MERGING SERVICE PROVIDER NETWORKS

Bastian Hoss

Markus Jungbluth

DKNOG16 – Mar 5th – 6th

MERGING SERVICE PROVIDERS?

xantaro

- In the past 15 years many German service providers were acquired by other service providers.

- This is usually done as part of a **„build-and-buy"** strategy.

- In many cases **both** the **acquiring** and the **acquired** service provider operate an **IP/MPLS network**!

- After the deal has been closed, the **real fun begins**!

- **Disclaimer**: There are many other challenges in this context. For example integrating optical networks, IT systems, teams, etc.
  This talk focuses mainly on IP/MPLS networks.

The sale is completed: HanseNet is now part of Telefónica

EQT acquires inexio, a leading provider of fiber-optic internet access in Germany

EQT, OMERS to buy Deutsche Glasfaser, increasing competition in fibre-optic rollout
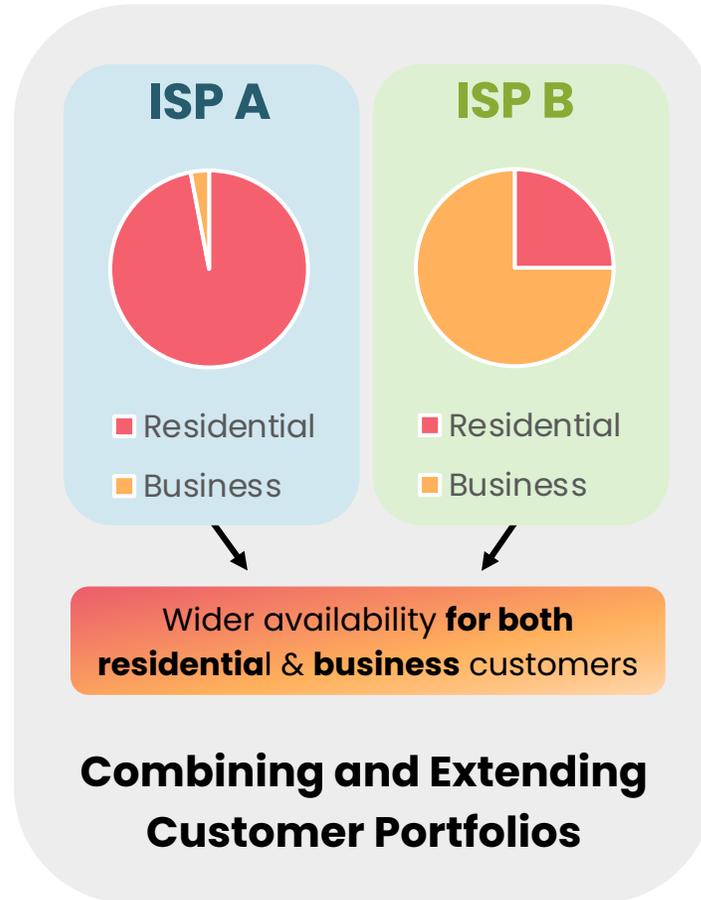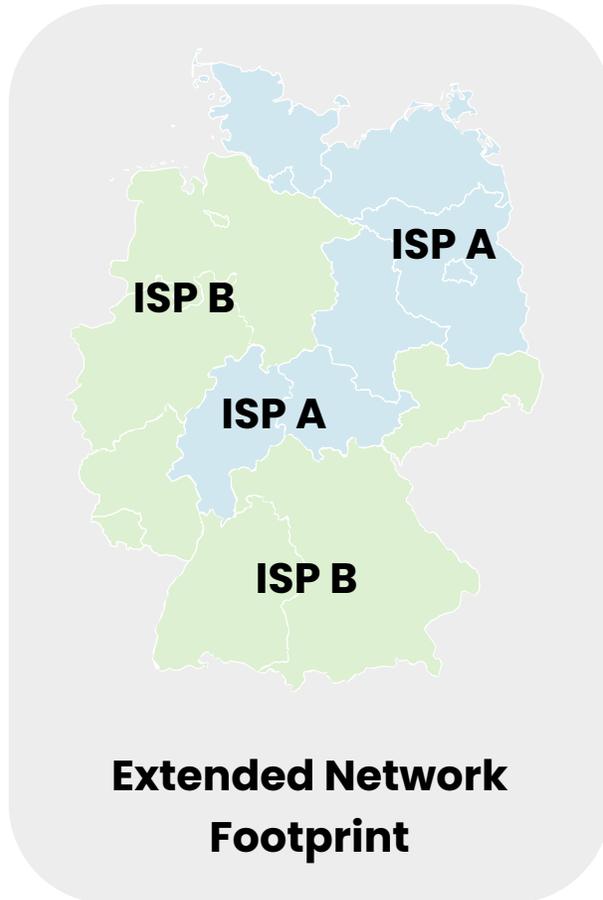
KEVAG Telekom und teliko werden eins!

Versatel AG übernimmt die KielNET GmbH

**Vodafone completes takeover of Unitymedia, UPC from Liberty Global**

UNSERE GRÜNE GLASFASER ÜBERNIMMT IN-FRAFIBRE GERMANY UND WIRD DAMIT ZU EINEM DER GRÖSSTEN GLASFASERANBIETER DEUTSCHLANDS

MOTIVATION FOR A MERGER

xantaro

Extended Network Footprint

ISP A

ISP B

ISP A

ISP B

Combining and Extending Customer Portfolios

ISP A

- Residential
- Business

ISP B

- Residential
- Business

Wider availability **for both residential** & **business** customers

- **Sharing Resources**
  - PoPs / DCs
  - Fibers
  - IPv4 Networks
- **Sharing Services**
  - Peering / Upstream
  - Caching
  - L3BSA

… and more!

Cost Savings

HOW DO WE MERGE THE NETWORKS?


