L3DL
Layer 3 Discovery & Liveness
## draft-ietf-lsvr-l3dl-03

LSVR WG

2019.11.19

randy@psg.com, sra@hactrn.net, keyur@arrcus.com

Primary Goal

Layer 3 Topology Discovery and Liveless for LSVR / BGP-SPF

Just a Reminder
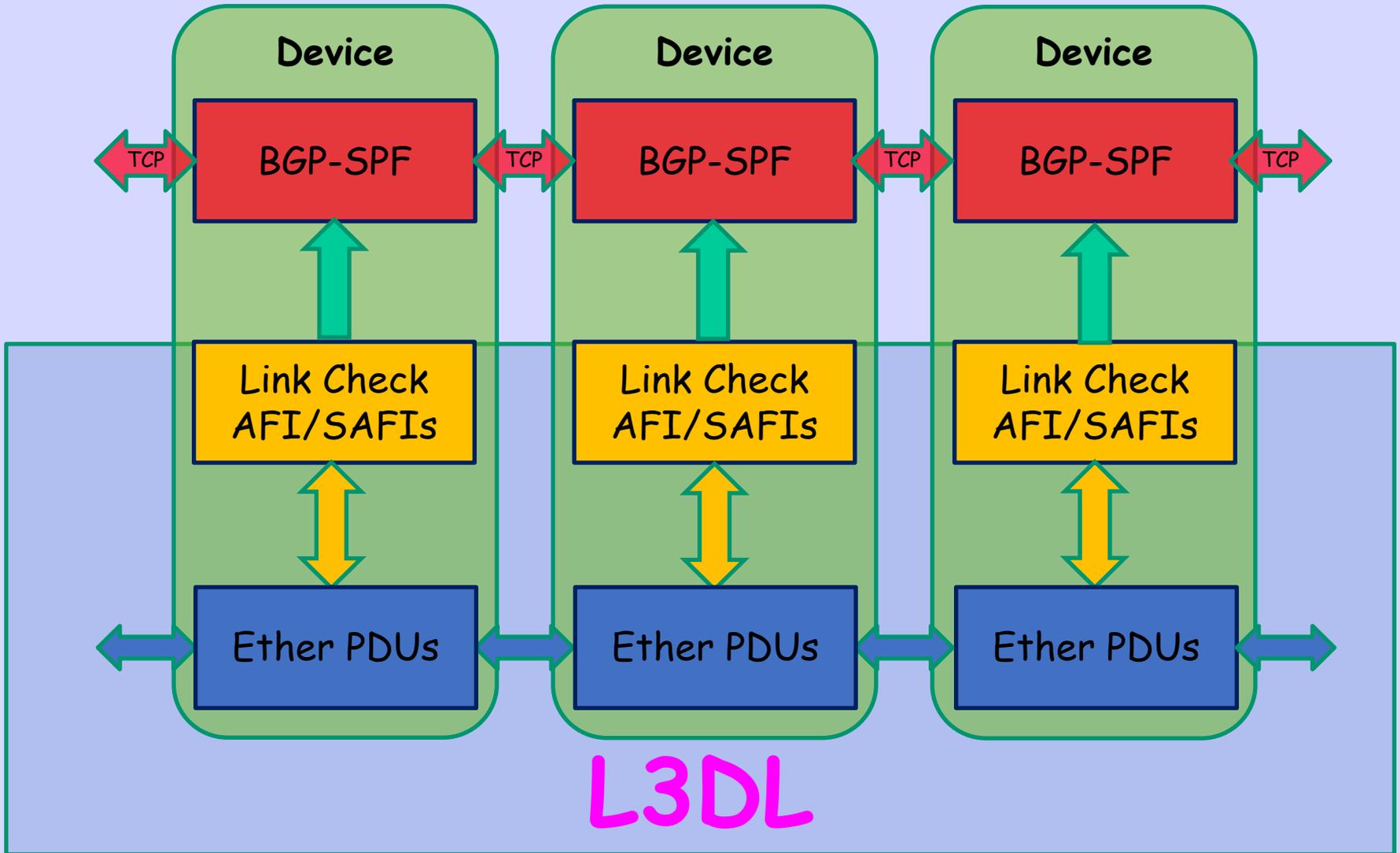
This is NOT a Routing Protocol

Discovers the Layer 3 Addresses on a PointToPoint Link

You Read the Draft

So What's New?

Devil is in the Details

# PDU Retransmission

Should a PDU need to be retransmitted, it MUST BE sent as the identical Datagram set as the original transmission.

The Transmission Sequence Number informs the receiver that it is the same PDU.

Transmission Seq No

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Version    | Transmission Sequence Number |L|            ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~       Datagram Number         |         Datagram Length       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Checksum                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Payload...                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
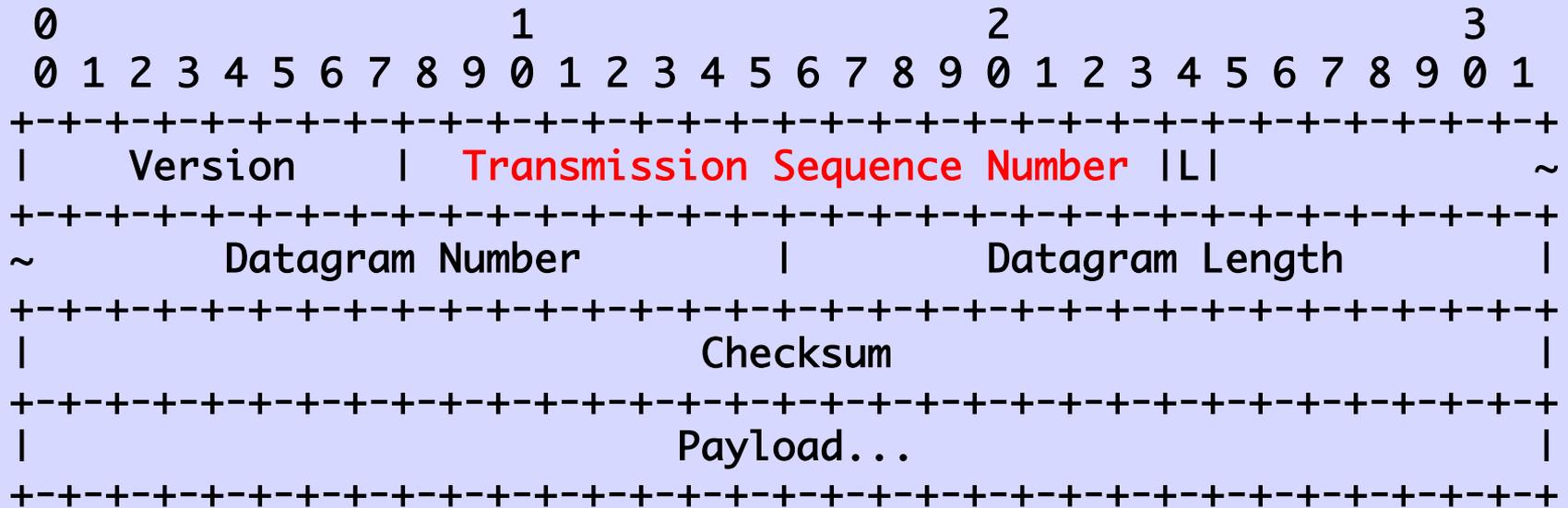
Clarification

If a Datagram fails checksum verification, the datagram is invalid and should be silently discarded. The sender will retransmit the PDU, and the receiver can assemble it.

