



Use cases and pitfalls in MPLS/VPLS networks



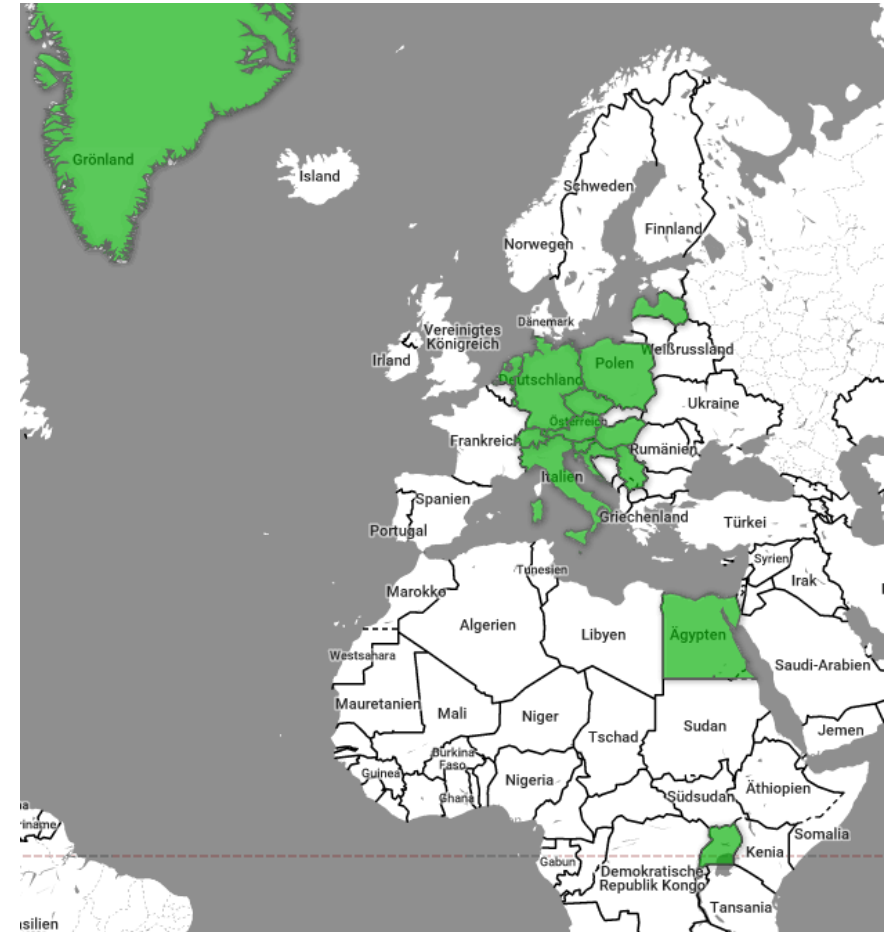
Contact

- Phone: +49 761 2926500
- Email: sales@fmsweb.de
- Shop: <https://www.mikrotik-shop.de>
- MikroTik Mirror: <http://www.mikrotik-software.de>
- Twitter: https://twitter.com/fmsweb_de
- Website: <http://www.fmsweb.de>
- Wiki: <http://wiki.fmsweb.de>
- Presentations: <http://wiki.fmsweb.de/wiki/MUM-Presentations>
- Facebook: <https://www.facebook.com/fmsinternetservice>



About me

- Sebastian Inacker
<inacker@fmsweb.de>
- FMS Internetservice GmbH,
Germany
- MikroTik Trainer TR0011
(May 2007)
- MTCNA, MTCRE, MTCTCE,
MTCUME, MTCWE,
MTCIPv6E, MTCINE





MikroTik trainings and workshops

- Own training center and on site
(Austria, Germany, Greenland, Hungary,
Luxembourg, Malta, Netherlands,
Switzerland, Uganda)
- Inquiries: sales@fmsweb.de





Overview / big picture

“Implementing and running a MPLS/VPLS network is easy.
As long as it is running well.”

Topics:

- Typical use cases of our (ISP) customers
- Typical pitfalls
- Surprising pitfalls
- Real world examples



Overview / big picture

“Implementing and running a MPLS/VPLS network is easy.
As long as it is running well.”

Not *main* topics

- Step-by-step guide for *each* setup (focus on pitfalls)
- Reason for MPLS/VPLS (You *should* know, why)



Reasons for MPLS/VPLS

Ok, very short and incomplete...

Benefits of MPLS

- Routing more complex than MPLS
- Some future setups (L3 VPN, TE) require MPLS

Benefits of VPLS vs. EoIP

- VPLS: No fragmentation (if done right)
- EoIP: Big overhead (42 bytes) & might cause fragmentation



Overview / big picture

Pitfalls:

- Incomplete (of course)
- Not limited to MPLS/VPLS

Needs for VPLS

- MPLS
- Routing (OSPF here)
- Physical infrastructure



Warning / heads-up / caution

This presentation will include errors, mistakes and wrong configuration attempts *to show resulting errors!*

Examples are simplified.

Keep that in mind.