ISP A

IP/MPLS

ISP B

IP/MPLS
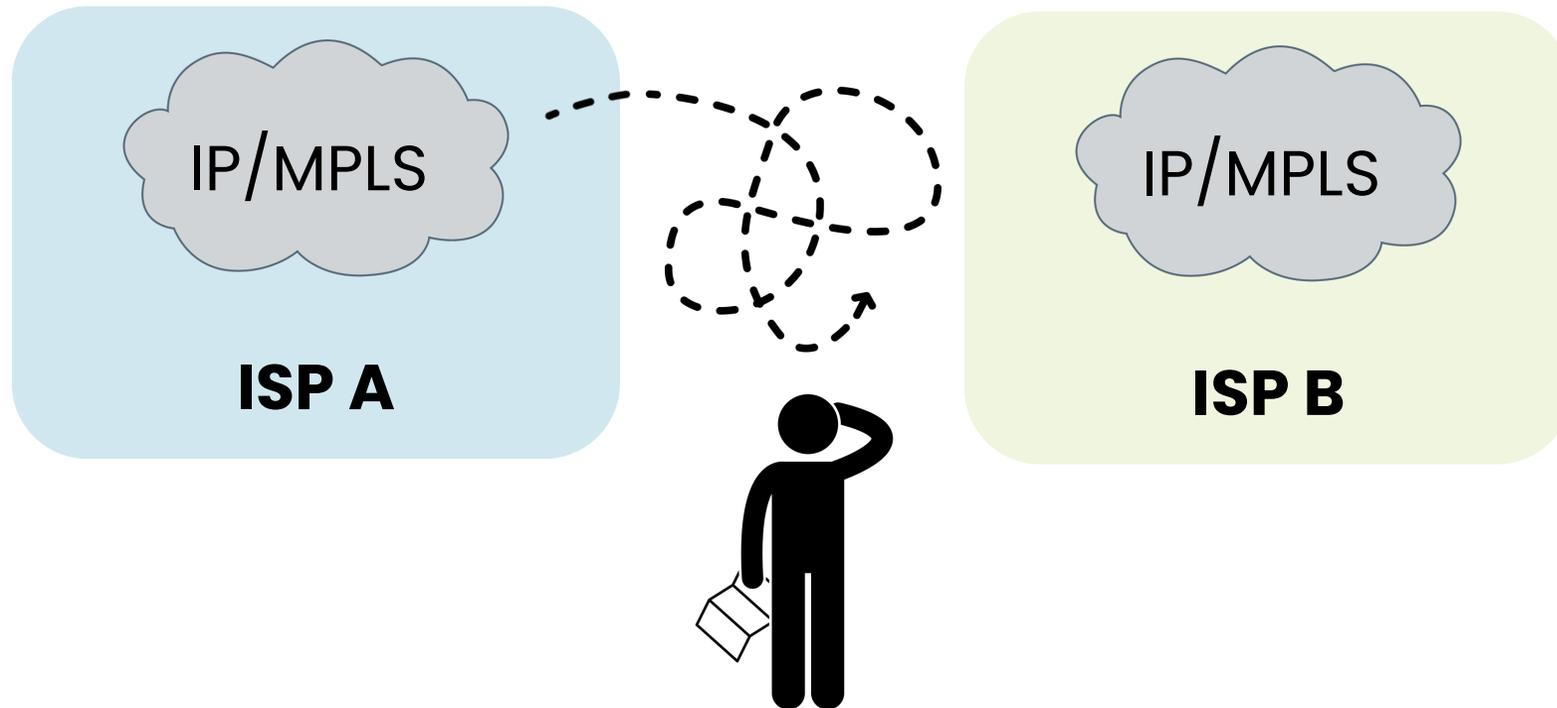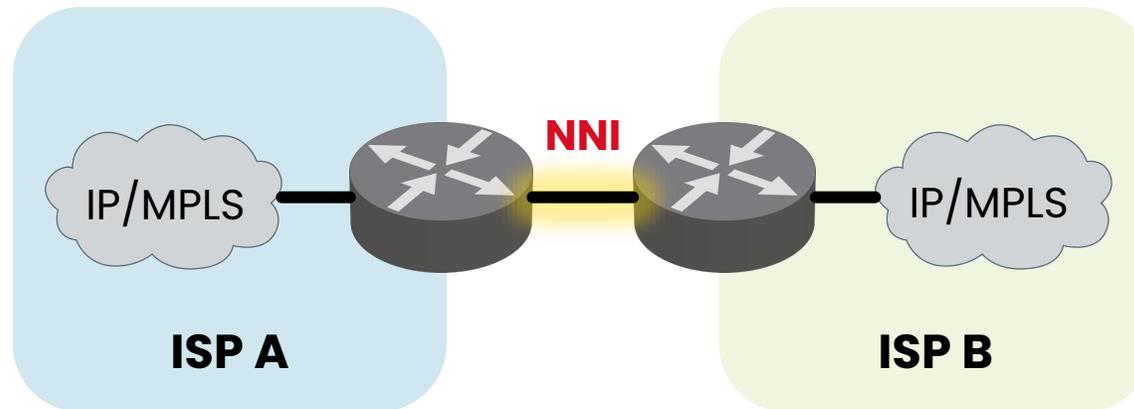
APPROACH 1: NNIs

APPROACH 1: NNIs
DAY1

- Many times as a first step **NNIs** are established between the two networks.

- **NNI** sounds fancy, but is **just a 802.1q tagged handover** at the end of the day.

- This allows the two ISPs to **establish services beyond the boundaries** of their respective networks.