Oracle Removed

Authorization Failure elsewhere no longer differentiated from Authorization Failure in OPEN

Transport Layer can handle 2^32 octet PDUs, and is now Clearly Delineated

2^24 Datagrams (in Ethernet Frames)
2^16 Octets/Datagram (except it is a Frame)

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Version    | Transmission Sequence Number |L|  Datagram   ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Number     |            Datagram Length                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Checksum                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Payload...                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Ether Frame   Ether Frame   Ether Frame   Ether Frame

Datagram   Datagram   Datagram   Datagram

PDU

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   PDU Type    |            Payload Length                    ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~               |            Payload ...                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Sig Type    |       Signature Length       |               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+               +
~                         Signature                            ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
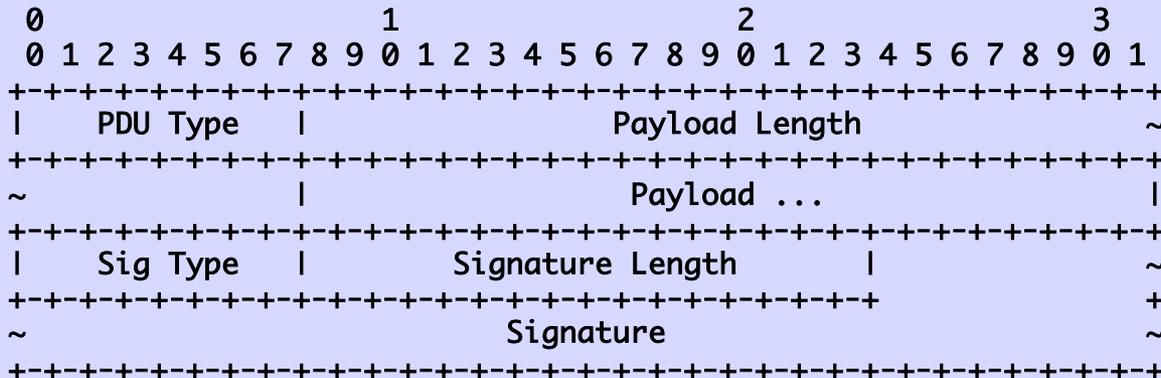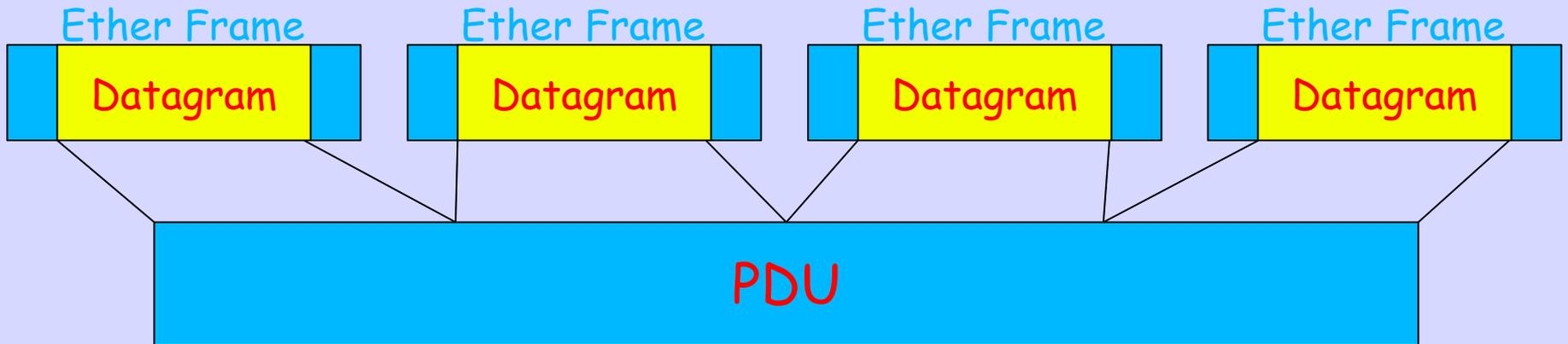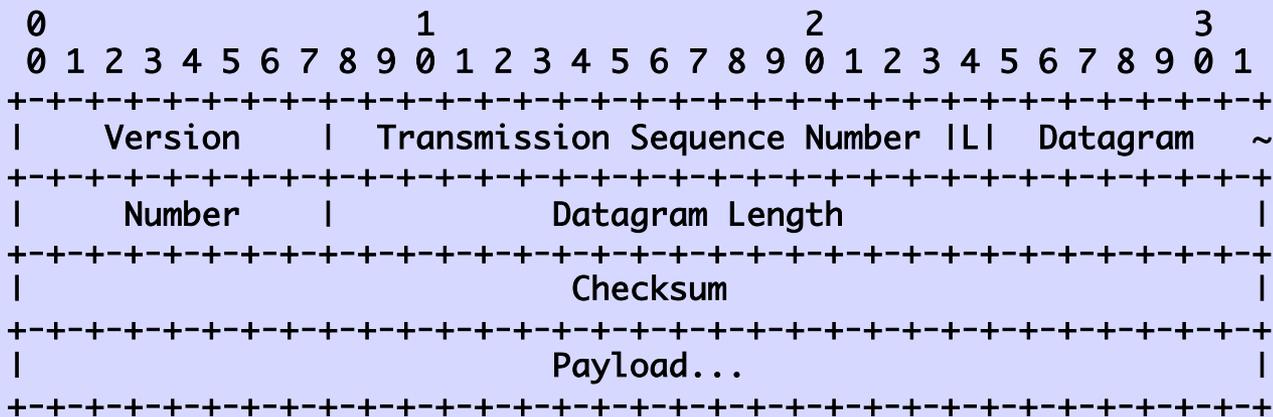
Why not TCP?

- When this runs, there are no IP Addresses

- This protocol is to <u>Learn IP Addresses</u>

- So it is a cheap TCP-like protocol

- Reassembly of out of order Datagrams

- Retransmission with Back-off

- PDUs are ACKnowledged

- Long Lived Sessions

- ...

Fully Stateful Session Per Peer

Graceful Restart

State May Be Resumed á la BGP

Encaps etc PDUs

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    PDU Type    |                Payload Length                ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~               |                  Count                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Serial Number                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Encapsulation List...           |     Sig Type     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Signature Length       |         Signature ...          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