The beginning

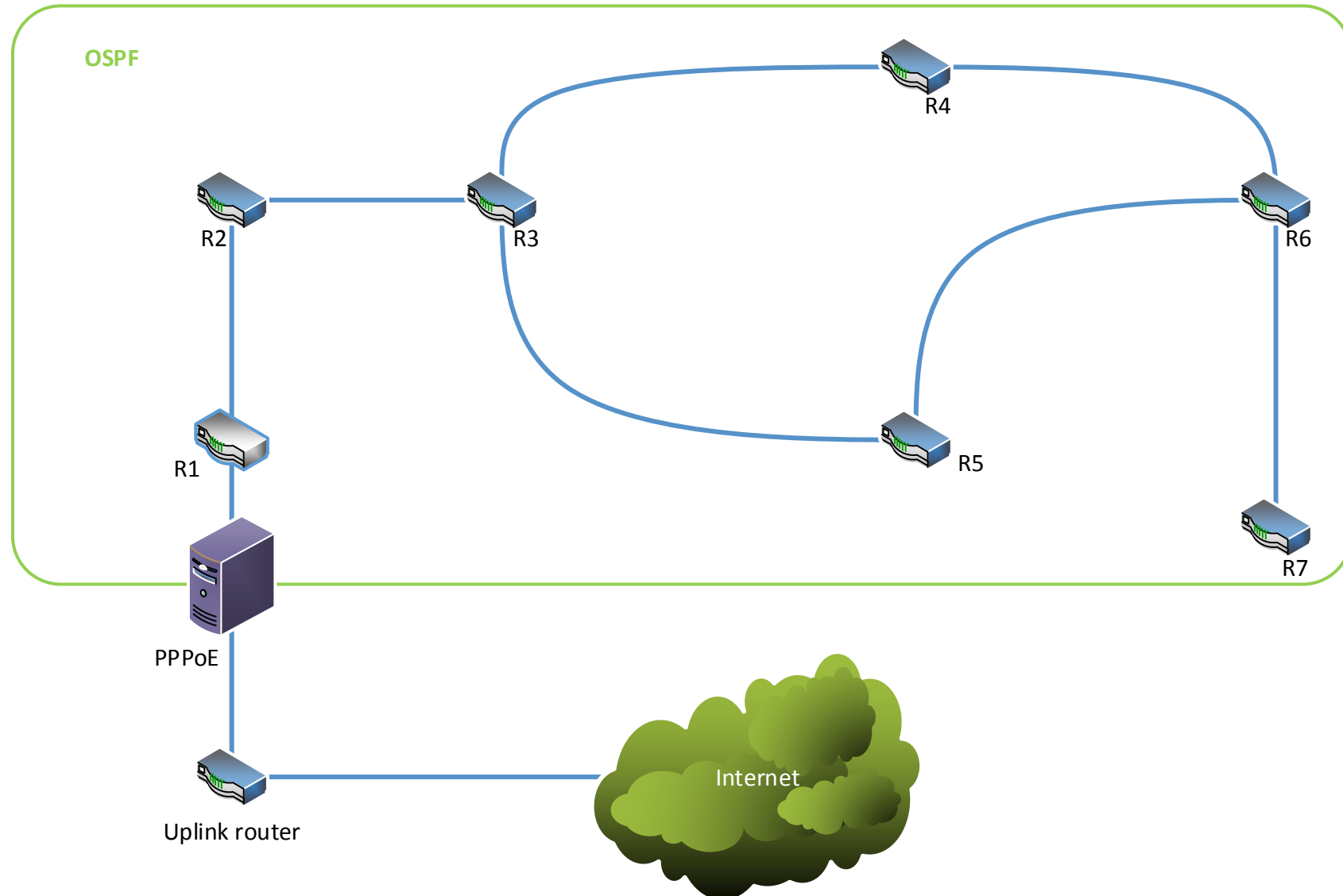


Existing setup

- Existing OSPF network
- One PPPoE server
- EoIP (L2 tunnel) for client connections