- The **concept** of an NNI is usually **well-understood** both by **engineers** and existing **OSS/BSS** systems.
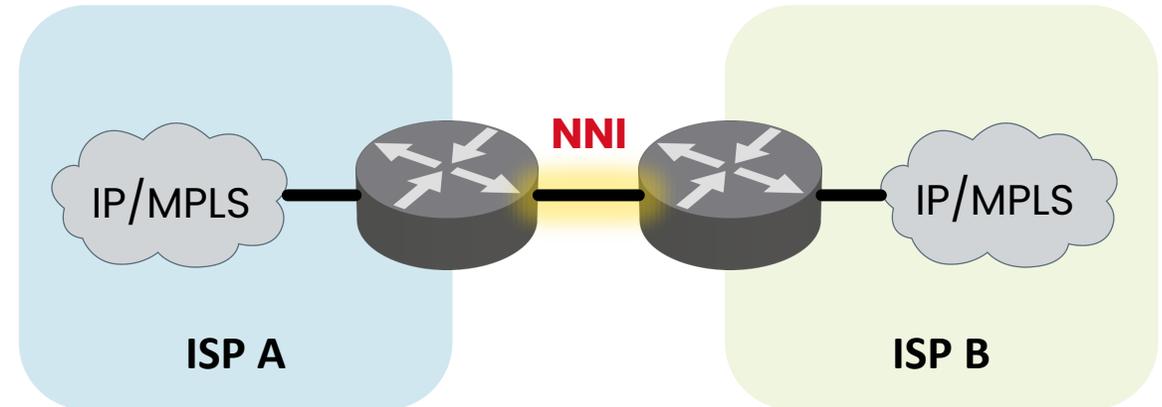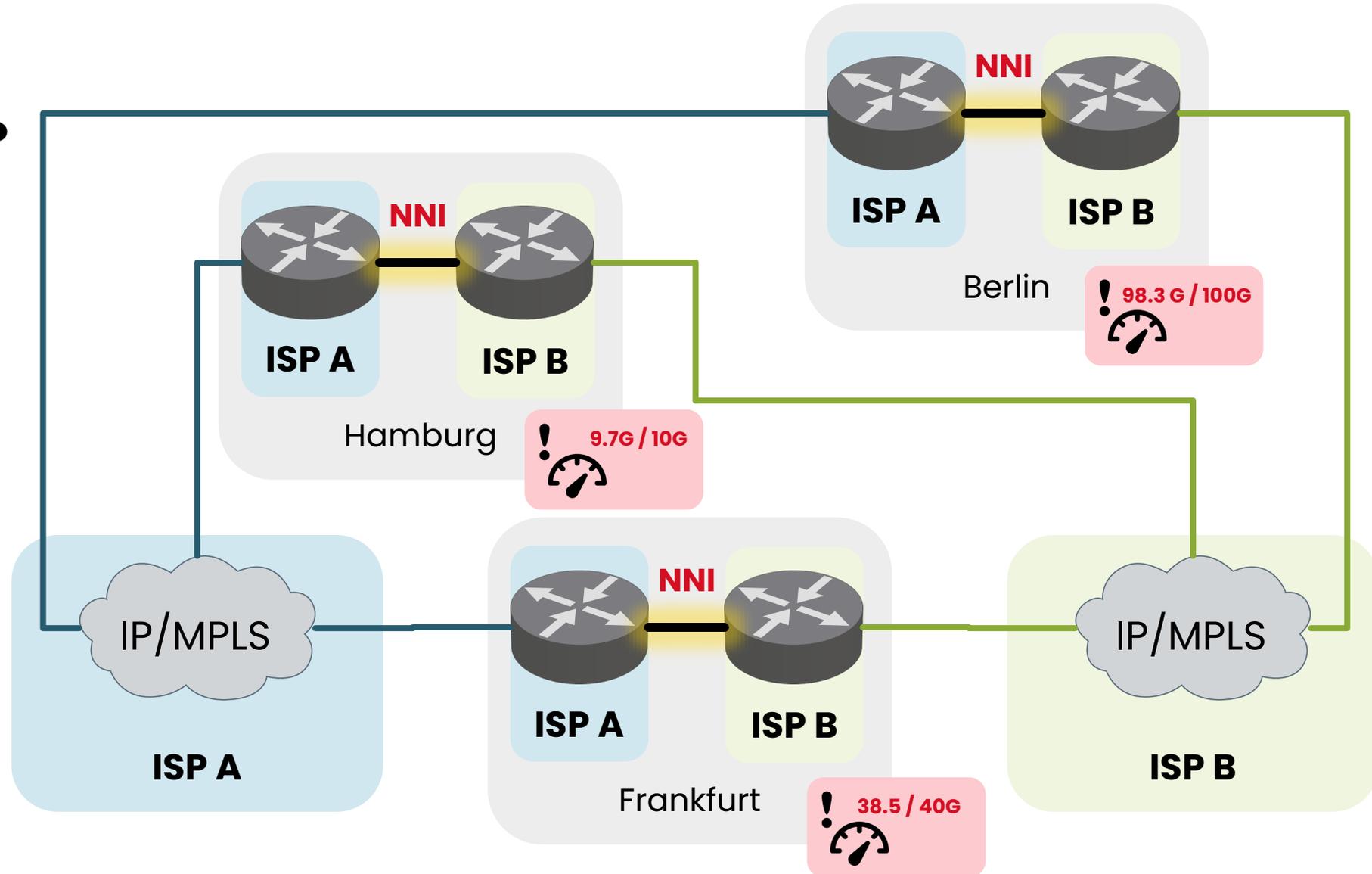
APPROACH 1: NNIs
DAY237

- Offers **no native redundancy** mechanism

- Traffic is always forced across a specific NNI, which could lead to **increased latency** or **traffic tromboning**

- **Capacity** of the links have to be monitored, as all **services share the NNI bandwidth**.

xantaro

NNI

ISP A ISP B

Berlin

98.3 G / 100G

NNI

ISP A ISP B

Hamburg

9.7G / 10G

IP/MPLS

ISP A

NNI

ISP A ISP B

Frankfurt

38.5 / 40G

IP/MPLS

ISP B

APPROACH 2: INTER-PROVIDER / INTER-AS
## **GENERAL IDEA**

- External BGP is mostly used with the **inet-unicast** and **inet6-unicast AFI/SAFI**

- As part of **Inter-Provider Setups**, external BGP sessions can also be used to signal information for various **MPLS Services across AS boundaries**.
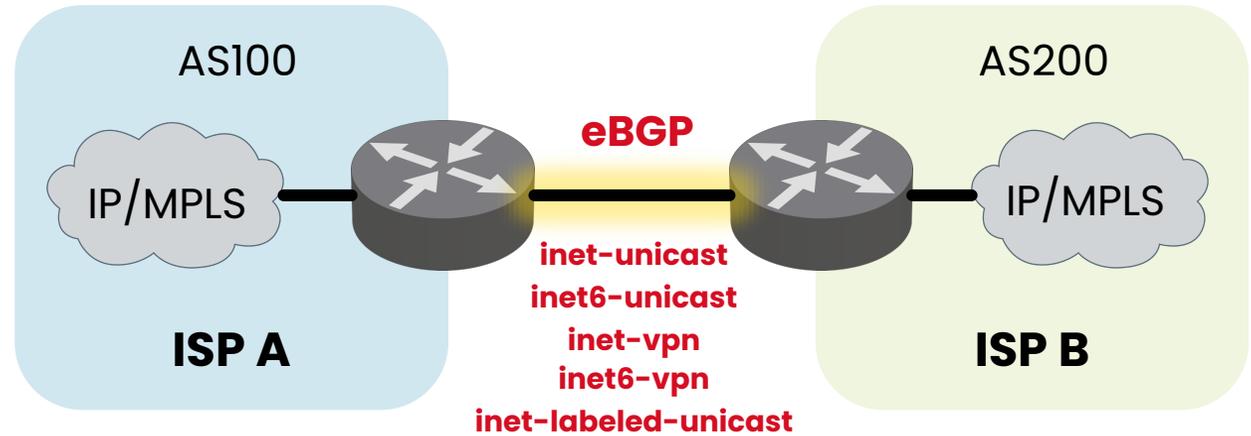
- **IETF** got us covered!