OPEN PDU

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| PDU Type = 1 |                  Payload Length              ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~             |                    Nonce                       ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~            | LLEI Length  |            My LLEI              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-~
~                           | AttrCount  |                     ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~      Attribute List ...   |   Auth Type    |   Key Length  ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~            |                 Key ...                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Serial Number                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Sig Type  |          Signature Length       | Signature ... |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
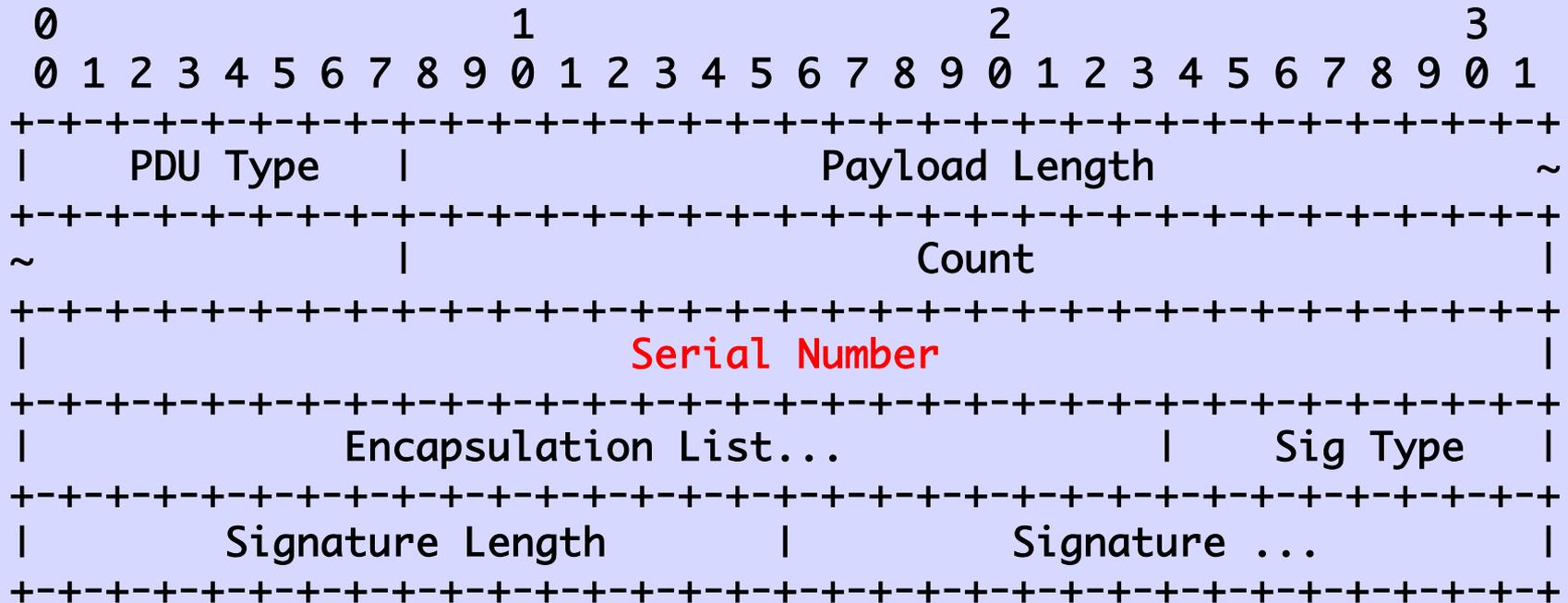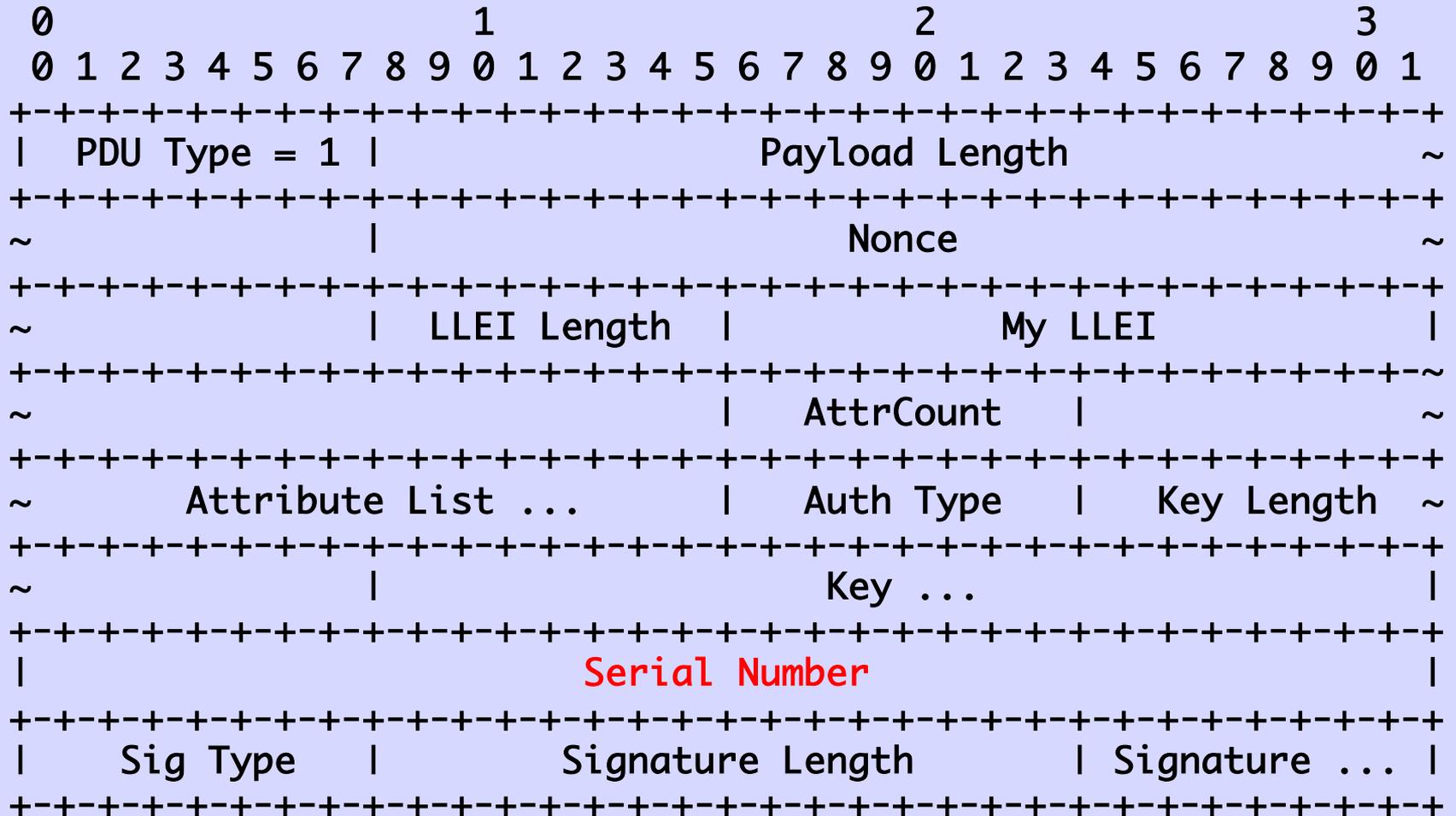
Duplicate Datagram and PDU Detection

Retransmission on Time-Out for ACK

Announce/Withdraw

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  PDU Type = 4 |                 Payload Length                ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~               |                    Count                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Serial Number                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Encaps Flags  |                  IPv4 Address                 ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~               |    PrefixLen    |    more ...    |   Sig Type  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Signature Length        |          Signature ...      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+


