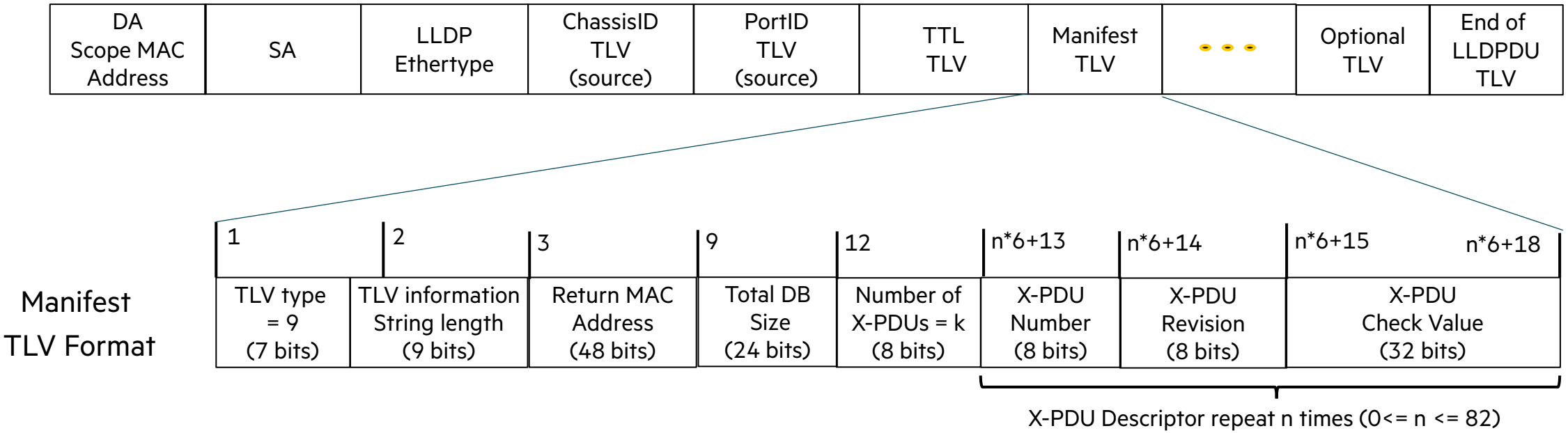
LLDP uses periodic LLDPDU transmissions to assure liveliness