- **RFC4364** defines **three options** for this in ascending order of **maturity** and **scalability**:
  - **Option A**: One eBGP Session per Service (per instance)
  - **Option B**: One eBGP Session for all Services and Transport
  - **Option C**: One eBGP Session for all Services **+** One separate eBGP Session for Transport

AS100

AS200

eBGP

IP/MPLS

IP/MPLS

ISP A

ISP B

inet-unicast
inet6-unicast
inet-vpn
inet6-vpn
inet-labeled-unicast

WHAT OPTION SHOULD YOU CHOOSE?

- **All options** that we have **shown so far**, only get you **halfway there!**

- They are individual **quick fixes** at the end of the day:
  - Distinct networks will **continue to be distinct**
  - Both networks will continue to **live on their administrative islands** with controlled **ferry services** in between them
  - **Teams** will also **not merge**, since they are **busy doing things on their islands**
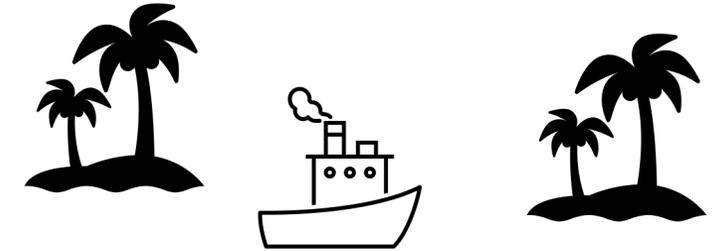
- **Real value add** can only be achieved through **merging both networks** and **building bridges** between them (permanent structures)
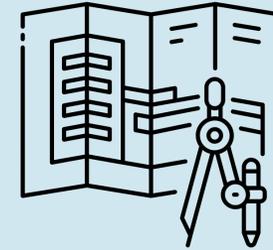
CoC VPN

NNI

Inter-Provider Option B

Inter-Provider Option A

Inter-Provider Option C
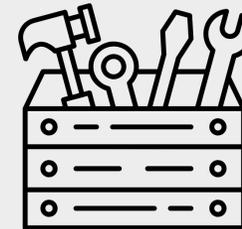
HOW DO WE BUILD THE BRIDGES?

xantaro

Blueprint

- **Seamless MPLS**

Toolbox

- **IGP-independent Label Distribution**: BGP-LU / BGP-CT / BGP-CAR
- **IGP-specific Label Distribution**: LDP, RSVP-TE, Segment Routing
- **Traffic-Engineering**: SRTE, RSVP-TE
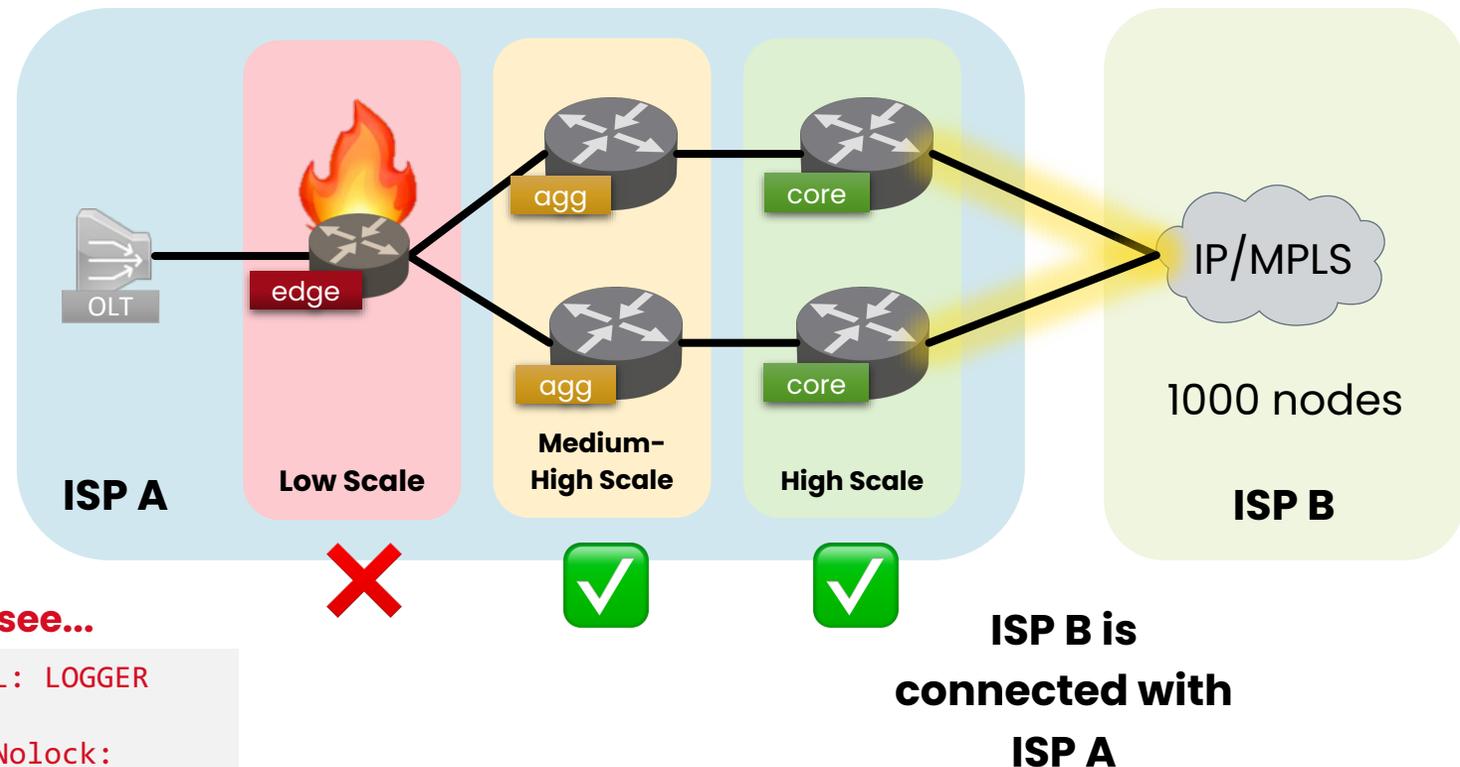- **IGP Design** (Merger, Split)
- **Static Routing**

WHAT IS SEAMLESS MPLS?

