A Carrier-Grade L2BSA Gateway with Netgraph

Lutz Donnerhacke

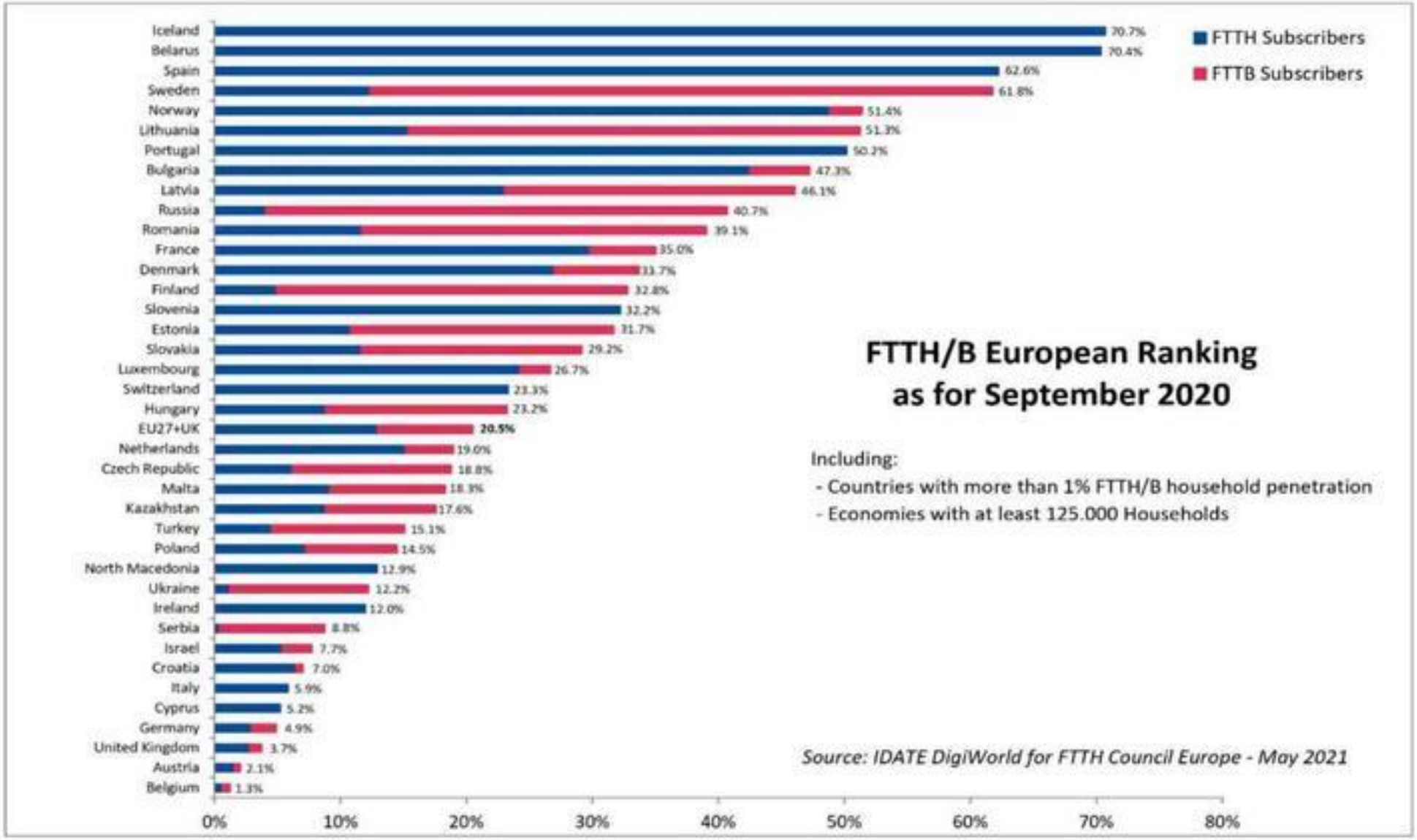
IKS Service GmbH

Personal background

- Grown up with a C64 in Eastern Germany
- Short excurse to MSDOS and Win3.11
- Switched to Linux in 1992
- Individual Network e.V. in Germany
  - grassroot Internet for ordinary people
- Build regional ISP in 1996
  - infrastructure provider of two regional Broadband-ISPs in Thuringia
- Switch to FreeBSD in 2013
  - Replace commercial LNS by netgraph/MPD