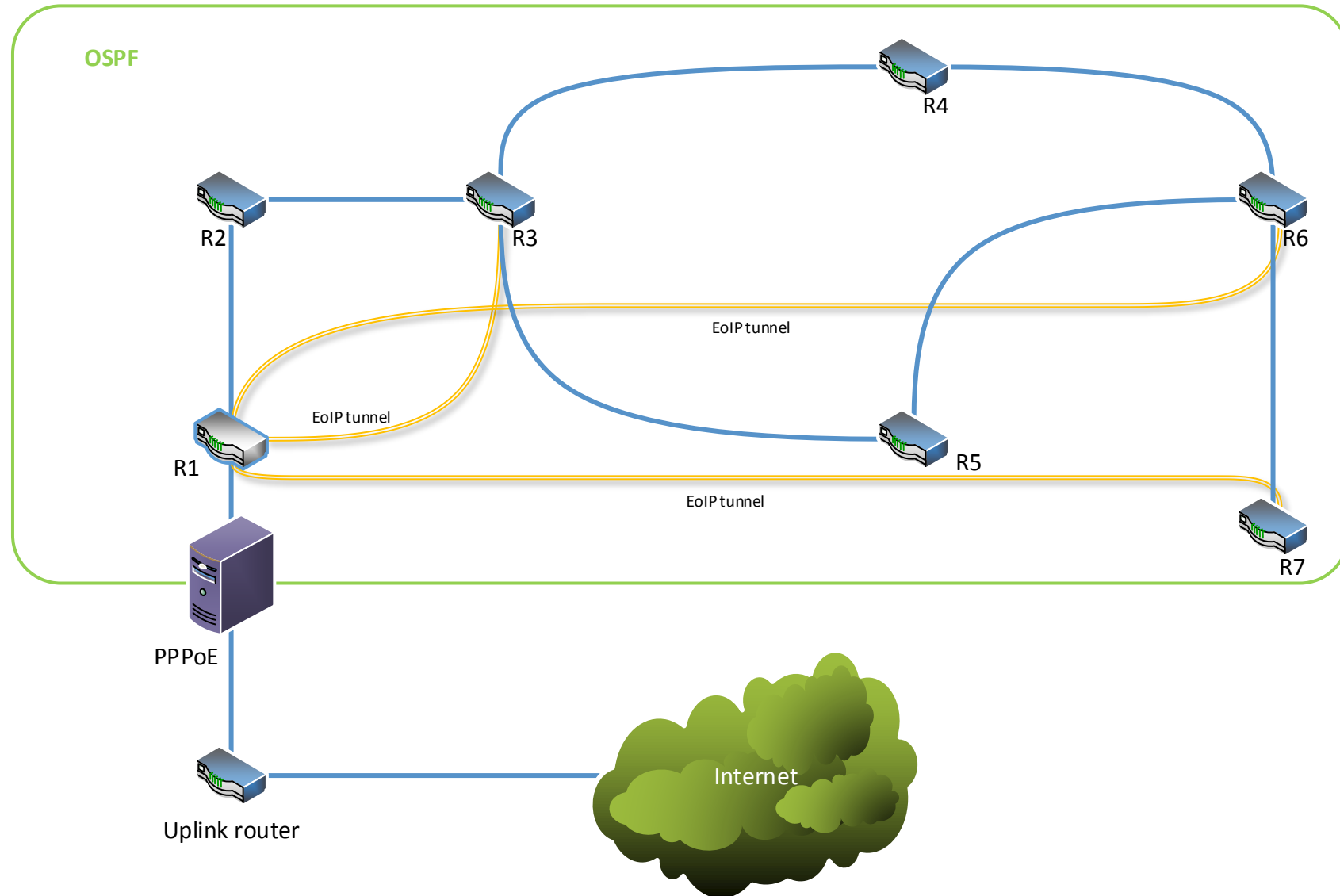


Existing setup





Existing setup





Requirements for MPLS

MPLS can be integrated without service disruption

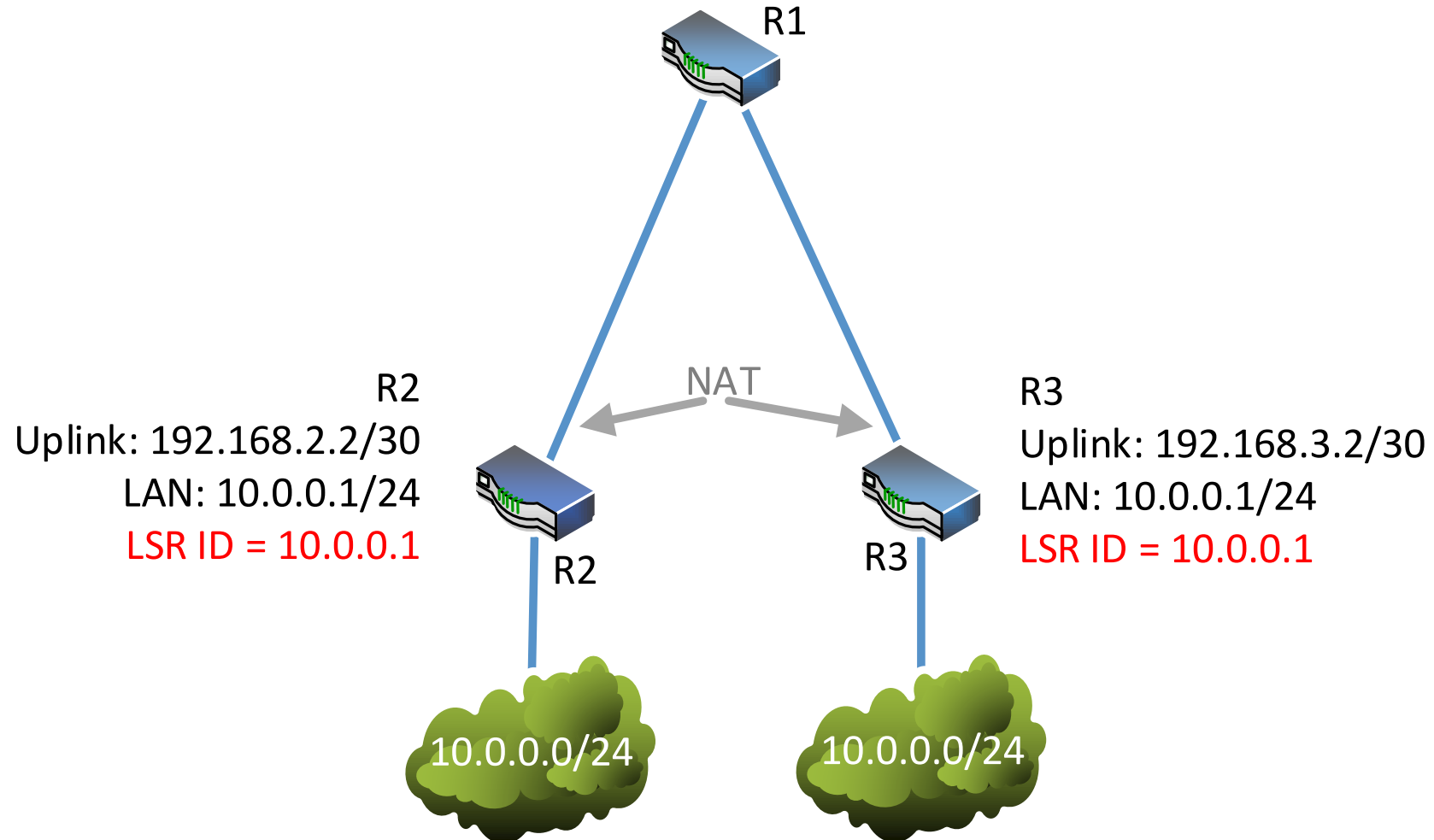
Running MPLS on top of OSPF:

- Enable LDP (Label Distribution Protocol)
- Set
 - LSR ID (Label Switching Router's ID)
 - Transport Address

If left unset: Lowest IP of router will be used
- Create LDP interfaces



1st possible issue: LSR ID not unique





Unique IP for LDP

Unique IP for

- LDP (LSR ID and Transport Address)

Let's *try* 10.255.255.<Router>/32 on physical interface

Address	Network	Interface
10.14.2.3/28	10.14.2.0	ether3
10.14.31.3/28	10.14.31.0	ether4
10.14.32.3/28	10.14.32.0	ether5
10.255.255.3	10.255.255.3	ether4

4 items

Enabled

LSR ID: 10.255.255.3

Transport Address: 10.255.255.3

Path Vector Limit: 255

Hop Limit: 255

Loop Detect

Use Explicit Null

Distribute For Default Route

OK Cancel Apply



Unique IP for OSPF

Unique IP for

- OSPF (Router ID) – same issue as with LSR ID

Take care: Setting of Router ID

- Restart of OSPF
 - Loss of routing table
- Service affecting action!

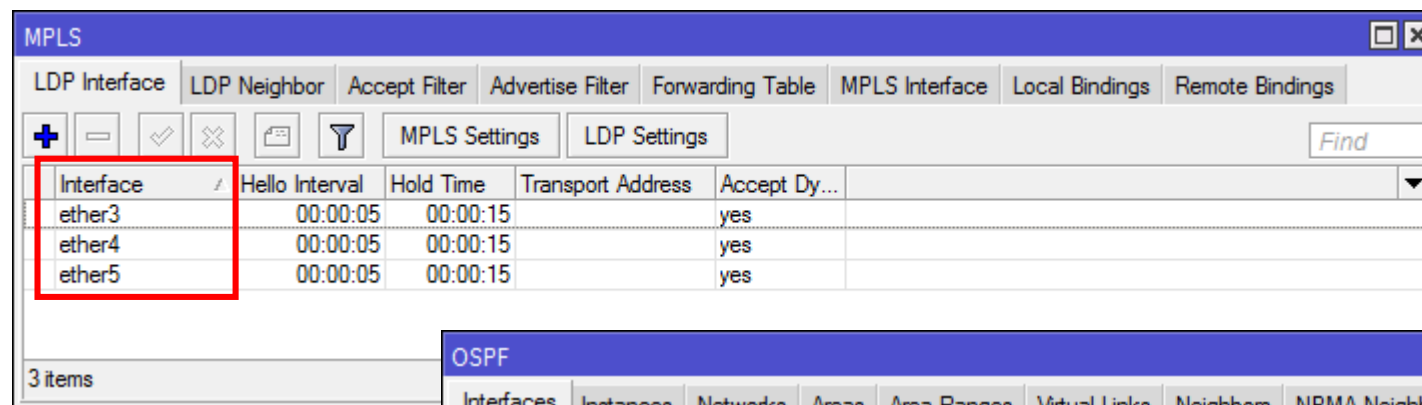
The screenshot shows the 'OSPF Instance <default>' configuration window. The 'General' tab is selected. The 'Name' field is set to 'default'. The 'Router ID' field is set to '10.255.255.3' and is highlighted with a red border. Below this, there are several dropdown menus for redistribution settings: 'Redistribute Default Route' (never), 'Redistribute Connected Routes' (no), 'Redistribute Static Routes' (no), 'Redistribute RIP Routes' (no), 'Redistribute BGP Routes' (no), and 'Redistribute Other OSPF Routes' (no). At the bottom, there are 'In Filter' (ospf-in) and 'Out Filter' (ospf-out) dropdown menus. The window has 'OK', 'Cancel', 'Apply', 'Disable', 'Comment', 'Copy', and 'Remove' buttons on the right side. The status bar at the bottom shows 'enabled' and 'default'.