- While there is an IETF Draft on Seamless MPLS, it is **not a specific protocol**.

  - It is rather **a collection of design ideas** behind making it possible to **inter-connect different domains or sections** of your network, while maintaining **end-to-end** IP/MPLS connectivity.

  - Thus enabling the **possibility to scale** a network to **100,000 of nodes.**

  - One of its key elements to achieve this, is the **separation** of **service and transport** planes

- Seamless MPLS **uses a toolbox of existing protocols** to achieve the desired architectures.
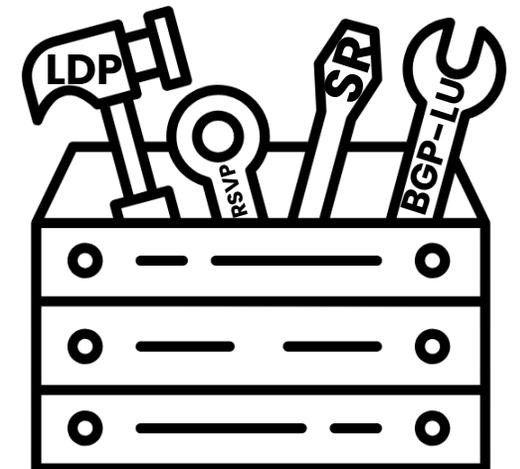
CHALLENGE 1:
ACCESS NETWORKS AND LOW-SCALE DEVICES (2)

xantaro

- Often SPs operate a **flat IGP** topology along with a a Label Distribution Protocol , because there is **simply no need** to **divide the network**.

- This is **completely fine**, until you **try to merge**!

- The **low scale devices** become the **weakest link**!

ISP A   **Low Scale**   **Medium-High Scale**   **High Scale**

IP/MPLS

1000 nodes

ISP B

ISP B is connected with ISP A

The log message you do not want to see...

```
584 1970/01/01 00:43:02.32 UTC CRITICAL: LOGGER
#2002 Base B:NPAPISYS:UNUSUAL_ERROR
"Slot B: npCreateL3TunnelNetwInterfaceNolock:
Error: bcmRc = Table full"
```

CHALLENGE 1:
ACCESS NETWORKS AND LOW-SCALE DEVICES (3)

xantaro

- These scenarios can usually be adressed by **separating** the **low-scale devices** from the rest of the network.

  - The separation can be achieved through **Level / Area concepts** within an IGP or **(preferred)** by using **different IGPs**

  - They will **not be part** of a **large** IGP domain.

  - They will connect to the rest of the network through **route summarization.**
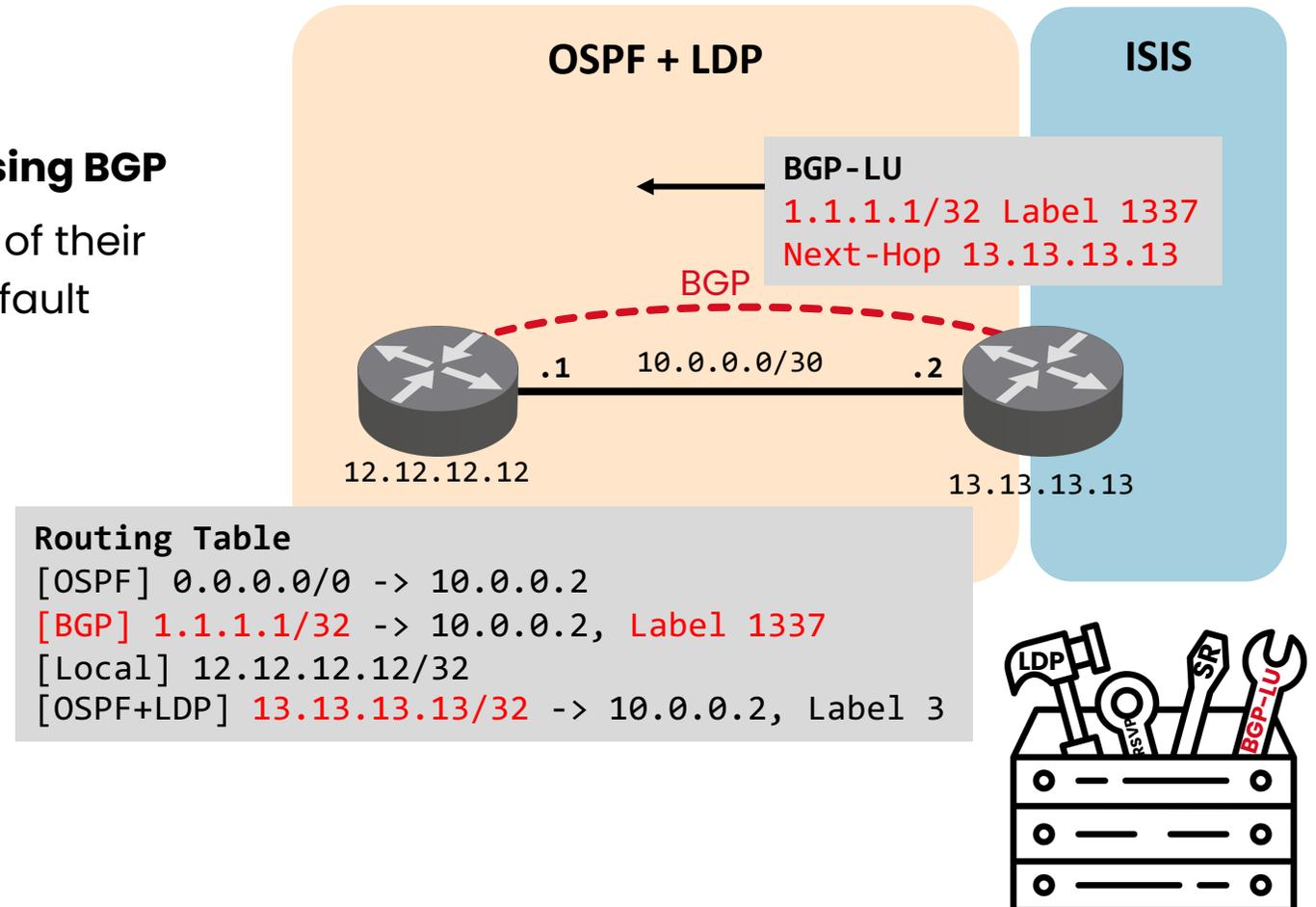
  - RIB / FIB size is drastically reduced!