Broadband in Germany

- 1983 – CDU (German conservatives) choose copper over fiber
  - Probably due to corruption via the minister Schwarz-Schilling
- 1995 – CDU privatization created Deutsche Telekom
  - Got the whole cable network as gift, hence a (copper) monopoly
- 1998 – Unbundling of last mile to competitors
- 2009 – Vectoring for increased bandwith over copper (DSL)
  - All lines need to be handled by a single DLSAM
  - Unbundling not longer possible
  - To break the local monopoly, regulated bitstream access was introduced
- We are still a copper DSL based country!
  - Fiber is rolled out 2022 but expected to be available not before 2024


FTTH Subscribers
FTTB Subscribers
Iceland 70.7%
Belarus 70.4%
Spain 62.6%
Sweden 61.8%
Norway 51.4%
Lithuania 51.3%
Portugal 50.2%
Bulgaria 47.3%
Latvia 46.1%
Russia 40.7%
Romania 39.1%
France 35.0%
Denmark 33.7%
Finland 32.8%
Slovenia 32.2%
Estonia 31.7%
Slovakia 29.2%
Luxembourg 26.7%
Switzerland 23.3%
Hungary 23.2%
EU27+UK 20.5%
Netherlands 19.0%
Czech Republic 18.8%
Malta 18.3%
Kazakhstan 17.6%
Turkey 15.1%
Poland 14.5%
North Macedonia 12.9%
Ukraine 12.2%
Ireland 12.0%
Serbia 8.8%
Israel 7.7%
Croatia 7.0%
Italy 5.9%
Cyprus 5.2%
Germany 4.9%
United Kingdom 3.7%
Austria 2.1%
Belgium 1.3%
0%
10%
20%
30%
40%
50%
60%
70%
80%
FTTH/B European Ranking
as for September 2020
Including:
- Countries with more than 1% FTTH/B household penetration
- Economies with at least 125.000 Households
Source: IDATE DigiWorld for FTTH Council Europe - May 2021

Layer2 Bitstream Access (regulated)

- Solution for the layer1 monopoly created by vectoring
  - DSL port in sync: Hand over all layer2 traffic to a different ISP
  - Requires a transport layer which maps 1:1 to the port
  - No opportunity for smart networks (like Multicast distribution)
  - Problem: How to handle bandwidth discrepancy?
- Technical approach (A10NSP)
  - Add a VLAN tag to all traffic from the port (IEEE 802.1ad)
    Can only handle up to ~4000 customers per aggregation line
  - Add current line speed in DHCP option or PPP LCP option
    Shape downstream traffic to signaled speed (or pay fines)

Commercial aspects

- Numbers
  - 1 Mio households in Thuringia => ~260 interconnection points
  - About 50 interconnections needed for DTAG operated areas
  - Most rural regions are operated by us (not DTAG)
- You need a line to each area, where your customer is
  - pay for a 1G/10G WAN + interconnection fee
  - get a couple of customers switching to you
- We need a very cheap termination device
  - Commercially available: Large PPPoE appliances
  - We like to have DHCP and IPv6 in multiple VLANs (i.e. separate VoIP)

What happens on wire?

Customer
Payload (IP)
C-VLAN Tag
Source-MAC
Destination-MAC

A10NSP Interconnection
Payload (IP)
C-VLAN Tag
S-VLAN Tag
Source-MAC
Destination-MAC

What really happens on wire

Customer
Payload (IP)
140
01:02:03:04:05:06
01:02:03:07:08:09

A10NSP Interconnection
Payload (IP)
140
27
01:02:03:04:05:06
01:02:03:07:08:09

What really happens on wire?