Time To Live TLV format



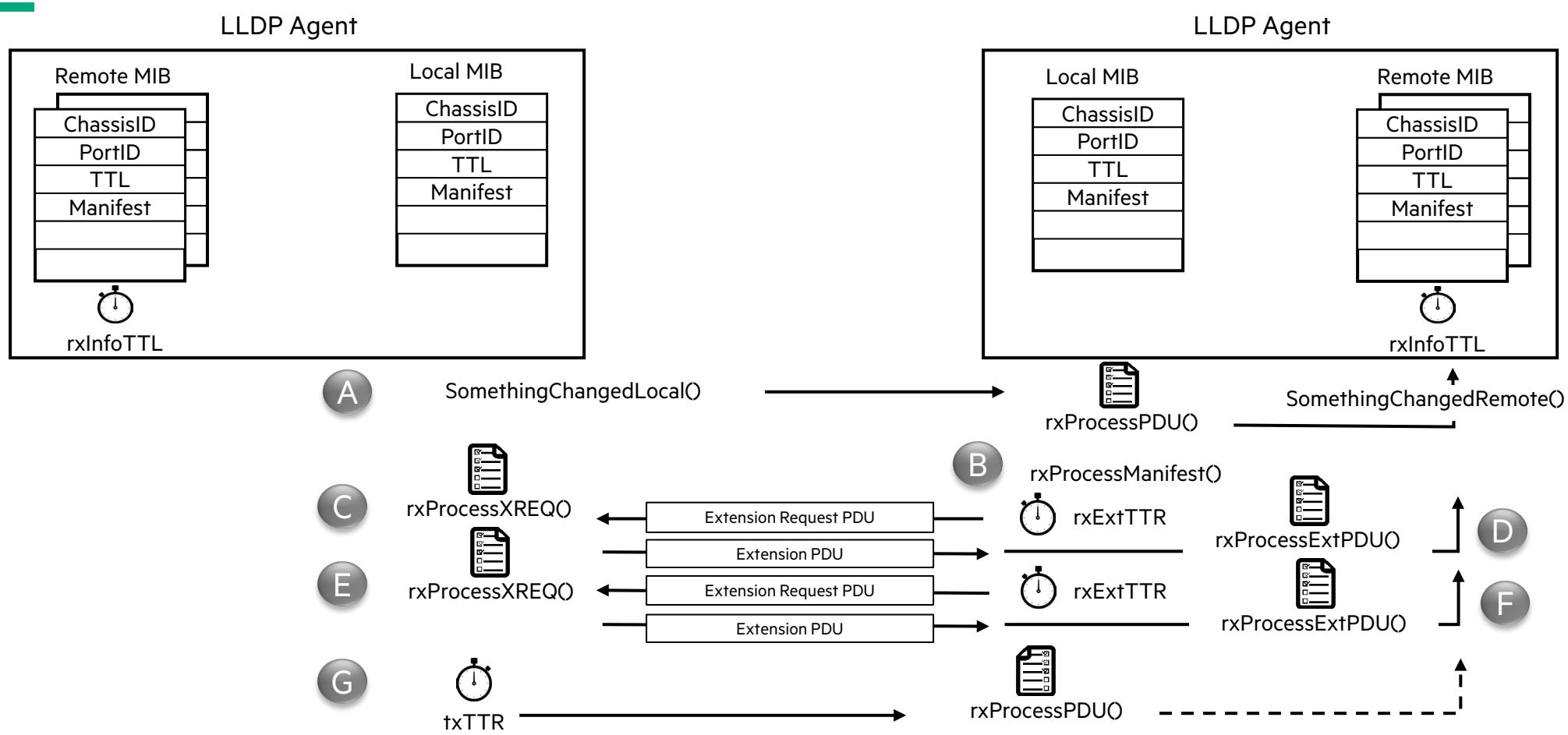
- Periodic transmissions can be set at 1-3600 seconds intervals.
  - Default is 30 seconds.
- The LLDP data base is invalid and discarded if a periodic transmission is not received within the time to live interval.
  - Default is 4 transmission times.
- A 0 time to live invalidates the database immediately.

A Manifest TLV is used to extend the LLDP database (IEEE Std 802.1ABdh)



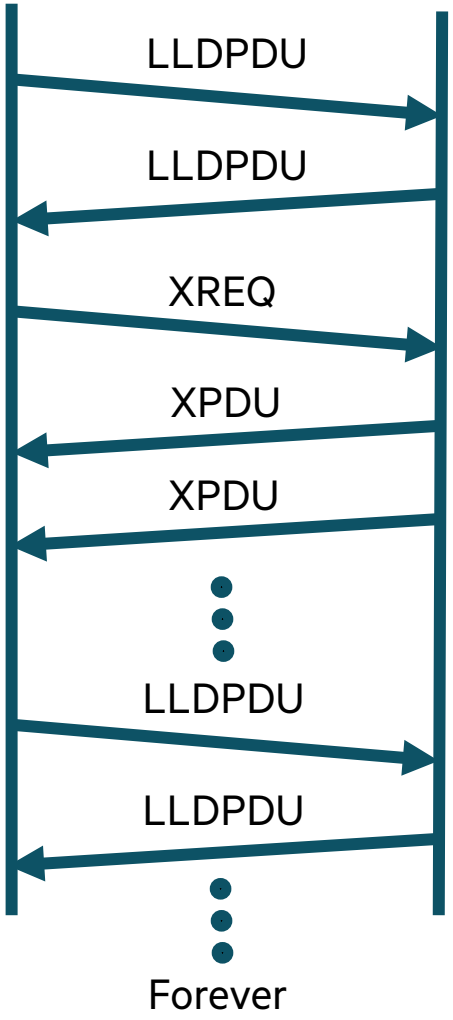
- Here the LLDPDU contains a Manifest TLV indicating an extended LLDP database
- The Manifest TLV specifies up to 83 extension XPDUs each containing LLDP TLVs
- Each of the extension XPDUs are requested using a unicast extension request XREQ
- Since the manifest contains a digest of each extension TLV the receiver can decide which if any extension TLV have changed. If none have changed then none need to be requested.
- Since the receiver controls the rate of XPDU transmissions it can manage congestion at input buffers.