- **There is a catch though:**

  - **Route summarization** and **MPLS** LSPs are **not best friends**!

  - Label Distribution usually **relies on exact match** /32 or /128 from IGP routes!
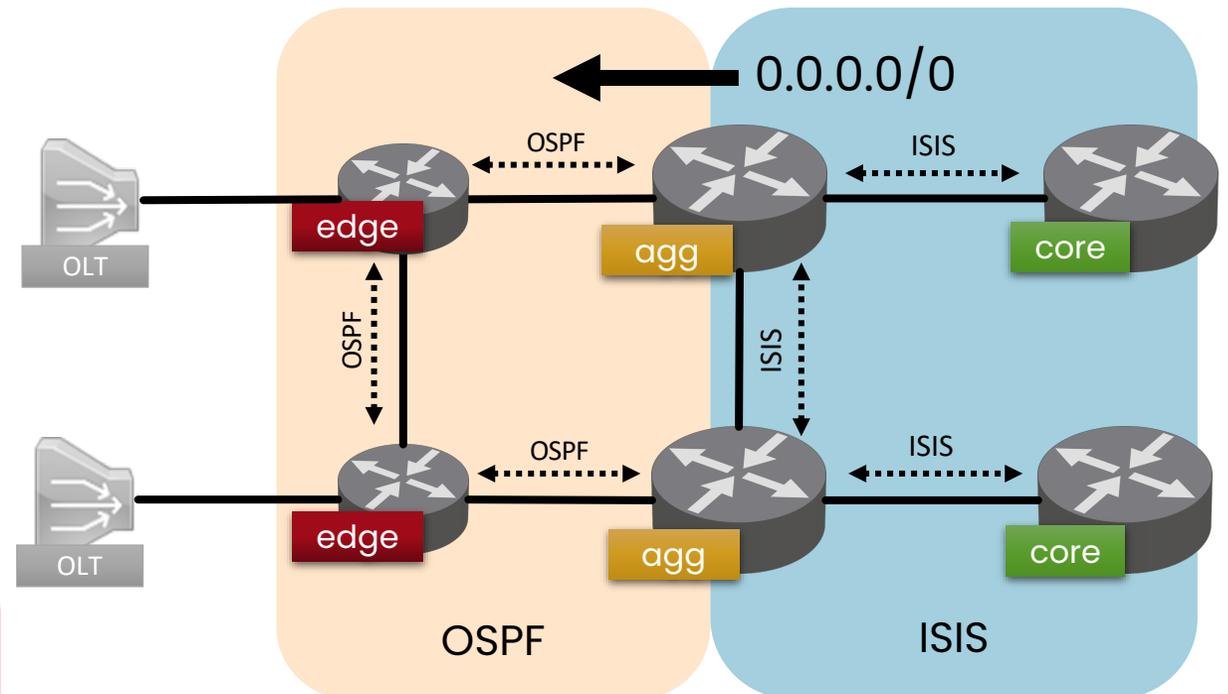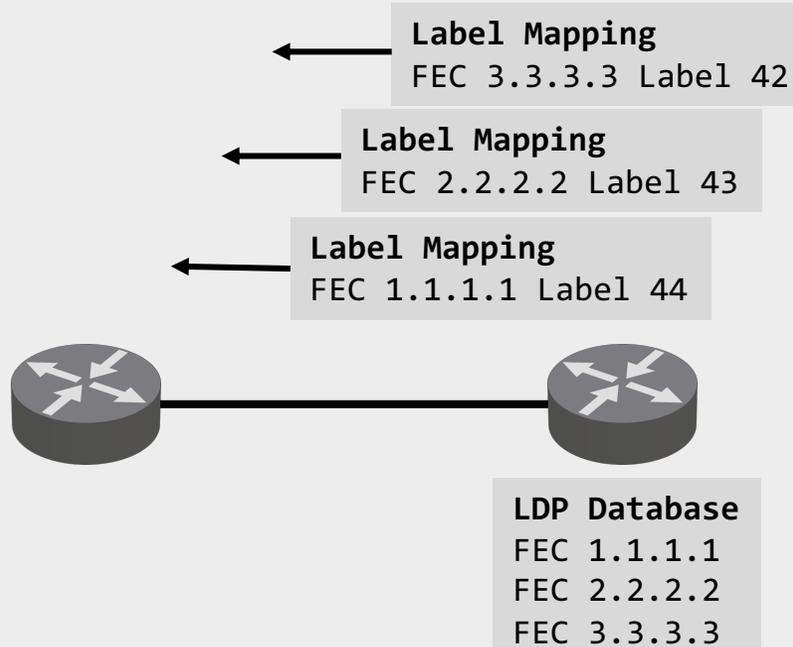
CHALLENGE 1:
ACCESS NETWORKS AND LOW-SCALE DEVICES (4)

- One solution for this can be **BGP Labeled Unicast (BGP-LU)**

  - It allows to distribute **Transport Labels using BGP**

  - The routes usually contain the **Next-Hop** of their **upstream router** and can work with a default route for NH resolution