LDP interfaces

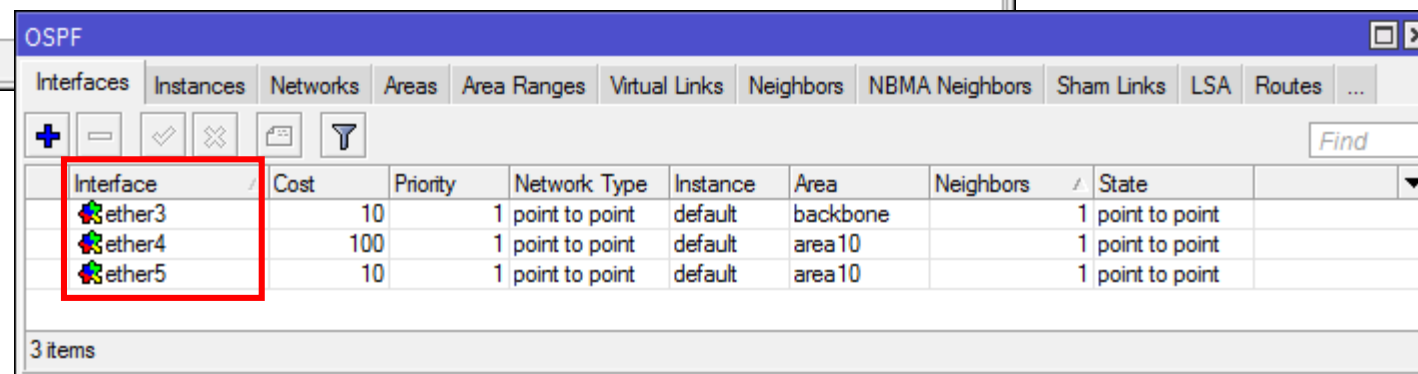
Set LDP interfaces

- Don't forget your backup path!
- Compare OSPF interfaces and LDP interfaces



MPLS LDP Interface configuration window showing a table of interfaces. The table has columns: Interface, Hello Interval, Hold Time, Transport Address, and Accept Dy... The rows are ether3, ether4, and ether5. All Hello Intervals are 00:00:05 and all Hold Times are 00:00:15. The 'Accept Dy...' column shows 'yes' for all three interfaces. A red box highlights the first three rows of the table.

Interface	Hello Interval	Hold Time	Transport Address	Accept Dy...
ether3	00:00:05	00:00:15		yes
ether4	00:00:05	00:00:15		yes
ether5	00:00:05	00:00:15		yes



OSPF Interfaces configuration window showing a table of interfaces. The table has columns: Interface, Cost, Priority, Network Type, Instance, Area, Neighbors, and State. The rows are ether3, ether4, and ether5. ether3 has a cost of 10, priority of 1, and network type of point to point. ether4 has a cost of 100, priority of 1, and network type of point to point. ether5 has a cost of 10, priority of 1, and network type of point to point. A red box highlights the first three rows of the table.

Interface	Cost	Priority	Network Type	Instance	Area	Neighbors	State
ether3	10	1	point to point	default	backbone		1 point to point
ether4	100	1	point to point	default	area10		1 point to point
ether5	10	1	point to point	default	area10		1 point to point



Create VPLS tunnels

Interface <vpls-to-r3>

General Status Traffic

Name: vpls-to-r3

Type: VPLS

MTU: 1500

Actual MTU: 1500

L2 MTU: 1500

MAC Address: 02:DF:35:57:11:55

ARP: enabled

ARP Timeout:

Remote Peer: 10.255.255.3

VPLS ID: 1:3

Cisco Style

Cisco Style ID: 0

Advertised L2MTU: 1500

PW Type: tagged ethernet raw ethernet

OK Cancel Apply Disable Comment Copy Remove Torch

enabled running slave BGP signaled Cisco BGP Sig...

Interface <vpls-to-r1>

General Status Traffic

Name: vpls-to-r1

Type: VPLS

MTU: 1500

Actual MTU: 1500

L2 MTU: 1500

MAC Address: 02:EA:AF:6B:4D:41

ARP: enabled

ARP Timeout:

Remote Peer: 10.255.255.1

VPLS ID: 1:3

Cisco Style

Cisco Style ID: 0

Advertised L2MTU: 1500

PW Type: tagged ethernet raw ethernet

OK Cancel Apply Disable Comment Copy Remove Torch

enabled running slave BGP signaled Cisco BGP Sig...



Check VPLS interface

Interface <vpls-to-r1>

General Status Traffic

Last Link Down Time:

Last Link Up Time:

Link Downs:

Remote Label:

Local Label:

Remote Status:

Remote Group:

Transport:

Transport Nexthop:

Imposed Labels:

OK
Cancel
Apply
Disable
Comment
Copy
Remove
Torch

enabled **running** slave BGP signaled Cisco BGP Sig...