 0           1           2           3           4  ...      7
+-----------+-----------+-----------+-----------+-----------+
|  Ann/With |  Primary  | Under/Over|  Loopback | Reserved ..|
+-----------+-----------+-----------+-----------+-----------+
```

Explicit ACK/ERROR

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  PDU Type = 3 |                Payload Length = 5            ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~               |   ACKed PDU   | EType |        Error Code    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Error Hint          |    Sig Type   |Signature Leng.~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~               |               Signature ...                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
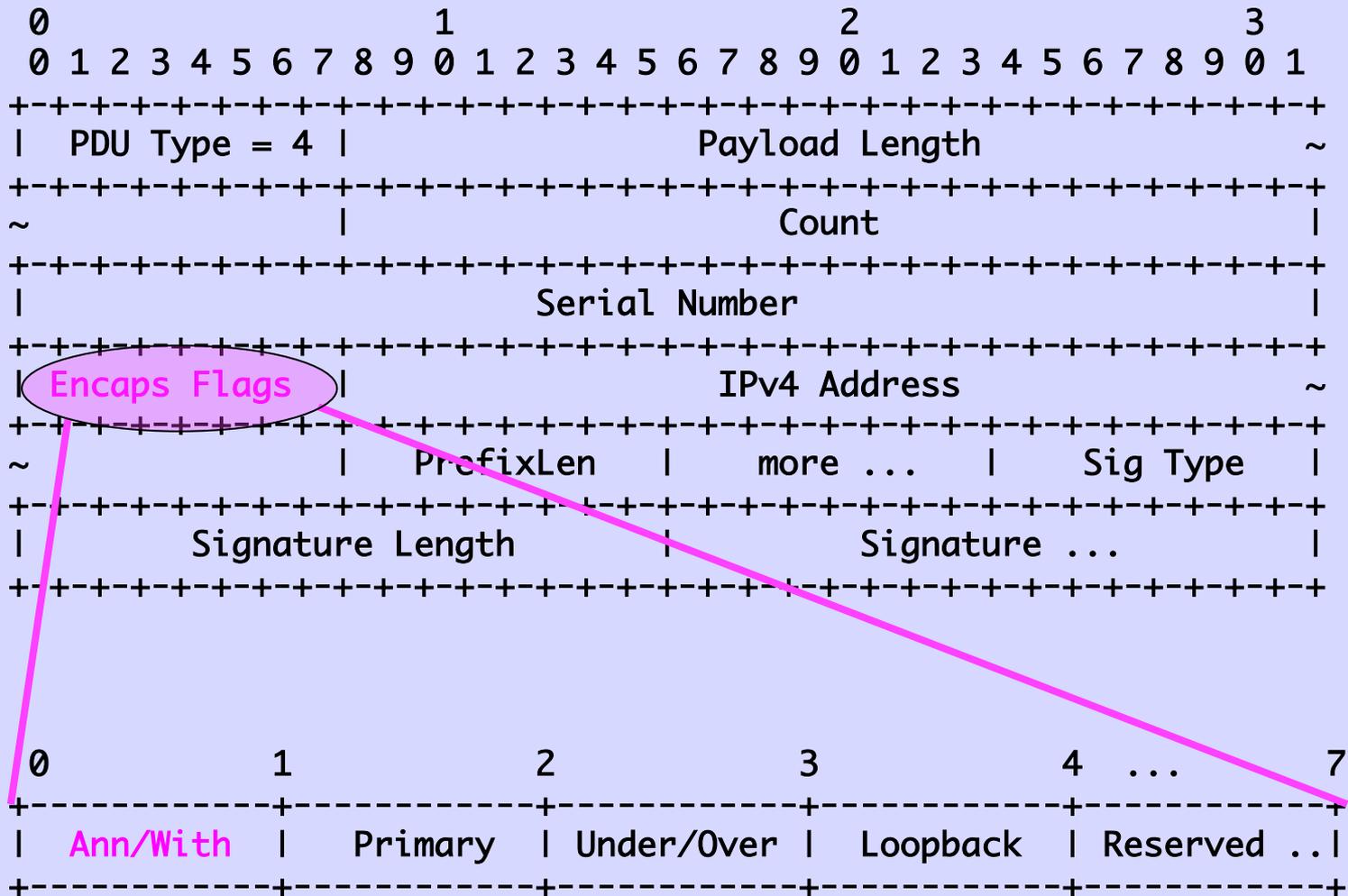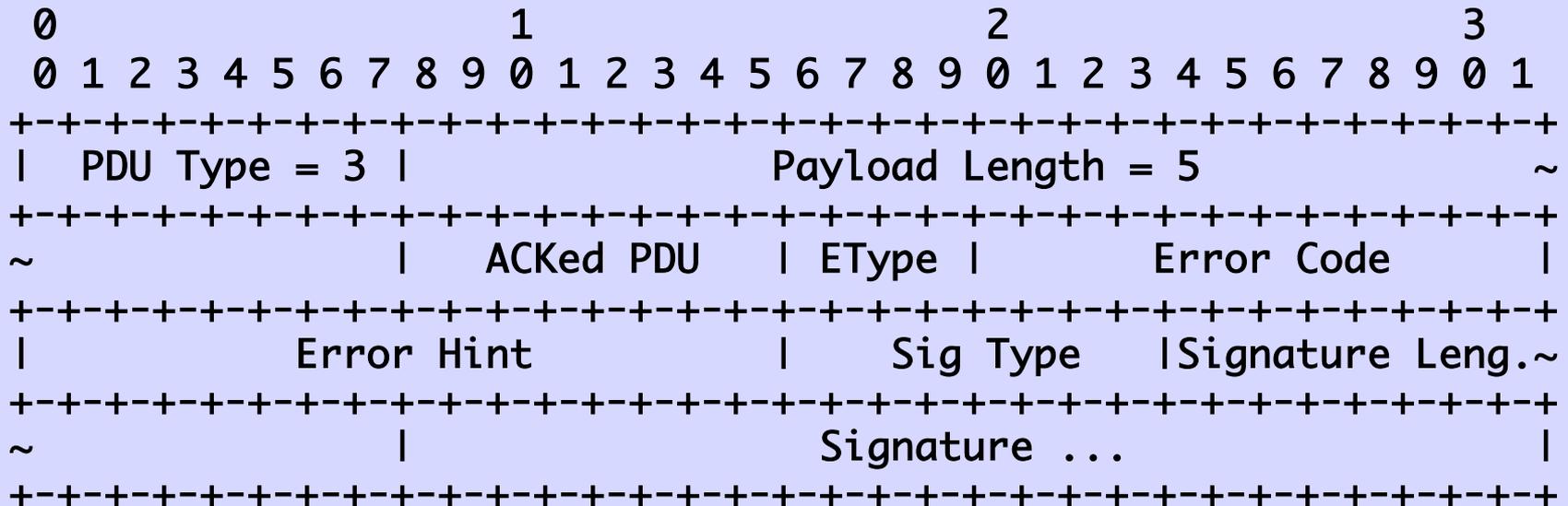
EType, Error Code, Error Hint = 0   /* no error, just an ACK */

Vendor Extension

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Type = 255   |             Length            |    ...      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Enterprise Number            |    Ent Type      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Enterprise Data ...                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
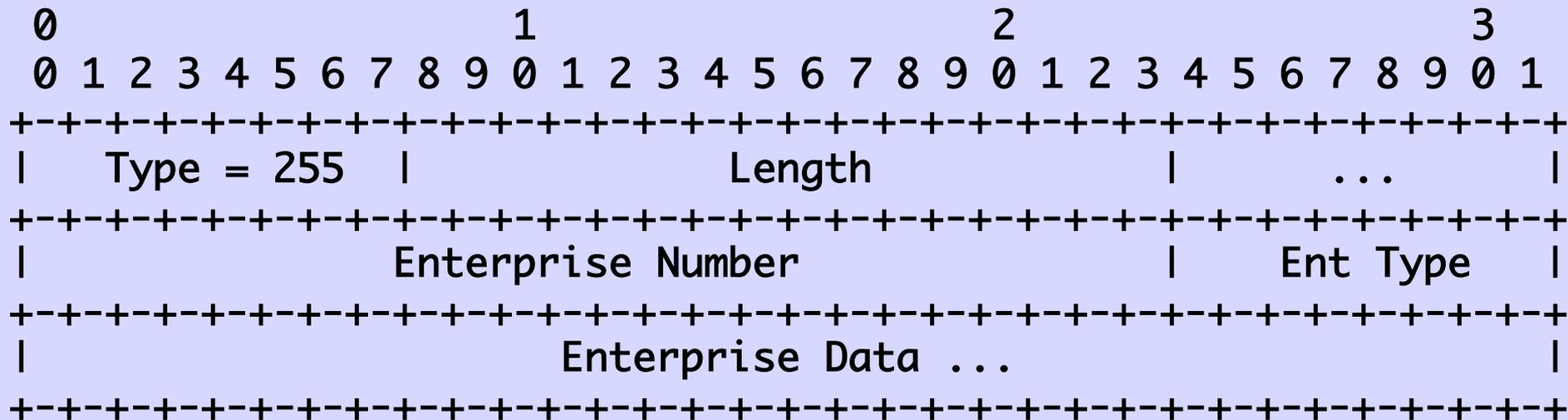
We now have Two Implementations
One Python3 (LSOE)
One in Golang

Going into ArcOS & Maybe a FOSS, with Interop in Vancouver

Time to WGLC?

Neeraj has Extended it
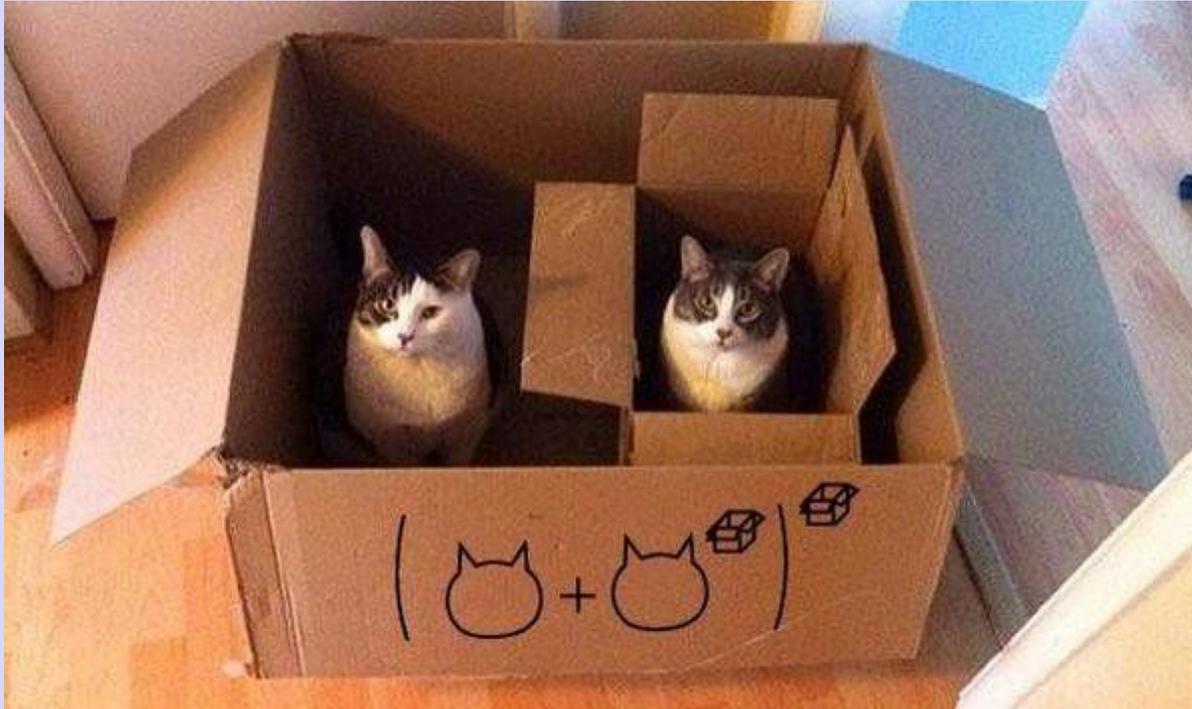
Always a Good Sign

`draft-malhotra-bess-evpn-pe-ce-00`

L3DL-Signing
Layer 3 Discovery and Liveness Signing

`draft-ymbk-lsvr-l3dl-signing`

LSVR WG

2019.11.19

randy@psg.com, sra@hactrn.net, keyur@arrcus.com

OPEN PDU

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| PDU Type = 1 |                 Payload Length                ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~              |                    Nonce                       ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~             | LLEI Length  |              My LLEI             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-~