    - **One next-hop for all** BGP-LU routes

    - IGP **redistribution** no longer necessary

    - BGP routes **can easily be filtered**, to reduce LIB/LFIB load

- Can optionally implemented **purely with BGP-LU** only (not relying on LDP and IGP at all)
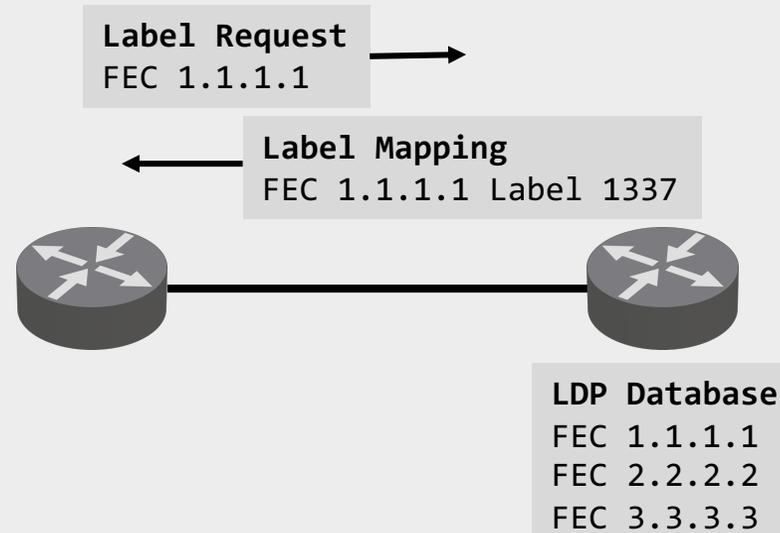
OSPF + LDP

ISIS

BGP-LU
1.1.1.1/32 Label 1337
Next-Hop 13.13.13.13

BGP

.1   10.0.0.0/30   .2

12.12.12.12

13.13.13.13

Routing Table
```
[OSPF] 0.0.0.0/0 -> 10.0.0.2
[BGP] 1.1.1.1/32 -> 10.0.0.2, Label 1337
[Local] 12.12.12.12/32
[OSPF+LDP] 13.13.13.13/32 -> 10.0.0.2, Label 3
```

CHALLENGE 1:
ACCESS NETWORKS AND LOW-SCALE DEVICES (5)

LDP Downstream Unsolicited (DU)

```
Label Mapping
FEC 3.3.3.3 Label 42

Label Mapping
FEC 2.2.2.2 Label 43

Label Mapping
FEC 1.1.1.1 Label 44
```

LDP Database
FEC 1.1.1.1
FEC 2.2.2.2
FEC 3.3.3.3

- Upstream router sends all FECs known to it, regardless of whether the downstream router actually needs them

LDP Downstream on Demand (DoD)

```
Label Request
FEC 1.1.1.1

Label Mapping
FEC 1.1.1.1 Label 1337
```

LDP Database
FEC 1.1.1.1
FEC 2.2.2.2
FEC 3.3.3.3

- Upstream router sends FECs only upon explicit request from the downstream router

- In any case LDP must **support** the **Longest Match Extension** to work with route summarization

- With some vendors, the **list of requested FECs** can be **auto-derived** from the configured services

- Similar to BGP-LU this approach can also be **used IGP-less** with a **static default route** if desired.
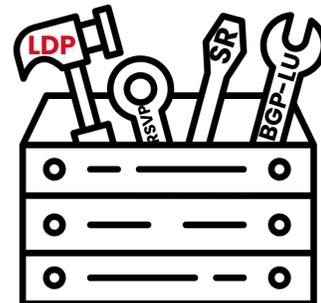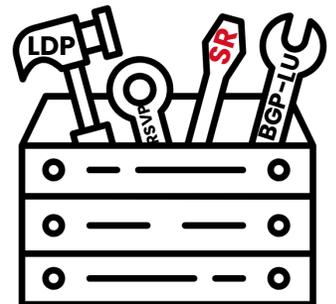
CHALLENGE 2:
BRINGING TOGETHER LARGE IGP DOMAINS (1)

- For many ISPs it is often **desirable to merge their IGPs** at some point
  - Unification of a „inner-core network"
  - Easier to maintain
  - Slimmer protocol stack
  - Network-wide BGP-LU support not required

- As a **prerequisite** some of the **tools we have introduced** can be used, to make sure that the network will scale with the additonal nodes.

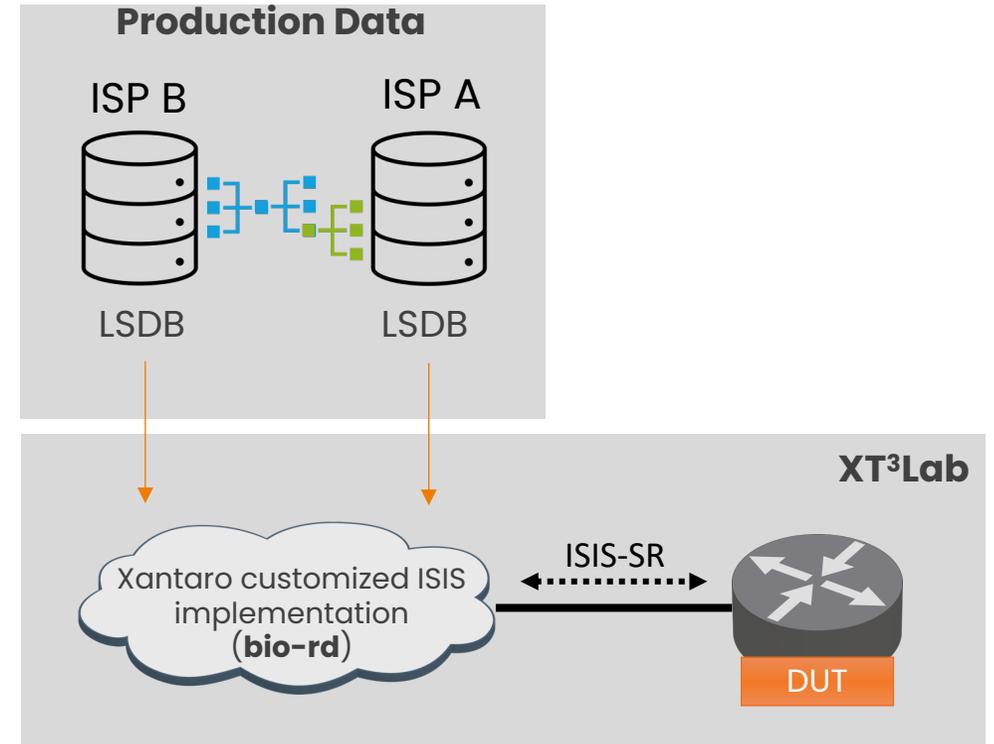- This merge can also already be done, when the Service Plane (the AS) has not been unified yet.