VPLS interface not running!

```
Terminal
[admin@S:R03] > /interface vpls print
Flags: X - disabled, R - running, D - dynamic,
B - bgp-signaled, C - cisco-bgp-signaled
0  name="vpls-to-r1" mtu=1500 l2mtu=1500 mac-address=02:EA:AF:6B:4D:41
   arp-enabled arp-timeout=auto disable-running-check=no
   remote-peer=10.255.255.1 vpls-id=1:3 cisco-style=no cisco-style-id=0
   advertised-l2mtu=1500 pw-type=raw-ethernet use-control-word=default
[admin@S:R03] >
```



Check MPLS

Empty:

- MPLS Local Bindings
- MPLS Remote Bindings
- MPLS Forwarding Table

MPLS Local Bindings window showing an empty table with columns: Dst. Address, Label, Advertised Path, Peers.

Dst. Address	Label	Advertised Path	Peers
0 items			

MPLS Forwarding Table window showing one item in the table.

In Label	Out Labels	Interface	Nexthop	Destination	Bytes	Packets	Hw. Bytes	Hw.Pack...
expl-null					0 B	0	0 B	0

1 item



Check routing

IP routes to 10.255.255.x are missing

Ping

General Advanced

Ping To: 10.255.255.3

Interface:

ARP Ping

Packet Count:

Timeout: 1000 ms

Start
Stop
Close
New Window

Seq # /	Host	Time	Reply Size	TTL	Status
0		timeout			no route to host
1		timeout			no route to host
2		timeout			no route to host
3		timeout			no route to host
4		timeout			no route to host
5		timeout			no route to host
6		timeout			no route to host

7 items 0 of 7 packets rec... 100% packet loss

Route List

Routes Nexthops Rules VRF

+ - ✓ ✗ 📄 🔍 Find all

	Dst. Address /	Gateway	Distance /
DAC	▶ 10.14.0.0/28	vlan5.ether3 reachable	0
DAC	▶ 10.14.1.0/28	ether2 reachable	0
DAo	▶ 10.14.2.0/28	10.14.1.2 reachable ether2	110
DAo	▶ 10.14.5.0/28	10.14.1.2 reachable ether2	110
DAC	▶ 10.14.11.0/28	ether4 reachable	0
DAo	▶ 10.14.31.0/28	10.14.1.2 reachable ether2	110
DAo	▶ 10.14.32.0/28	10.14.1.2 reachable ether2	110
DAo	▶ 10.14.41.0/28	10.14.1.2 reachable ether2	110
DAo	▶ 10.14.42.0/28	10.14.1.2 reachable ether2	110
DAC	▶ 10.255.255.1	ether2 reachable	0

10 items



Routing ok, VPLS ok

The image displays three overlapping screenshots from a network management interface:

- Interface <vpls-to-r1>:** Shows configuration for the interface. The status is **running** (highlighted in red). Other fields include Last Link Up Time: Mar/26/2018 14:31:07, Link Downs: 0, Remote Label: 57, Local Label: 28, Transport: 10.255.255.1, and Transport Nexthop: 10.14.2.2.
- Route List:** Shows a table of routes. A red box highlights the following entries:

	Dst. Address	Gateway	Distance
DAo	10.255.255.1	10.14.2.2 reachable ether3	110
DAo	10.255.255.2	10.14.2.2 reachable ether3	110
DAC	10.255.255.3	ether4 reachable	0
DAo	10.255.255.4	10.14.31.4 reachable ether4	110
DAo	10.255.255.5	10.14.32.5 reachable ether5	110
DAo	10.255.255.6	10.14.32.5 reachable ether5	110
DAo	10.255.255.7	10.14.32.5 reachable ether5	110
- VPLS:** Shows the VPLS configuration for the interface. A red box highlights the entry:

RS	Name	Type	Actual MT
	vpls-to-r1	VPLS	



Lesson learned

MPLS is based on routing

- Broken/incomplete routing, broken/incomplete MPLS
- Broken MPLS, broken VPLS

Debugging:

Consider dependencies!

The screenshot shows a 'Route List' window with a table of routing entries. The table has columns for 'Dst. Address', 'Gateway', and 'Distance'. The entries include DAC and DAo routes for various IP addresses and their corresponding gateways and distances.

	Dst. Address	Gateway	Distance
DAC	10.14.0.0/28	vlan5.ether3 reachable	0
DAC	10.14.1.0/28	ether2 reachable	0
DAo	10.14.2.0/28	10.14.1.2 reachable ether2	110
DAo	10.14.5.0/28	10.14.1.2 reachable ether2	110
DAC	10.14.11.0/28	ether4 reachable	0
DAo	10.14.31.0/28	10.14.1.2 reachable ether2	110
DAo	10.14.32.0/28	10.14.1.2 reachable ether2	110
DAo	10.14.41.0/28	10.14.1.2 reachable ether2	110
DAo	10.14.42.0/28	10.14.1.2 reachable ether2	110
DAC	10.255.255.1	ether2 reachable	0

10 items



Working traceroute

Traceroute (Running)

Traceroute To: Start

Packet Size: Stop

Timeout: ms Close

Protocol: New Window

Port:

Use DNS

Count: ▼

Max Hops: ▼

Src. Address: ▼

Interface: ▼

DSCP: ▼

Routing Table: ▼

Hop	Host	Loss	Sent	Last	Avg.	Best	Worst	Std. Dev.	History	Status
1	10.14.5.6	0.0%	49	0.7ms	0.7	0.7	0.8	0.0		<MPLS:L=48,E=0>
2	10.14.42.5	0.0%	49	0.5ms	0.5	0.5	0.7	0.0		<MPLS:L=50,E=0>
3	10.14.32.3	0.0%	49	0.5ms	0.5	0.5	0.6	0.0		<MPLS:L=42,E=0>
4	10.14.2.2	0.0%	49	0.5ms	0.5	0.5	0.5	0.0		<MPLS:L=48,E=0>
5	10.255.255.1	0.0%	49	0.4ms	0.4	0.4	0.5	0.0		