Extended LLDP Operation (IEEE 802.1ABdh-2021)



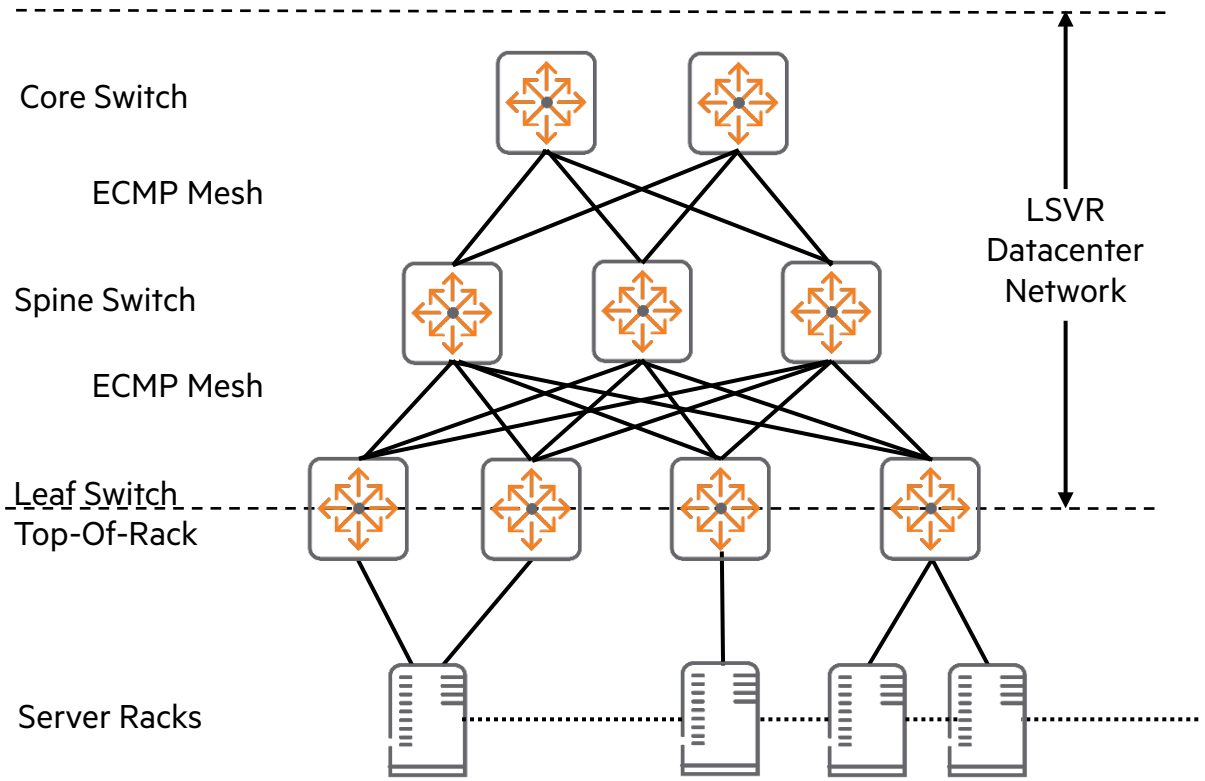
- Send LLDPDU as specified when periodically and whenever something changes
- Only send extension LLDPDU when explicitly requested by a XREQ
- Only issue XREQ when manifest shows the local copy is out of date

Extended LLDP (multi-frame) ladder: (802.1ABdh-2021)