~                            | AttrCount     |                  ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~       Attribute List ...   |  Auth Type    |   Key Length   ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~             |                     Key ...                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Serial Number                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Sig Type   |       Signature Length     | Signature ... |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
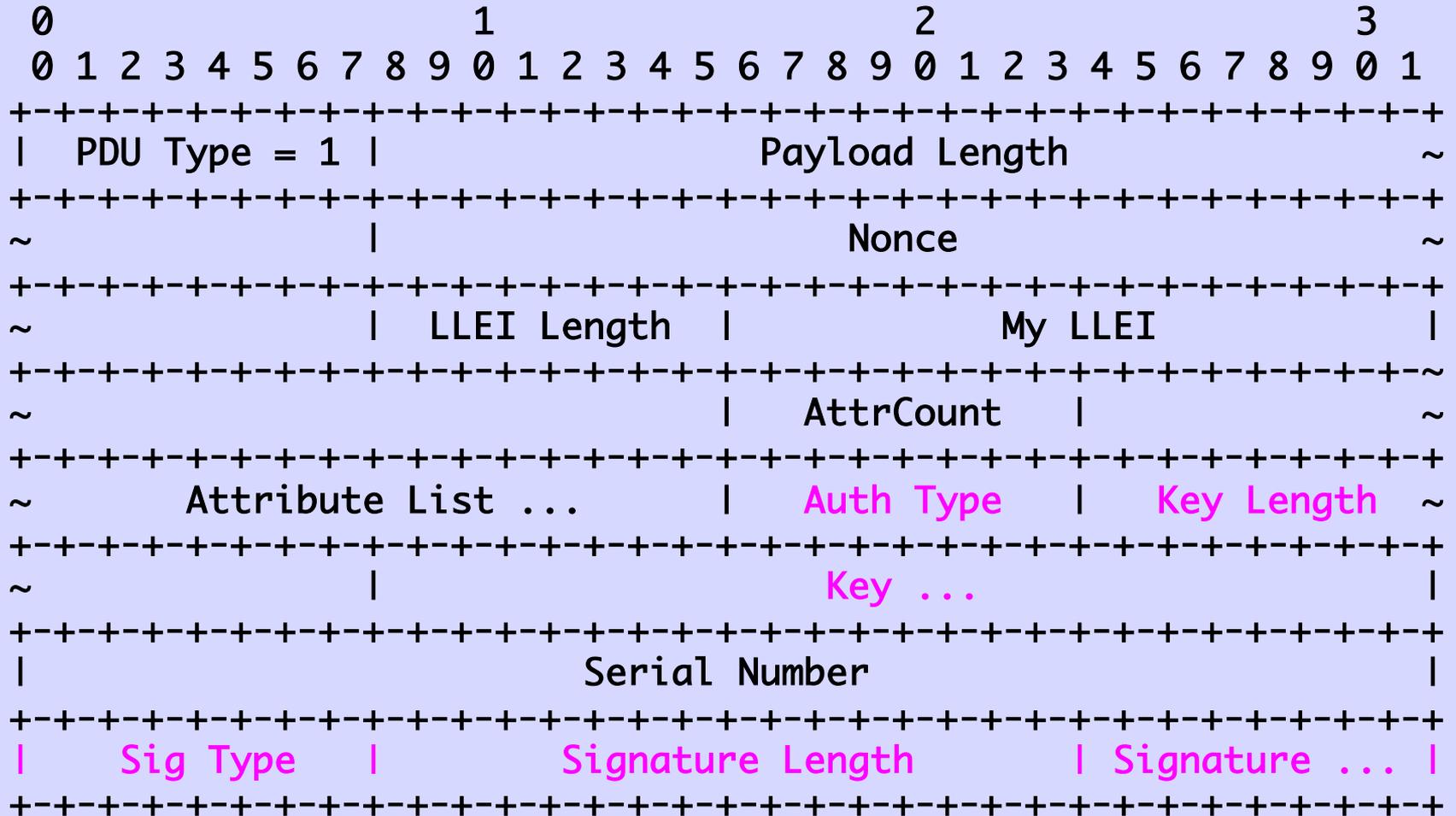
PDU Sender Signing

- The Key in the OPEN PDU SHOULD be the public key of an asymmetric key pair.

- The sender signs with the private key, of course.

- The device sending the OPEN may use one key for all links, a different key for each link, or some aggregation(s) thereof

Trust on First Use (TOFU)

- The OPEN key is generated on the sending device

- It is believed without question by the receiver

- Used to verify all subsequent PDUs from the same sender with the same Key Type.

PKI-Based Keying

- An enrollment step is performed.

- The public key is put into a certificate, which is signed by the the operational environment's trust anchor.

- The relying party can be confident that the public key is under control of the identified L3DL protocol entity.

Do Not Be Afraid

Creative Commons: Attribution & Share Alike

This is NOT X.509

- These need not be X.509 certificates

- X.509 is much more complicated than we need

- They are just signatures of one key (the session key supplied in the Key field of the OPEN PDU) by another key (the trust anchor)

- Every device must have TA burned in

Verify is the Same

- The two methods are indistinguishable

- The key provided in the OPEN PDU is used to verify the signatures of subsequent PDUs.

- The difference that PKI-based keys may be verified against the trust anchor when the OPEN PDU is received.

The Choice of Which Keying is Left to the Operator