Customer
Payload (IP)
140
01:02:03:04:05:06
01:02:03:07:08:09

A10NSP Interconnection
Payload (IP)
140
35
01:02:03:04:05:06
01:02:03:07:08:09

What really happens on wire!

Customer
Payload (IP)
140
01:02:03:04:05:06
01:02:03:07:08:09

A10NSP Interconnection
Payload (IP)
140
35
01:02:03:04:05:06
01:02:03:07:08:09

- On every DSL resync
  - A *new, dynamic* S-VLAN is used
  - Traffic send to the old S-VLAN is lost immediately
- PPP will drop the line and resync
- DHCP does not even notice

Simplify by adding complexity

static

Customer	A10NSP	Interface	Rotate	Service	Interface	*Line*
Payload	Payload	Payload	Payload	Payload	Payload	Payload
C-Vlan	C-Vlan	C-Vlan	*S-Vlan*	*S-Vlan*	*S-Vlan*	Src-MAC
Src-MAC	*S-Vlan*	*S-Vlan*	A-Vlan	A-Vlan	Src-MAC	Dst-MAC
Dst-MAC	Src-MAC	A-Vlan	C-Vlan	Src-MAC	Dst-MAC	
	Dst-MAC	Src-MAC	Src-MAC	Dst-MAC		
		Dst-MAC	Dst-MAC			