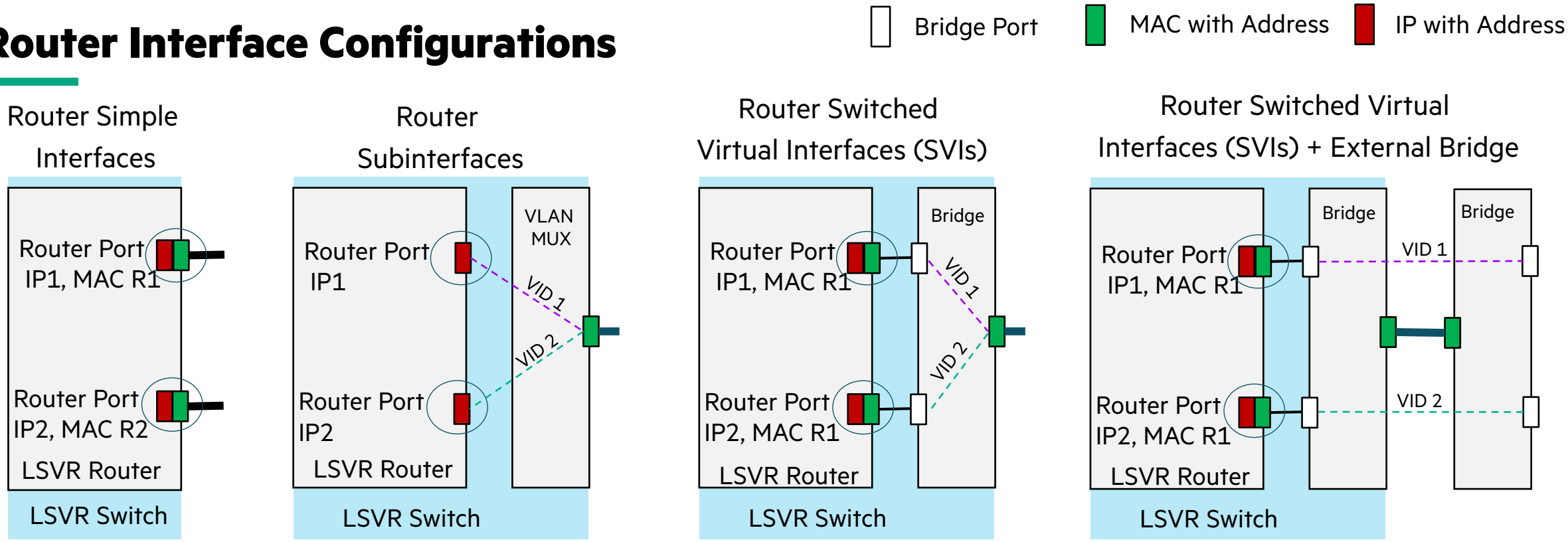
- Each Normal LLDPDU is a single frame containing the TLVs of the first frame of LLDP information from the sender
- LLDP is extended by adding a Manifest TLV to the Normal LLDPDU which describes additional extension XPDUs
- Each XPDU contains further source LLDP information encoded in LLDP format TLVs.
- Normal LLDPDU for extended LLDP are backward compatible with Classic LLDP (802.1AB-2016) and follow the same transmission rules.
- The receiver of a Normal LLDPDU containing a Manifest TLV makes requests for those XPDUs that are out of date from the receiver's database.

Datacenter Network Using LSVR



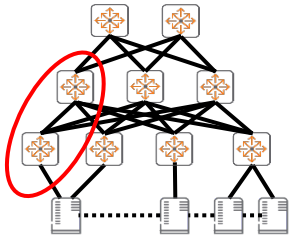
- Most datacenters are configured as 2-3 layer Clos networks using ECMP for distribution over the mesh and LAGs/M-LAGs for server attachment
- Typically these networks provide an IPv4/IPv6 topology organized with ToR and Spine switches within Pods (around 8-128 racks)
- Servers at the network edge manage virtual and tenant networks which are encapsulated into the IP packets for transmission over the data center
- The orchestrator controls the creation of the virtual and tenant networks along with coupling to services