Request
# WG Adoption

L3DL-ULPC
Upper Layer Protocol Configuration
draft-ymbk-lsvr-l3dl-ulpc-02

LSVR WG

2019.11.19

randy@psg.com, sra@hactrn.net, keyur@arrcus.com

L3DL PDU for ULPC

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Type = 9    |                 Payload Length             ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~               |    ULPC Type    |    AttrCount    |           ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~        Attribute List ...       |    Sig Type    | Signature Len ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~               |                 Signature ...                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
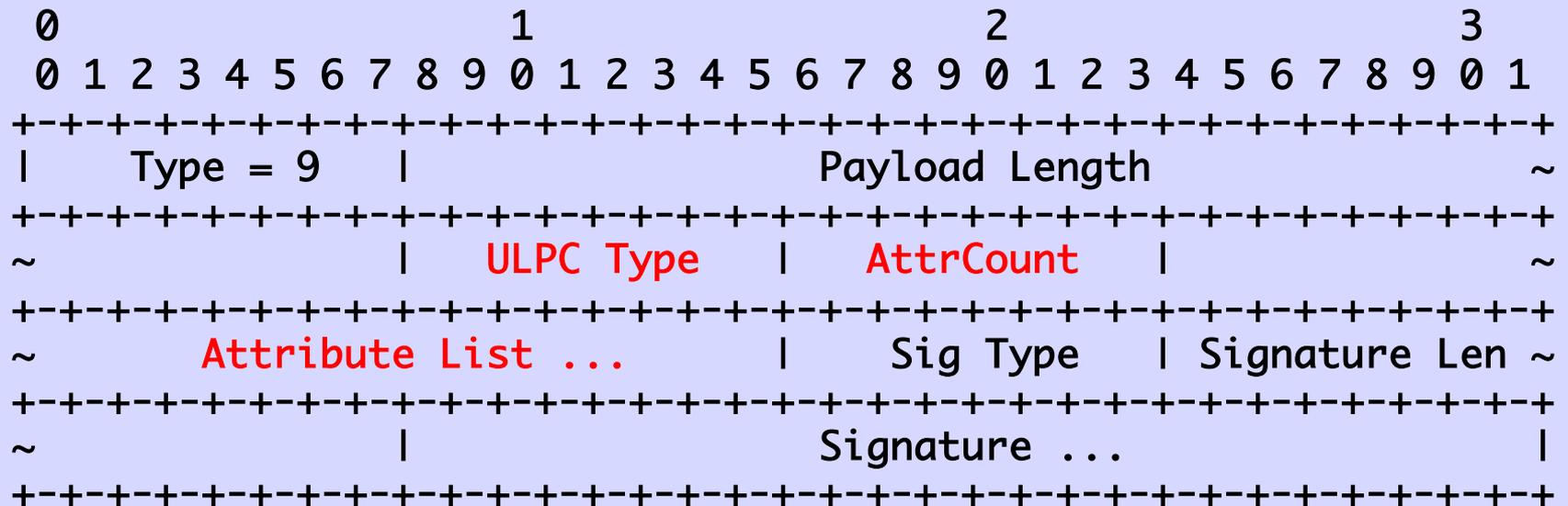
Provide the <u>minimal</u> set of configuration parameters for BGP OPEN to succeed

# Not to replace or conflict with data exchanged by BGP OPEN

# Multiple sources of truth are a recipe for complexity and pain

ULPC for BGP

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Attr Type = 1 | Attr Len = 48 |           My ASN             ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Attr Type = 2 | Attr Len = 56 |   My IPv4 Peering Address    ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                               |   Prefix Len  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Attr Type = 4 |    Attr Len   |                              ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+                              ~
~               BGP Authentication Data ...                    ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Yes, there is one for IPv6 ☺

Request
WG Adoption

And For Dessert

An Ops Hack (not a draft) At Layer 3 With L3DL

TOR

TOR

Server with MicroServices

veth0 veth0 veth0 veth0 veth0 veth0 veth0 veth0

sw0 sw1 sw2 sw3 sw4 sw5 sw6 sw7

L3 'Switch' AKA Router
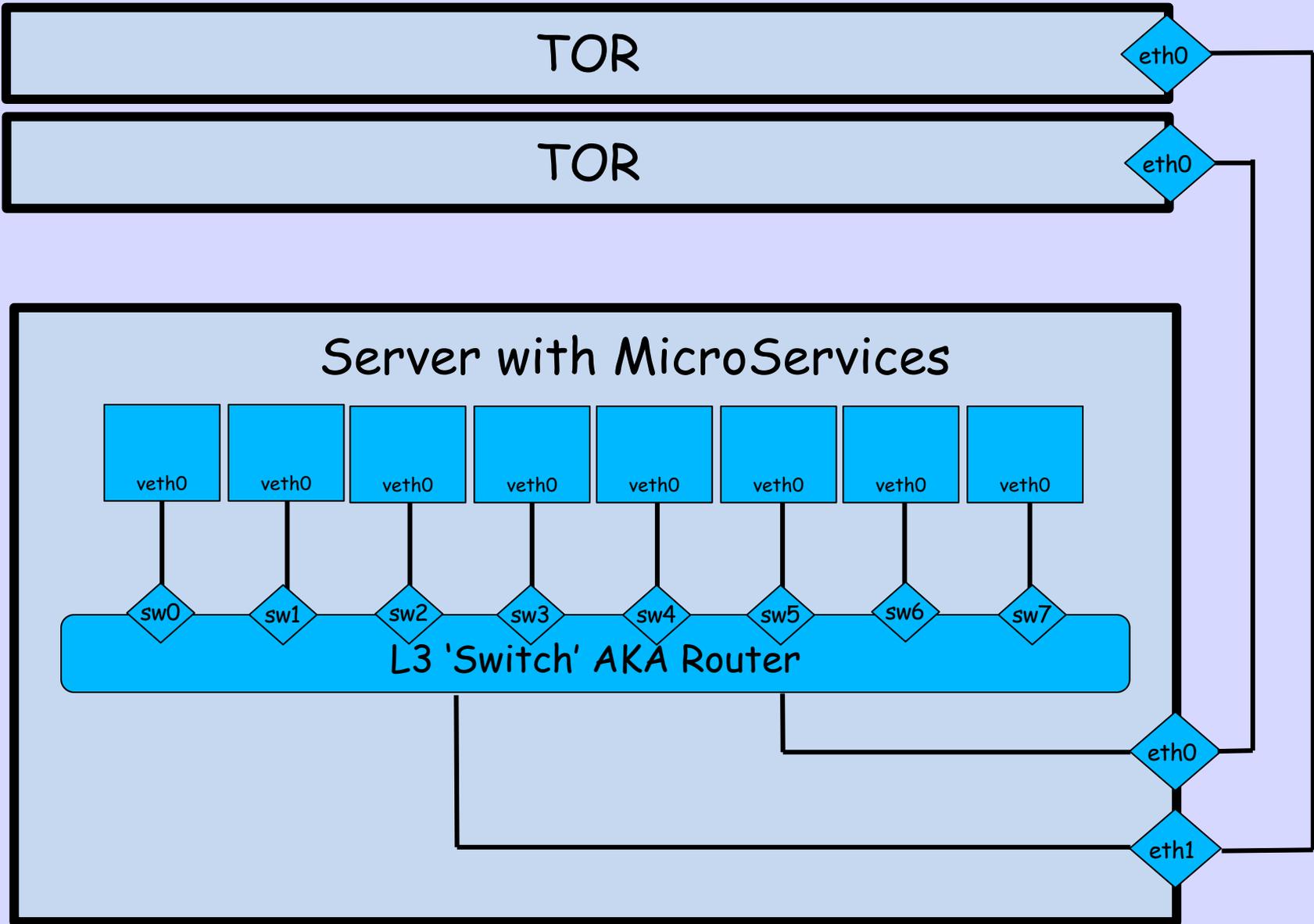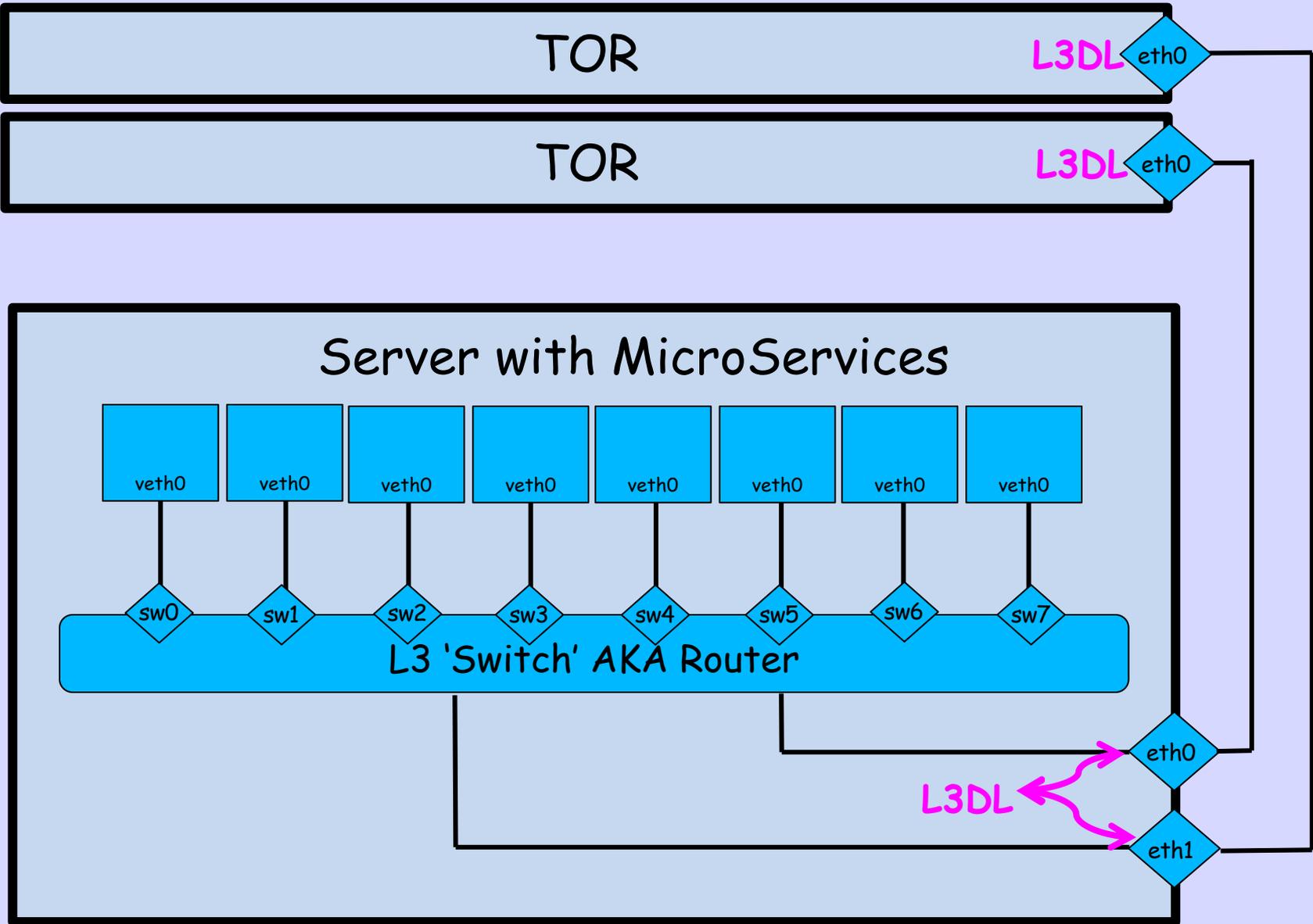
eth0

eth0

eth0

eth1

L3DL Config:
    Interface eth0
    Interface eth1

Active and Passive