ISIS

ISIS

ISP A

ISP B

LDP

SR

RSVP

BGP-LU
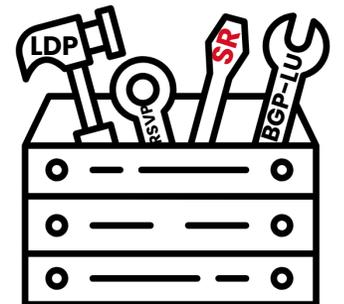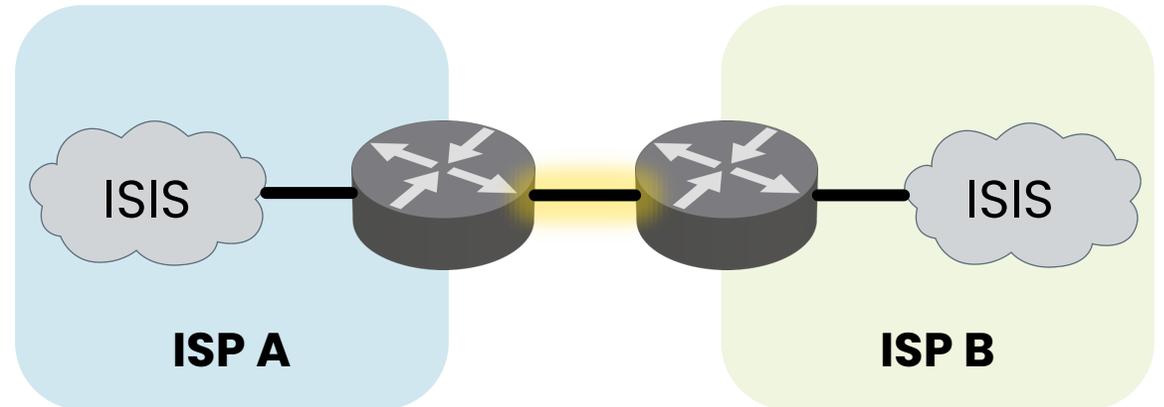
CHALLENGE 2:
BRINGING TOGETHER LARGE IGP DOMAINS (2)

- Comes with many challenges

  - **Renumbering** Duplicate Loopbacks, Transfer-Networks, Node-SIDs

  - Harmonizing **Metric** Values

  - Harmonizing IGP-specific **settings**

  - In case different IGPs are used, one network will have to **transition to a new IGP**

- Great opportunity to implement **Segment Routing (SR)**!

  - Fewer protocols for label distribution

  - Many **new features**: Flex Algo, TI-LFA, SR-TE

  - Our experience**: It is worth the change!**

    - Demands can change quickly

    - SR will have you covered with features like Flex Algo and SR-TE with changing demands

CHALLENGE 2:
# BRINGING TOGETHER LARGE IGP DOMAINS (3)

xantaro

- Before merging the ISIS-Domains of two networks, we can **simulate a merger of the networks** in our lab.

  - **Exported all ISIS Link-State-PDUs** (LSPs) from both networks

  - Manipulation of LSPs to **simulate a merger**

  - Transmission of LSDB to DUT

- As such a simulation is not feasible with commodity network testers, we **modified an existing software implementation of ISIS** to be able to transmit the LSDB towards a test device.

- We were **able to demonstrate** to our customer that the DUT **did not exhibit any undesired behavior** in a post-migration scenario.

Production Data

ISP B   ISP A

LSDB   LSDB

XT³Lab

Xantaro customized ISIS implementation (**bio-rd**)

ISIS-SR

DUT

```
352  func (l *lsdb) insertFake() {
353      l.lspsMu.Lock()
354      defer l.lspsMu.Unlock()
355      // Insert sample LSPs
356
357      dump := [][]byte{
358          /* Frame (398 bytes) */
359          {0x01, /* ........ */
360          0x67, 0x04, 0xa6, 0x17, 0x20, 0x16, 0x12, 0x80, /* g... ... */
361          0x21, 0x00, 0x00, 0x00, 0x02, 0x91, 0x4e, 0x79, /* !.....Ny */
362          0xa9, 0x03, 0x01, 0x04, 0x03, 0x49, 0x00, 0x01, /* .....I.. */
363          0x0e, 0x02, 0x05, 0xd4, 0x81, 0x02, 0xcc, 0x8e, /* ........ */
364          0x86, 0x04, 0xac, 0x10, 0x80, 0x15, 0x84, 0x04, /* ........ */
```

AND NOW IT'S UP TO YOU!
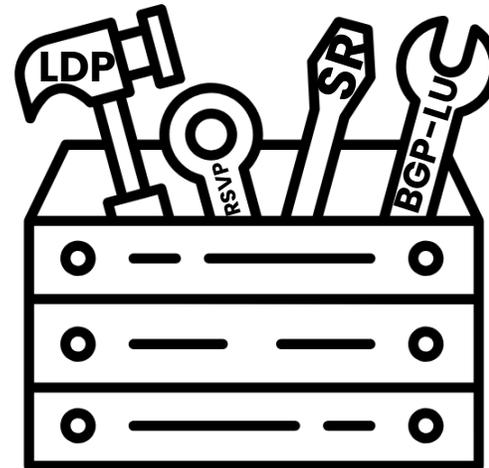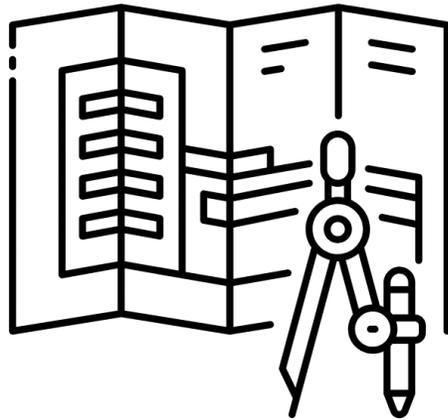
THANK YOU!

References

Rosen and Rekhter. *BGP/MPLS IP Virtual Private Networks (VPNs.) RFC4364.*
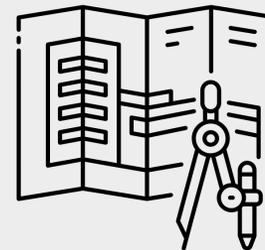Leyman, Decraene, Filsfils, Konstantynowicz and Steinberg. *Seamless MPLS Architecture. draft-ietf-mpls-seamless-mpls-07.*