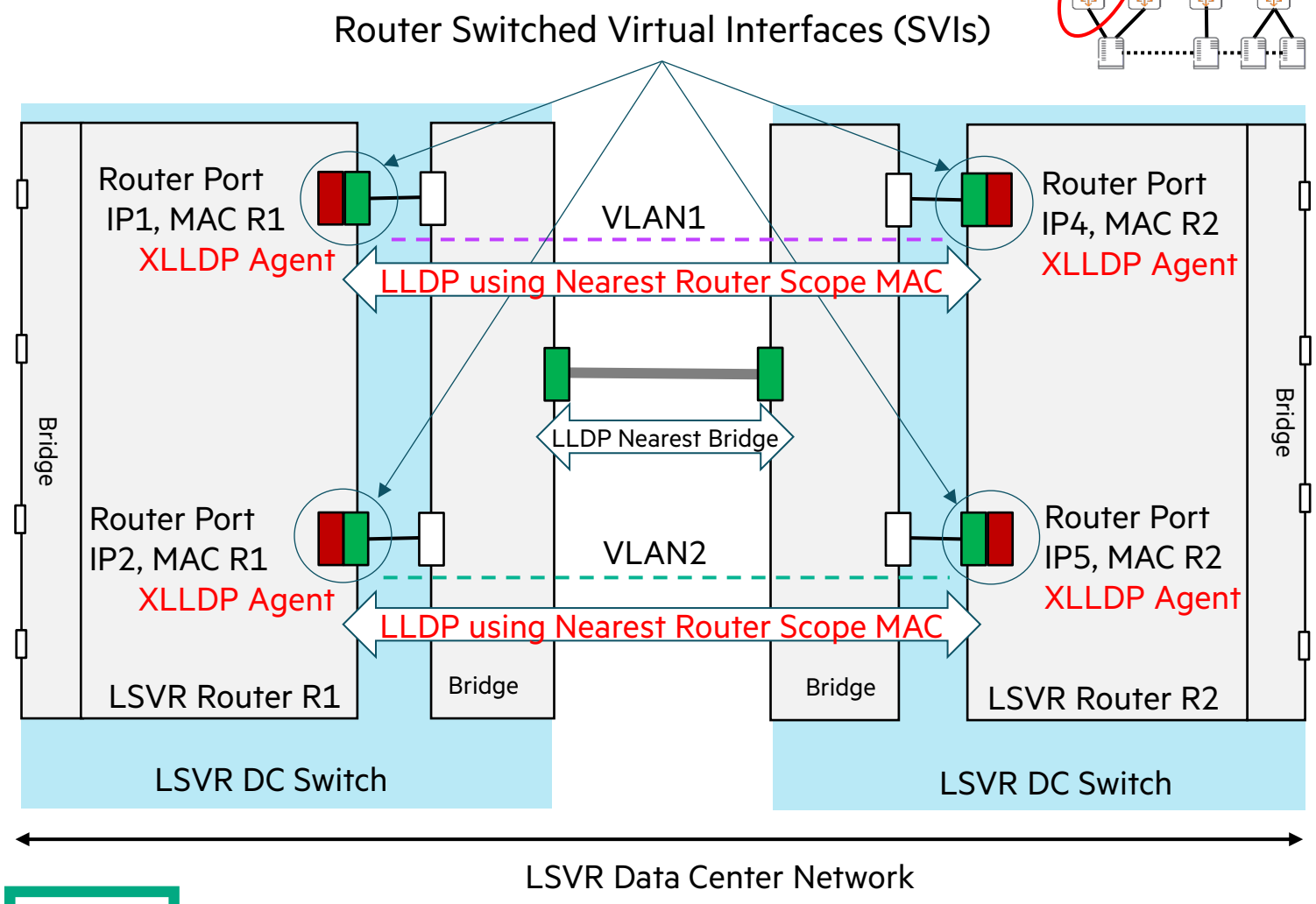
Router Interface Configurations



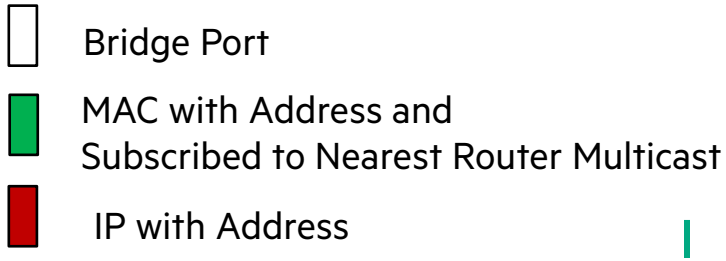
- Three distinct types of router interfaces are possible. These are:
  - Simple interface were the router has a direct attachment to a physical MAC port
  - Subinterfaces were multiple router interfaces share a physical MAC through a VLAN multiplexer
  - Switched Virtual Interfaces (SVIs) where multiple router interfaces have a virtual MAC connected to an internal bridge
- Any of the three router interface types can also be attached to other routers through an external bridge