```
protocols {
  isis {
      interface lo0.0 passive;
      interface ge-0/0/0.0 level 2 metric 2;
      interface ge-0/0/1.0 passive;
      interface ge-0/0/2.0 level 2 metric 42;
      interface gr-0/2/0.0 passive;
      }
  }
```

Active and Passive

interface ge-0/0/0.0 level 2 metric 2;

- It sends and receives protocol
- It injects its L3 into protocol

interface ge-0/0/1.0 passive;

- It DOES NOT send or receive protocol
- It injects its L3 into protocol

Announce/Withdraw

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| PDU Type = 4 |               Payload Length                 ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~              |               Count                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Serial Number                                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Encaps Flags |               IPv4 Address                    ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~             | PrefixLen    |   more ...    |   Sig Type     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Signature Length      |         Signature ...          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+


 0           1           2           3           4   ...     7
+-----------+-----------+-----------+-----------+-----------+
| Ann/With  |  Primary  | Under/Over | Loopback  | Reserved ..|
+-----------+-----------+-----------+-----------+-----------+
```
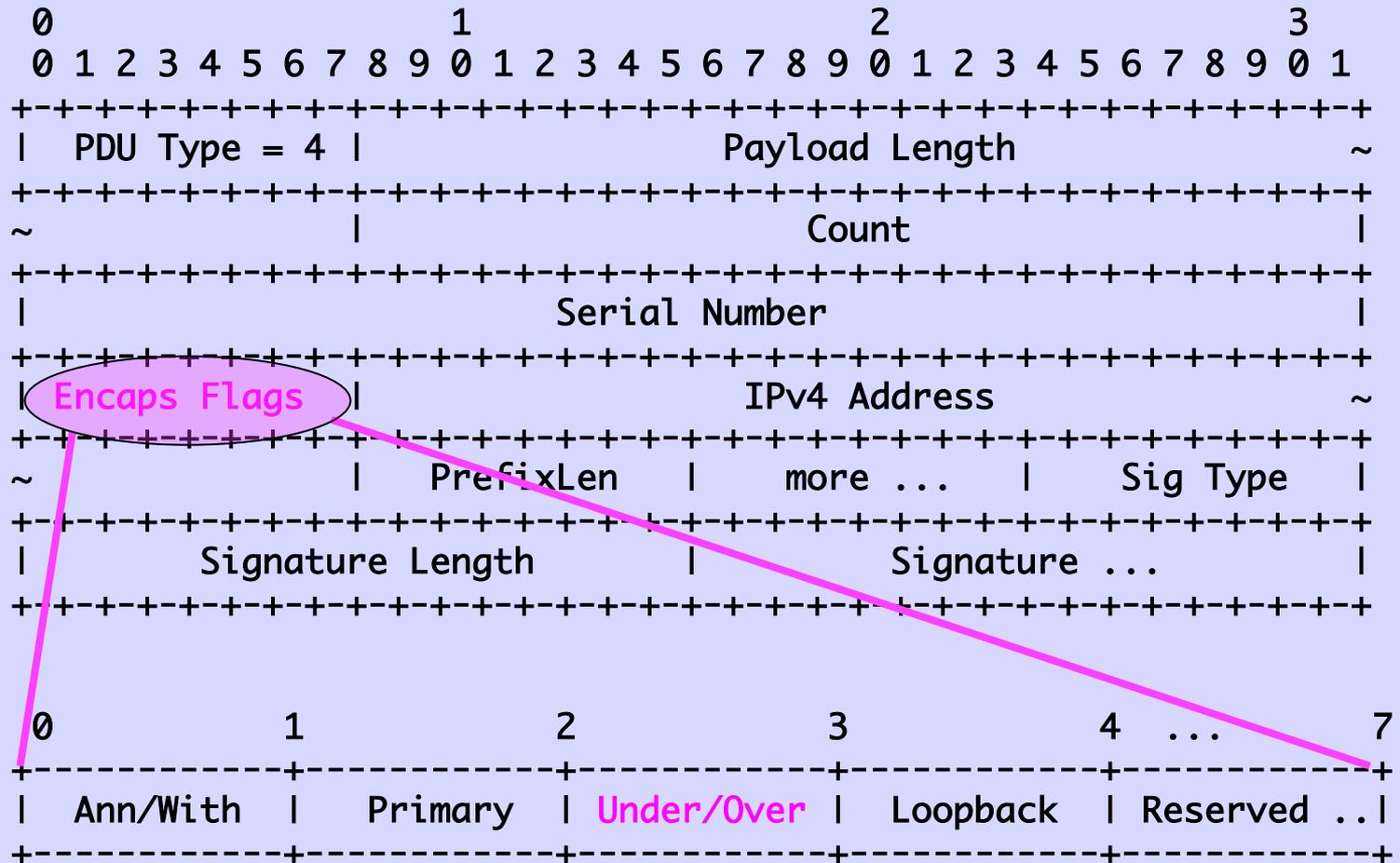
Underlay == Passive
Overlay == Active
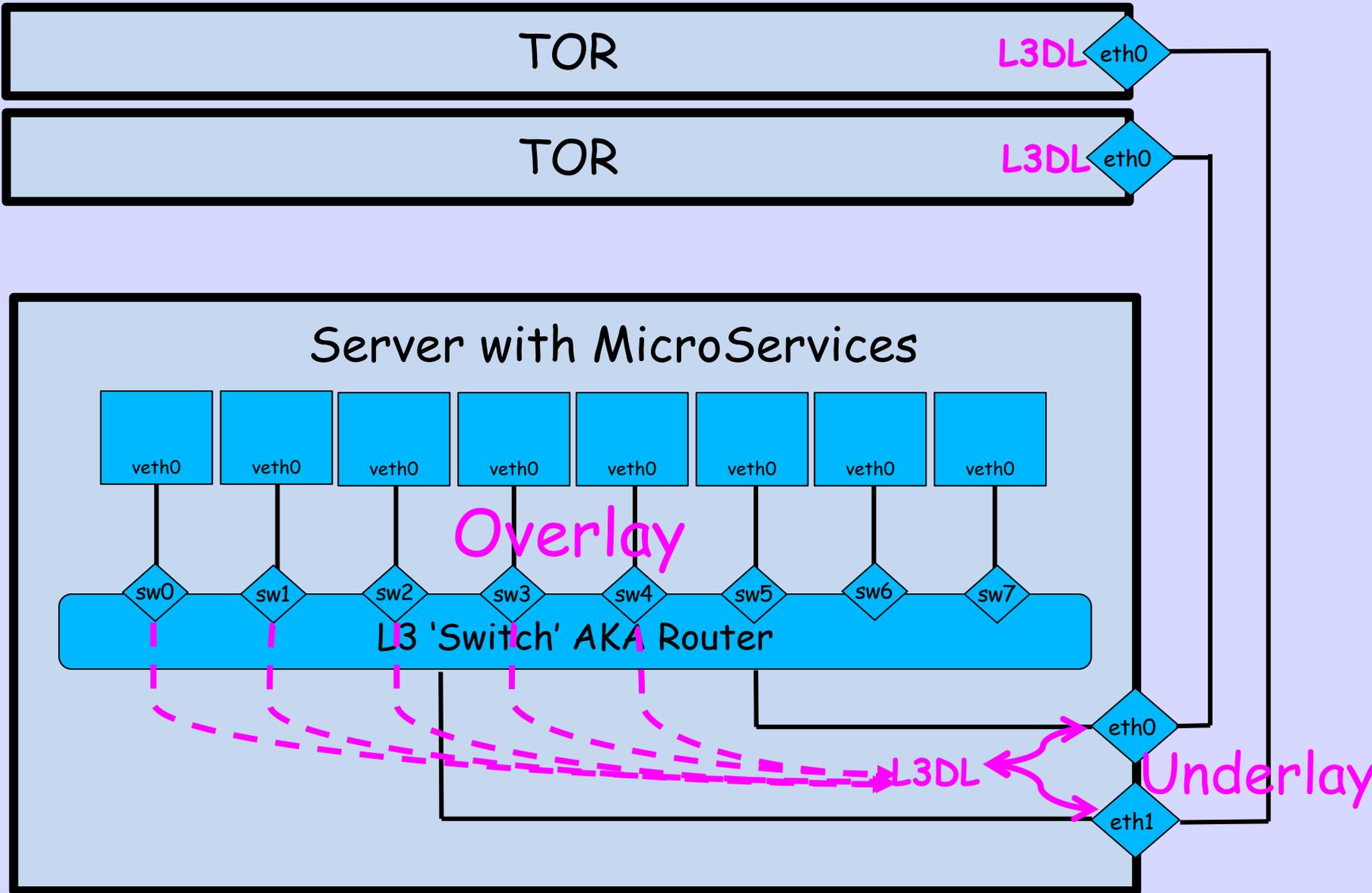
TOR L3DL eth0