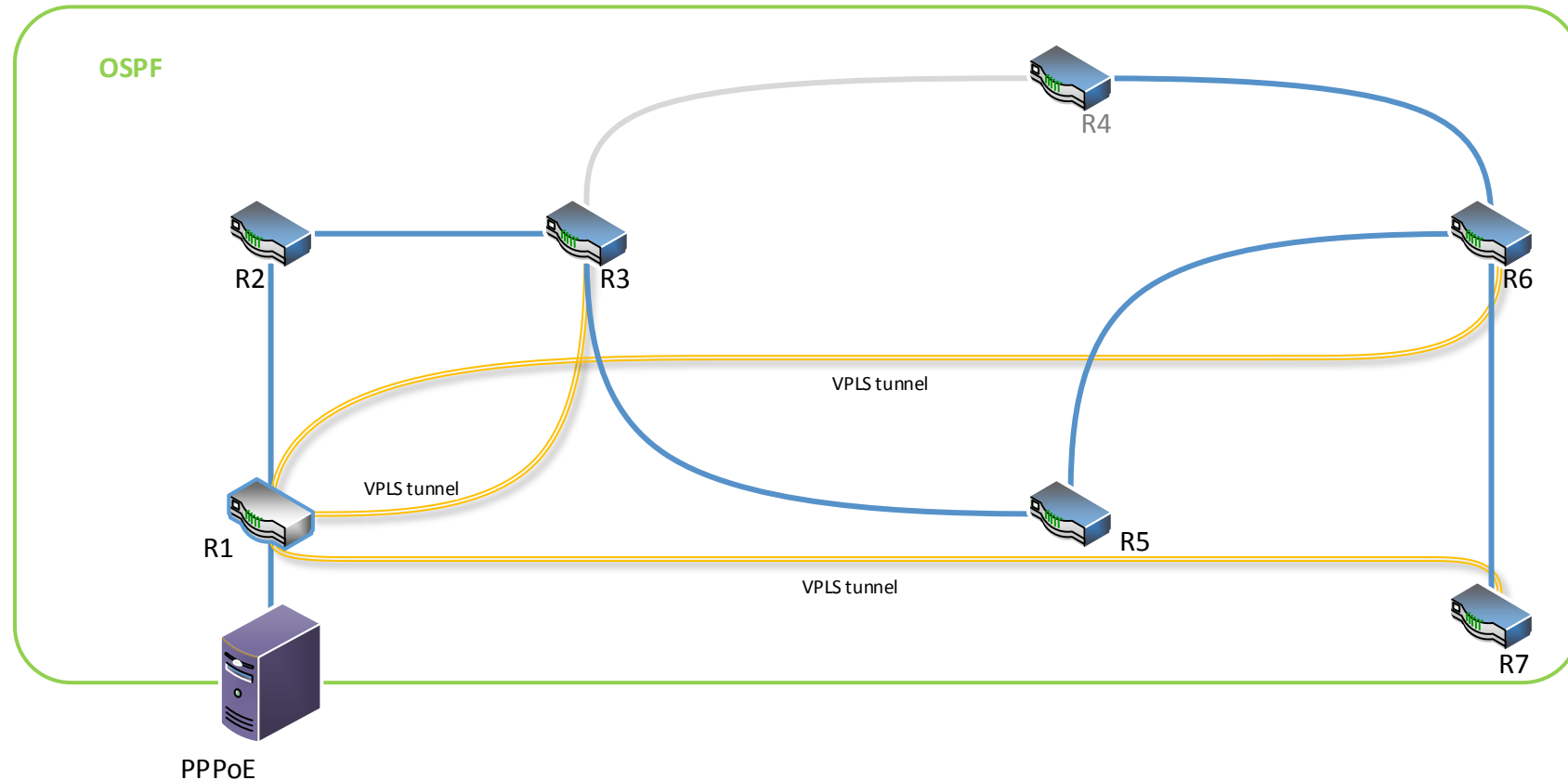
5 items



Let's break things



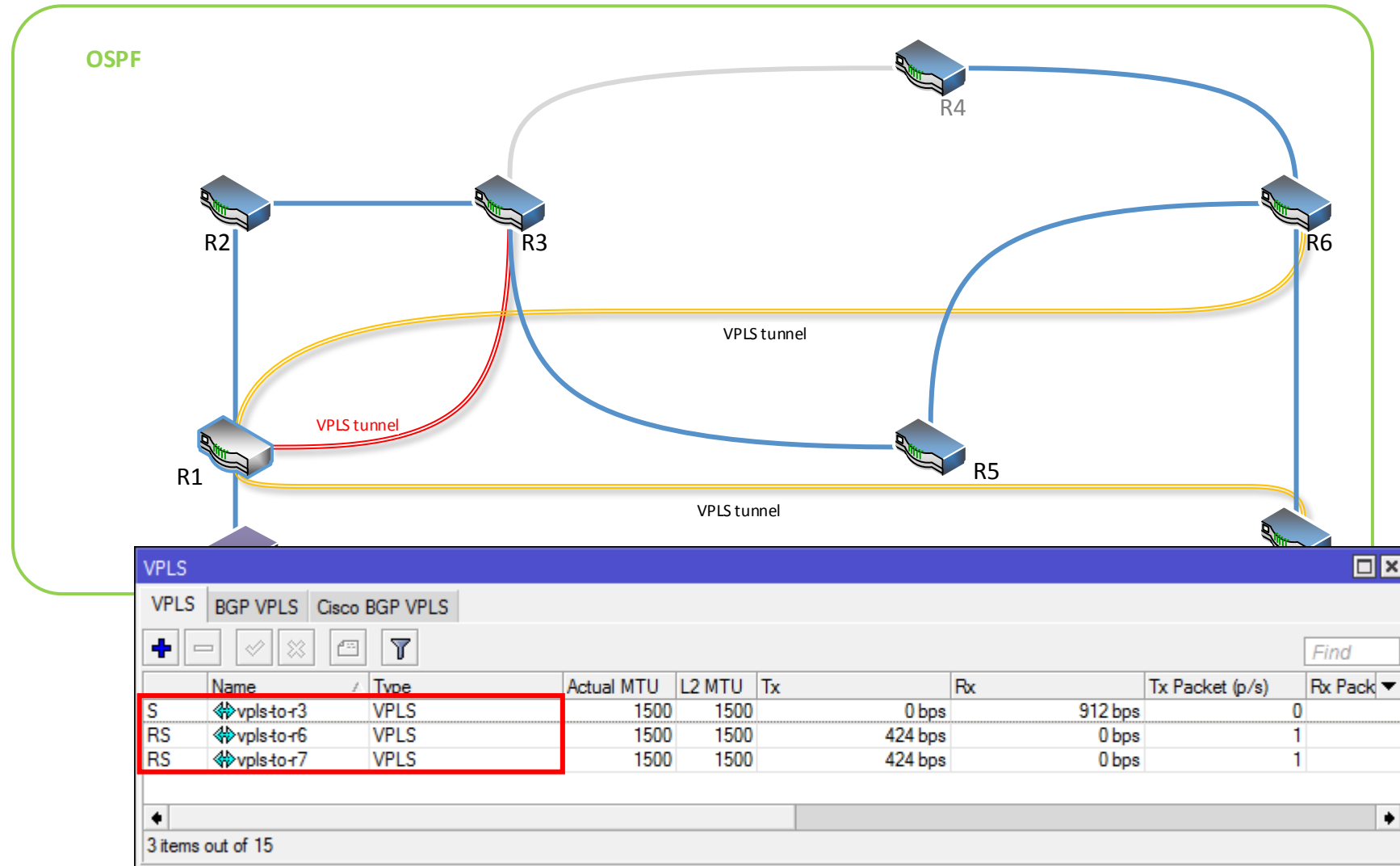
Maintenance at backup link



Maintenance at R4 (*backup link*). OSPF is going through R5.
Customers at R3 complain. Customers at R6, R7 are fine.