LLDP addressing for data center switch

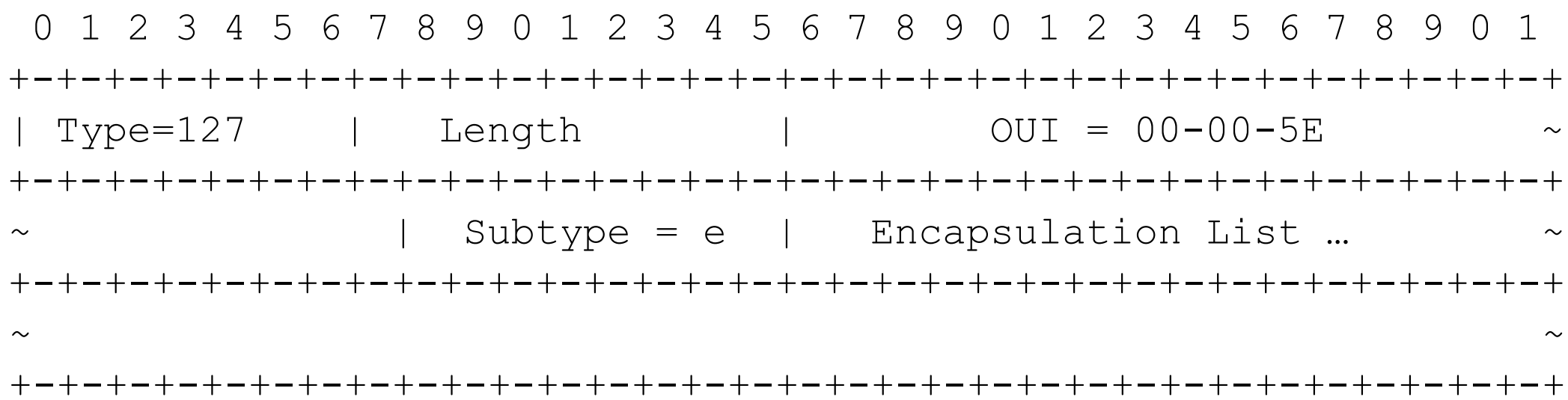




- Router Switched Virtual Interfaces (SVIs) each have MAC and IP addresses.
- VLANs are coupled to router SVIs by the bridging layer of the switch.
- LLDP can be extended between SVIs by using a “Nearest Router Multicast Address” for the LLDP destination address (Scope MAC Address).
- The Nearest Router Multicast Address must not be one of the bridge reserved addresses.
- The router SVIs subscribe to the Nearest Router Multicast Address, however it passes through the bridge layer as an ordinary multicast which is contained in a VLAN.
- Non-LSVR End Stations disregard the Nearest Router Multicast since they are not subscribed.
- The Nearest Router Multicast becomes a reserved address for routers.
- Router directly attached to tagged and untagged (subinterfaces) physical ports (no bridge layer) can use the same addressing mechanism.
- Nearest Router Multicast also works if there are bridges between the routers.