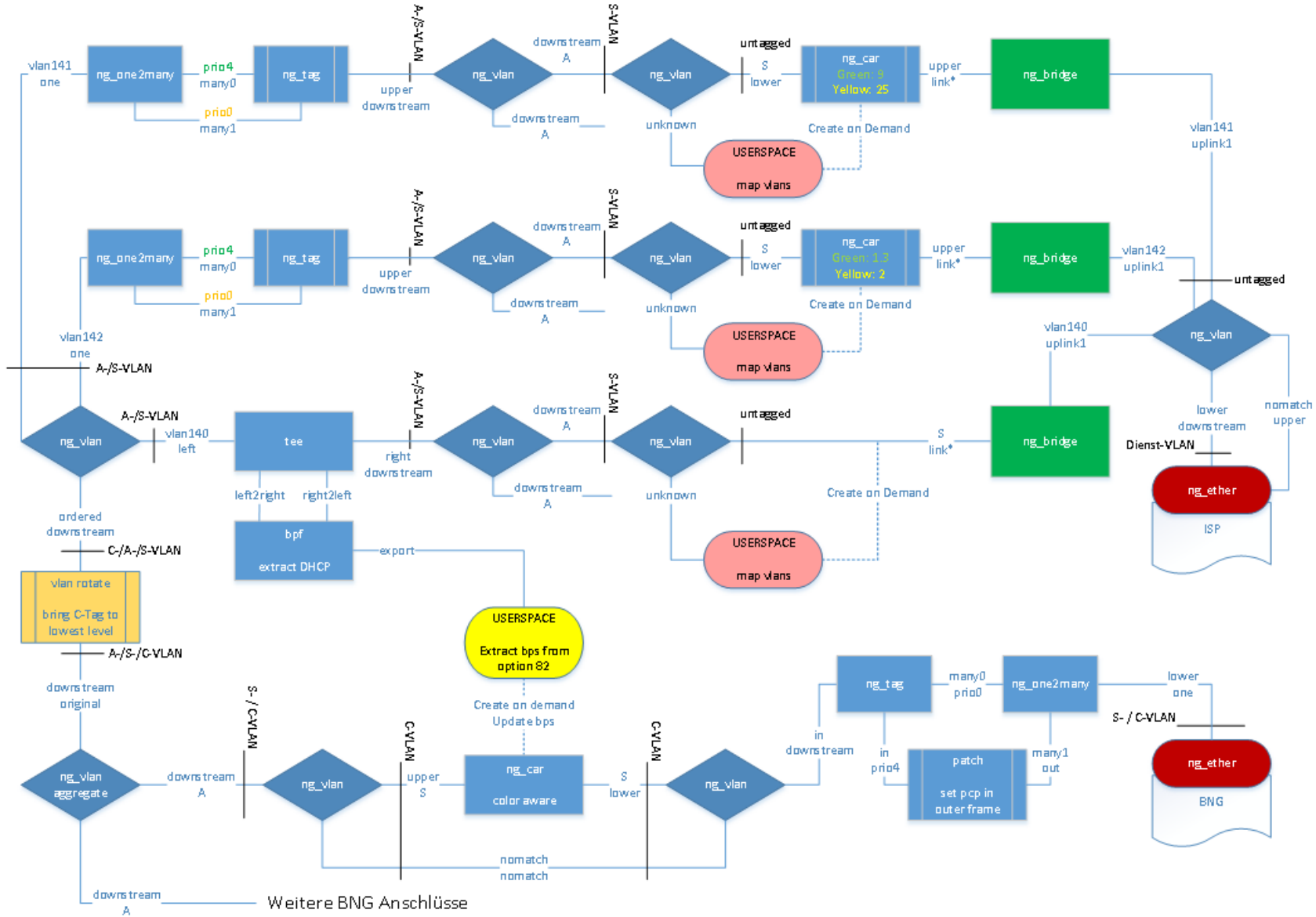
QoS per interface

QoS per service

Customer traffic handled per service

Todo list for netgraph

- Blue/Red: Existing nodes
- Green: Insufficient node
- Orange: Missing node
- Yellow/Rose: Userspace scripting