TOR L3DL eth0

Server with MicroServices

veth0 veth0 veth0 veth0 veth0 veth0 veth0 veth0

Overlay

sw0 sw1 sw2 sw3 sw4 sw5 sw6 sw7

L3 'Switch' AKA Router

L3DL

eth0

Underlay

eth1

L3DL Config:
  Interface eth0
    Redistribute sw[0-7} as overlay

TOR   VRFs   L3DL   eth0

TOR   VRFs   L3DL   eth0

Multi-Tennancy

Server with MicroServices

veth0 veth0 veth0 veth0 veth0 veth0 veth0 veth0

Overlay

sw0 sw1 sw2 sw3 sw4 sw5 sw6 sw7

L3 'Switch' AKA Router with VRFs

VXlan

eth0

L3DL

Underlay

eth1
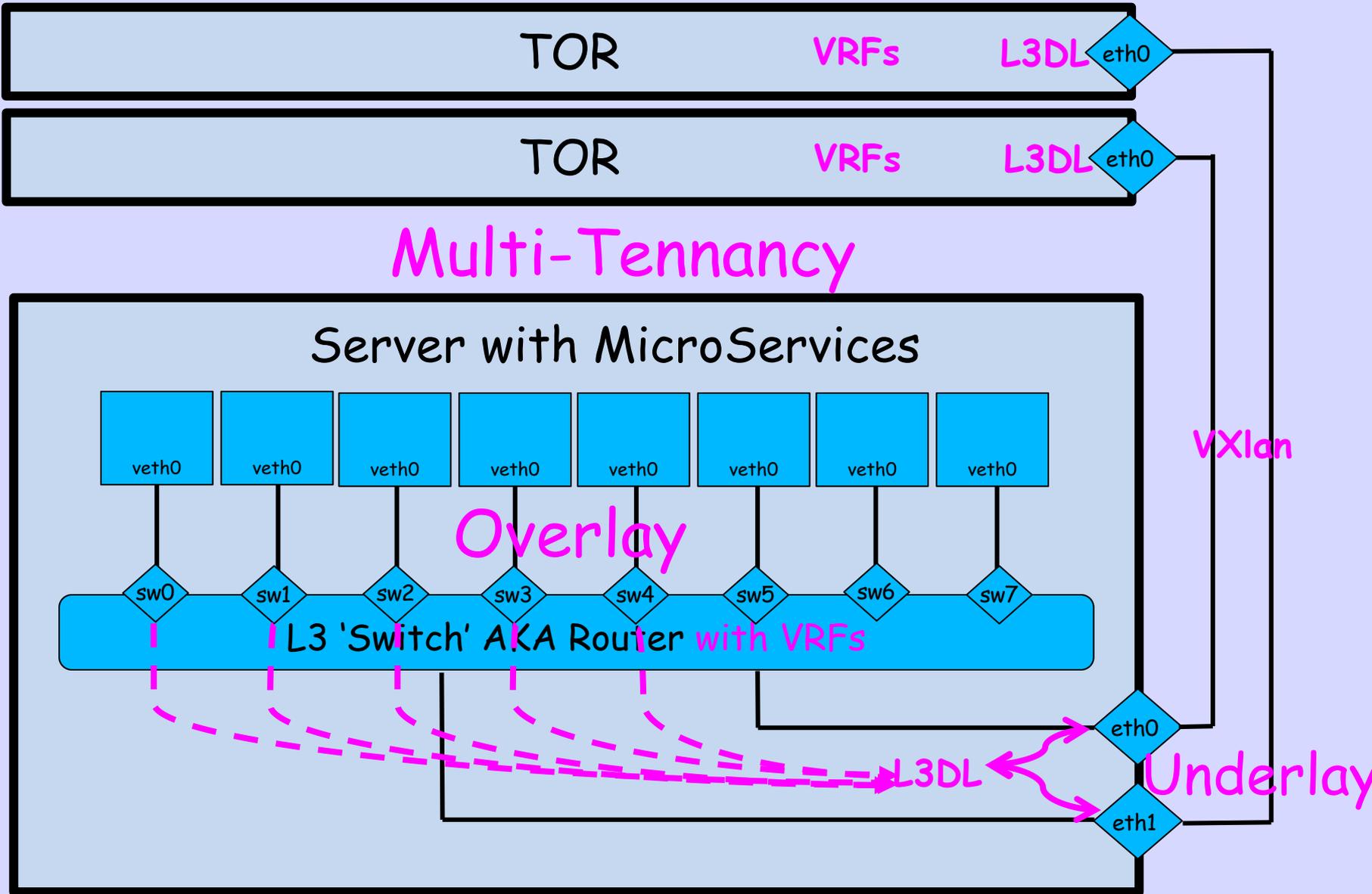
L3DL Config:
    Interface eth0
        Redistribute sw[0-7} as overlay

Ops Considerations

- Hundreds of μServices on a server

- 30-40 servers per rack

- Can cross racks, pods, …

- Service mobility

  - Creation and Deletion

  - μservice failure

  - Server failure

- It's Quiet: Only changes are propagated