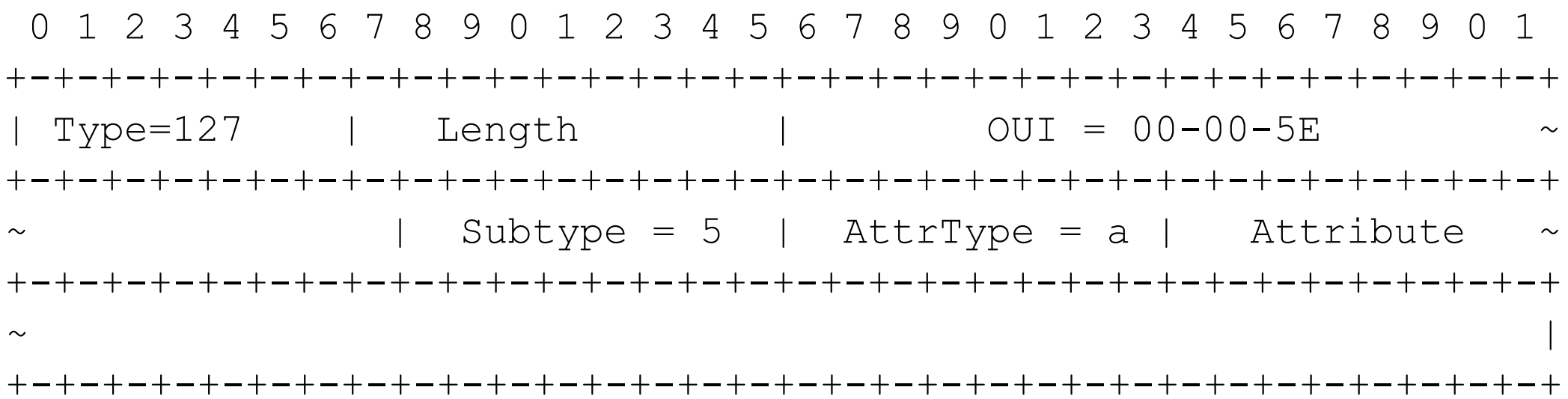


LSVR Encapsulation Types and Addressing TLVs



- Each encapsulation includes flags indicating primary, underlay/overlay, loopback
- The encapsulation subtypes are:
  - Subtype = 1: IPv4 encapsulation list each with flags, IPv4 address, prefix length
  - Subtype = 2: IPv6 encapsulation list each with flags, IPv6 address, prefix length
  - Subtype = 3: MPLS IPv4 encapsulation list each with flags, label list, IPv4 address + prefix length list
  - Subtype = 4: MPLS IPv6 encapsulation list each with flags, label list, IPv6 address + prefix length list

Upper-Layer Protocol Configuration TLV

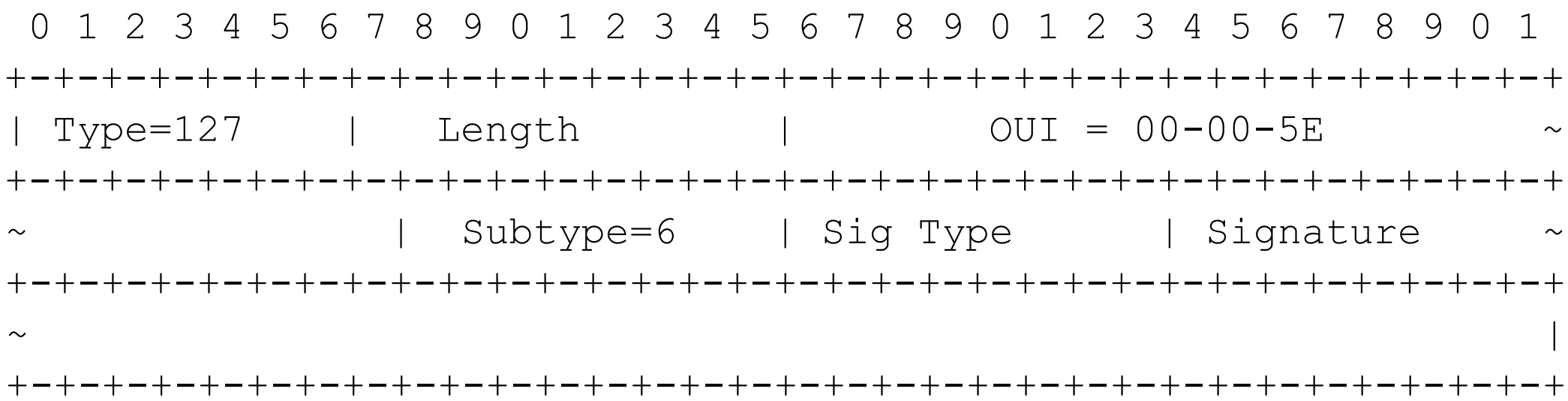


- Upper layer protocol configuration information subtype = 5
- The attribute types:
  - AttrType = 1: BGP ASN
  - AttrType = 2 : BGP IPv4 Peering Address
  - AttrType = 3 : BGP IPv6 Peering Address
  - AttrType = 4 : BGP Authentication Data
  - AttrType = 5: Misc Flags, Bit 0: GTSM, Bit 1: BFD, Bit 2-15 reserved (must be 0)