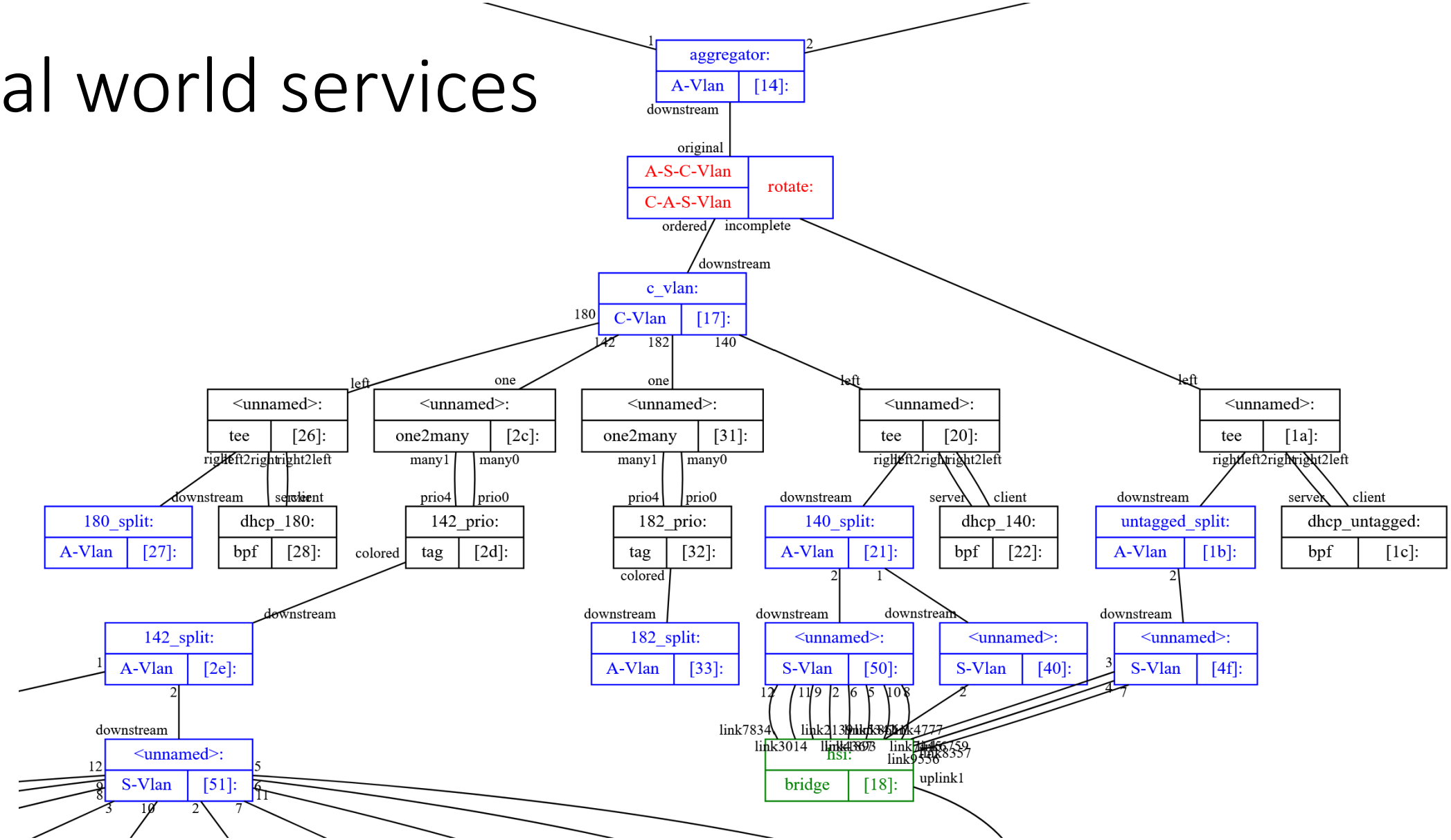
Necessary modifications to FreeBSD

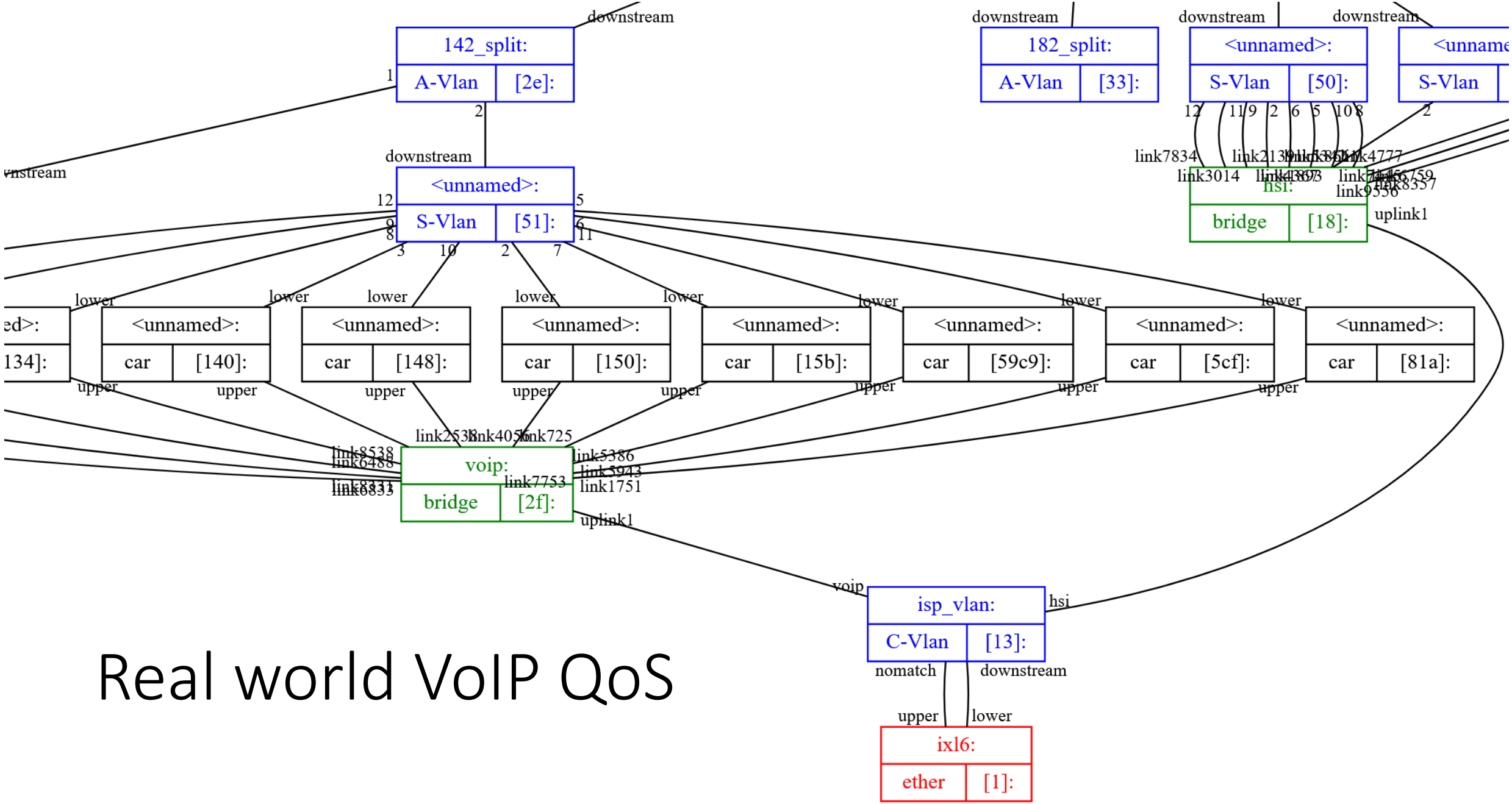
- Mandatory
  - D21846: Add 802.1Q ethertypes used in the wild
  - D24179: ixl: Permit 802.1ad frames to pass though the chip
  - D22076: netgraph/ng_vlan_rotate: IEEE 802.1ad VLAN manipulation netgraph node type
- Bridging
  - D21803: netgraph/ng_bridge: Replace NG_BRIDGE_MAX_LINKS with unlimited links
  - D23840: netgraph/ng_base: Allow larger BINARY2ASCII conversions
  - D23963: netgraph/ng_bridge: Introduce "uplink" ports without MAC learning
  - D28123: netgraph/ng_bridge: become multithreaded
- QoS
  - D22140: netgraph/ng_tag: Variable length data can not be set for all length
  - D22110: netgraph/ng_car: Add color marking code
- Debugging
  - D21968: netgraph/ng_source: Allow ng_source to inject into any netgraph network
  - D21965: usr.sbin/ngctl: Generate more compact GraphWiz output

Real world problems

- Same MAC in different VLANs
  - Tagged and untagged from the same customer device (CPE)
  - Overlapping MACs from different CPEs due to MAC increase per VLAN
- Different CPE configuration
  - Same service may come in different C-Vlans (or untagged)
  - Need a decision mechanism to choose the right one
- Erroneous transmission of data rates
  - No DHCP exchange due to signaled 0 bps (enforced!)
- Broken interface cards (ixl) do not transmit 802.1ad
  - Customer untagged is an IEEE violation, but works

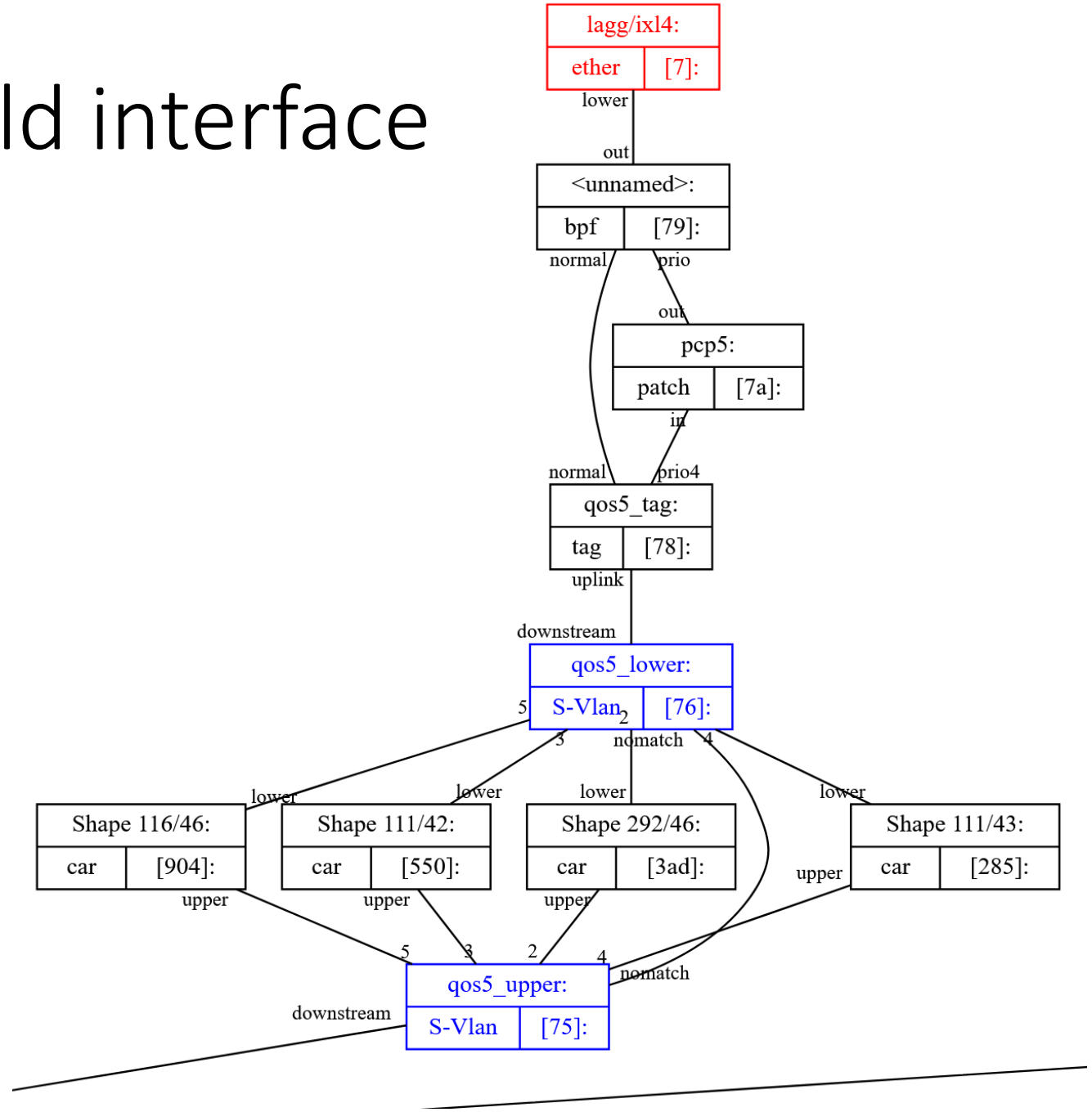
Real world services





Real world VoIP QoS

Real world interface



Real world bridge table

MAC	Link	Line	S-Vlan	C-Vlan
08:96:d7:46:fa:ec	hsi:link61	ixl0	2	140
08:96:d7:46:fa:ed	voip:link4249	ixl0	2	142
2c:91:ab:bb:4b:76	hsi:link7554	ixl10	10	140
2c:91:ab:bb:4b:77	voip:link3536	ixl10	10	142
3c:a6:2f:fe:ac:03	voip:link2299	ixl10	18	142
3c:a6:2f:fe:ac:02	hsi:link3507	ixl10	18	untagged
2c:3a:fd:dd:39:2a	hsi:link3588	ixl1	49	140
2c:3a:fd:dd:39:2b	voip:link5014	ixl1	49	142
00:01:21:01:44:2d	hsi:link9805	ixl8	5	180
00:01:21:01:44:2d	voip:link4426	ixl8	5	182

Questions?

https://lutz.donnerhacke.de/Blog

https://www.thyotec.de/

https://www.netkom.de/Privatkunden/Internet