Thank You

Backup

Signature TLVs for LLDP database

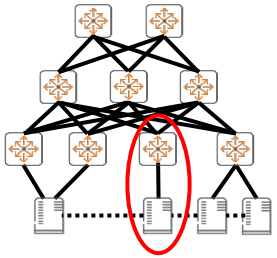


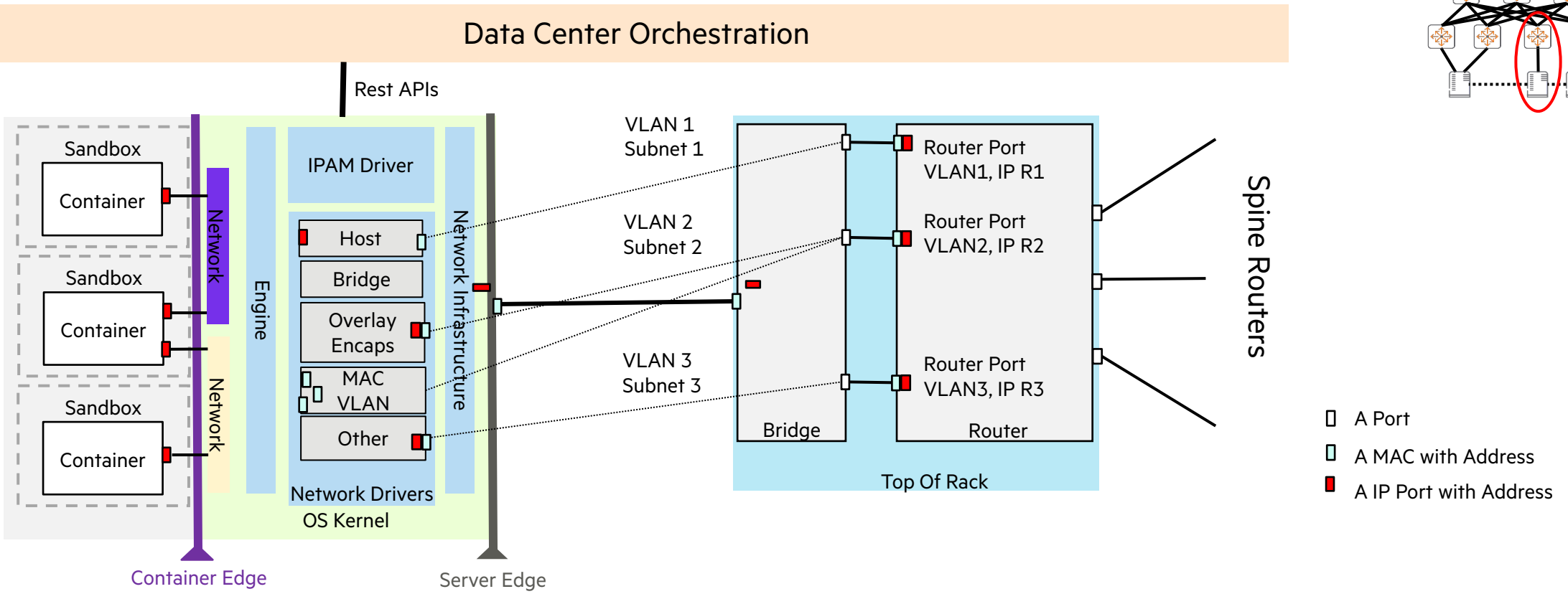
- Two signature TLV with the same format:
- Subtype=10: LLDPDU signature TLV
  - This TLV is used to sign individual LLPDUs. In particular it is used to sign the Normal LLDPDU

Chassis ID:

subtype	Basis	reference
0	Reserved	-
1	Chassis component	EntPhysicalAlias (RFC 6933)
2	Interface alias	IfAlias (RFC 2863)
3	Port component	EntPhysicalAlias (RFC 6933)
4	MAC address	MAC address (IEEE Std 802)
5	Network address	IP address string, first octet IANA family
6	Interface name	ifName (RFC 2863)
7	Locally assigned	Alpha-numeric string, local assigned
8-255	Reserved	-

Server Network Interfaces – Containers (i.e. Docker)





- Container Solutions use Linux Namespaces and Groups to isolate containers
- These solutions provide a variety of network connections, overlays for large scale datacenters
- DC network is a simple IP network. Scaling L3 encapsulations with Network Virtualization Edge like procedures