Maintenance at backup link





Maintenance at backup link

Interface <ether4>

General Ethernet Loop Protect Overall Stats Rx Stats ...

Name: ether4

Type: Ethernet

MTU: 1500

Actual MTU: 1500

L2 MTU: 1598

Max L2 MTU: 2028

MAC Address: 64:D1:54:4F:27:F3

ARP: enabled

ARP Timeout:

OK
Cancel
Apply
Disable
Comment
Torch
Cable Test
Blink
Reset MAC Address
Reset Counters

enabled running slave **no link**

R3: No link on ether4

10.255.255.3/32 on ether4

OSPF

Instances Networks Areas Area Ranges Virtual Links ...

Network	Area
10.14.31.0/28	area10
10.14.32.0/28	area10
10.14.2.0/28	backbone
10.255.255.3	backbone

4 items

Address List

Address	Network	Interface
10.14.2.3/28	10.14.2.0	ether3
10.14.31.3/28	10.14.31.0	ether4
10.255.255.3	10.255.255.3	ether4
10.14.32.3/28	10.14.32.0	ether5

4 items



Maintenance at backup link

No MPLS Forwarding / IP Route
for 10.255.255.3/32

MPLS							
LDP Interface		LDP Neighbor		Accept Filter		Advertise Filter	
Forwarding Table				MPLS Interface		Local Bindings	
In Label	Out Labels	Interface	Nexthop	Destination	Bytes	Pac	
expl-null					0 B		
L 49		ether2	10.14.1.2	10.14.2.0/28	0 B		
L 52	75	ether2	10.14.1.2	10.14.5.0/28	0 B		
L 59	81	ether2	10.14.1.2	10.14.32.0/28	0 B		
L 62	84	ether2	10.14.1.2	10.14.41.0/28	0 B		
L 58	80	ether2	10.14.1.2	10.14.42.0/28	0 B		
L 51		ether2	10.14.1.2	10.255.255.2	0 B		
L 65	86	ether2	10.14.1.2	10.255.255.4	0 B		
L 60	82	ether2	10.14.1.2	10.255.255.5	0 B		
L 50	70	ether2	10.14.1.2	10.255.255.6	0 B		
L 48	72	ether2	10.14.1.2	10.255.255.7	0 B		
V 56				vpls-to-r6	4403 B		
V 55				vpls-to-r7	4335 B		

13 items

Route List							
Routes		Nexthops		Rules		VRF	
+	-	✓	✗	🗨	🔍	Find	all
	Dst. Address	/	Gateway	Distance			
DAC	▶ 10.14.0.0/28		vlan5.ether3 reachable	0			
DAC	▶ 10.14.1.0/28		ether2 reachable	0			
DAo	▶ 10.14.2.0/28		10.14.1.2 reachable ether2	110			
DAo	▶ 10.14.5.0/28		10.14.1.2 reachable ether2	110			
DAC	▶ 10.14.11.0/28		ether4 reachable	0			
DAo	▶ 10.14.32.0/28		10.14.1.2 reachable ether2	110			
DAo	▶ 10.14.41.0/28		10.14.1.2 reachable ether2	110			
DAo	▶ 10.14.42.0/28		10.14.1.2 reachable ether2	110			
DAC	▶ 10.255.255.1		ether2 reachable	0			
DAo	▶ 10.255.255.2		10.14.1.2 reachable ether2	110			
DAo	▶ 10.255.255.4		10.14.1.2 reachable ether2	110			
DAo	▶ 10.255.255.5		10.14.1.2 reachable ether2	110			
DAo	▶ 10.255.255.6		10.14.1.2 reachable ether2	110			
DAo	▶ 10.255.255.7		10.14.1.2 reachable ether2	110			

14 items



Loopback bridge

Interface <loopback>

General STP VLAN Status Traffic

Name: loopback

Type: Bridge

MTU: []

Actual MTU: 1500

L2 MTU: 65535

MAC Address: 1A:F2:12:3E:E7:31

ARP: enabled

ARP Timeout: []

Admin. MAC Address: []

IGMP Snooping

Fast Forward

OK Cancel Apply Disable Comment Copy Remove Torch

enabled running slave

loopback bridge is a good idea

Loopback bridge:

Empty bridge with IP 10.255.255.x/32

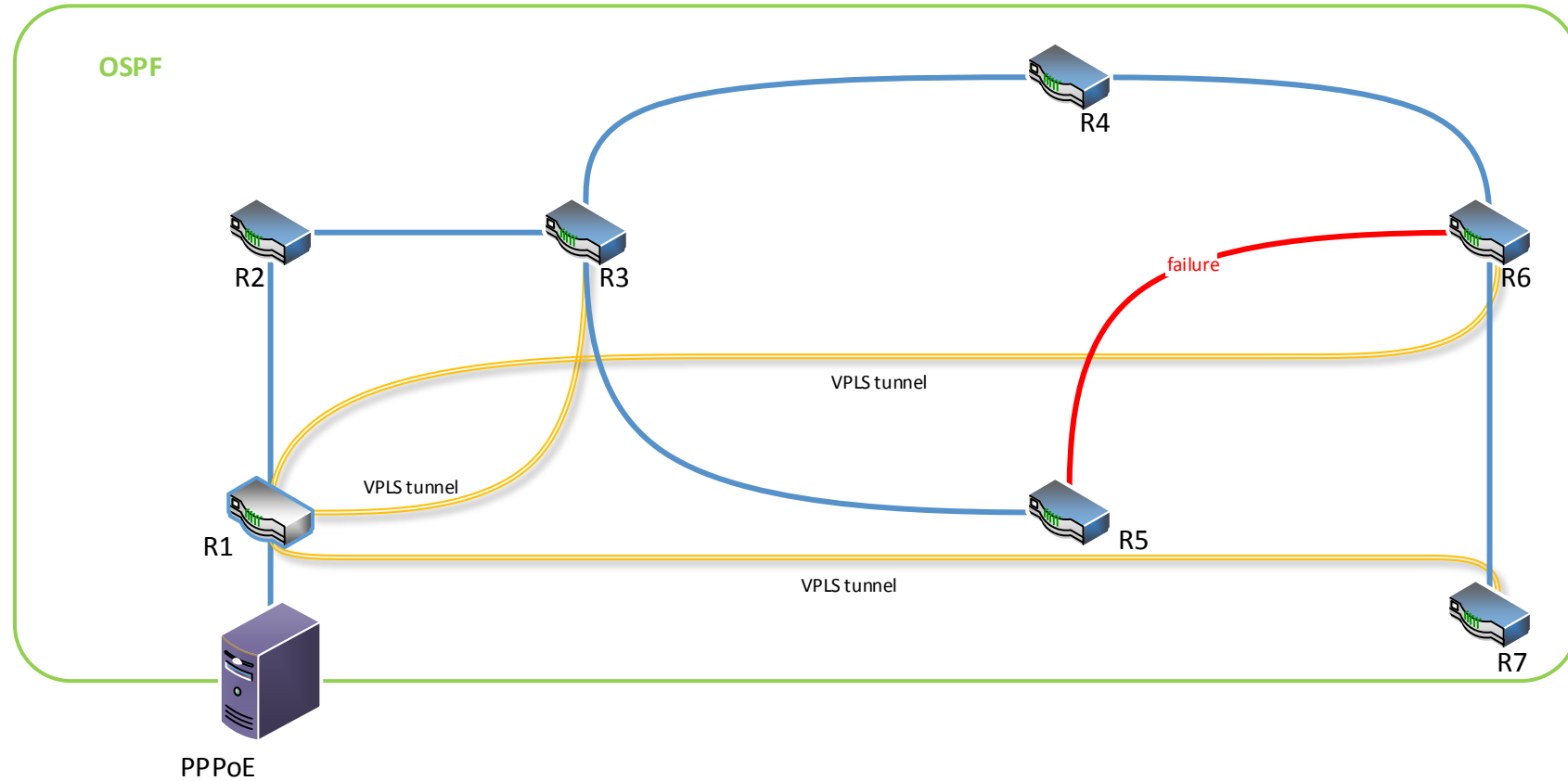
Address List

Address	Network	Interface
10.14.2.3/28	10.14.2.0	ether3
10.14.31.3/28	10.14.31.0	ether4
10.14.32.3/28	10.14.32.0	ether5
10.255.255.3	10.255.255.3	loopback

4 items



Failure at main link





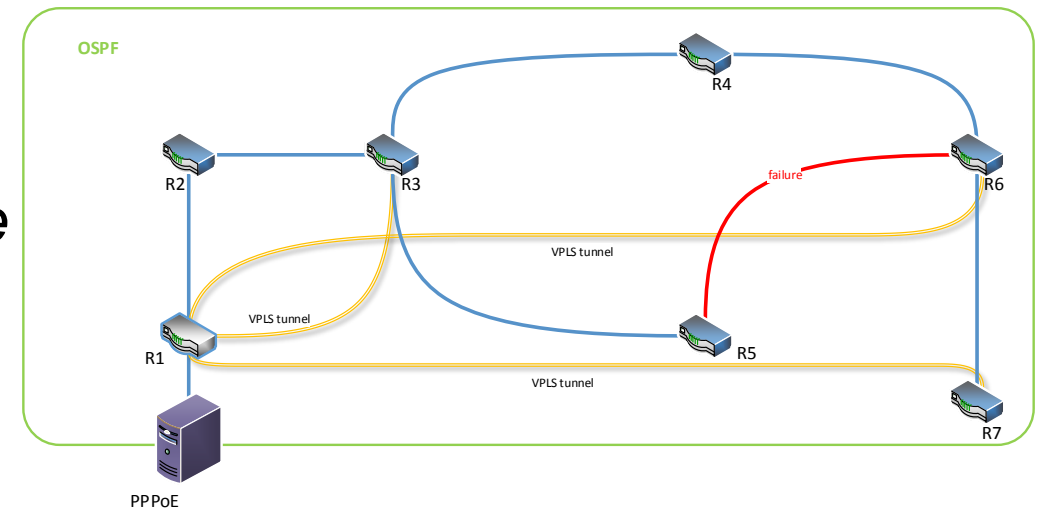
Failure at main link

Expected behaviour

- Routing through R4
- PPPoE customers at R3, R6, R7 online

Observed behaviour

- Routing through R4
- PPPoE customers at R6, R7 *offline*





Failure at main link

Ping from R1 to R7 ok

Traceroute (Running)

Traceroute To: 10.255.255.7

Packet Size: 56

Timeout: 1000 ms

Protocol: icmp

Port: 33434

Use DNS

Count: []

Max Hops: []

Src. Address: 10.255.255.1

Interface: []

DSCP: []

Routing Table: []

Hop	Host	Loss	Sent	Last	Avg.	Best	Worst	Std. Dev.	History	Status
1	10.14.1.2	0.0%	49	0.4ms	0.4	0.3	4.4	0.6		<MPLS:L=29,E=0>
2	10.14.2.3	0.0%	49	0.2ms	0.2	0.2	0.4	0.0		<MPLS:L=36,E=0>
3	10.14.31.4	0.0%	49	0.3ms	0.2	0.2	1.4	0.2		
4	10.14.41.6	0.0%	49	0.4ms	0.5	0.4	4.2	0.5		
5	10.255.255.7	0.0%	49	0.3ms	0.3	0.3	0.4	0.0		<MPLS:L=28,E=0>

5 items

Ping (Running)

General | Advanced

Ping To: 10.255.255.7

Interface: []

ARP Ping

Packet Count: []

Timeout: 1000 ms

Start

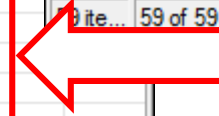
Stop

Close

New Window

Seq #	Host	Time	TTL
47	10.255.255.7	0ms	60
48	10.255.255.7	0ms	60
49	10.255.255.7	0ms	60
50	10.255.255.7	0ms	60
51	10.255.255.7	0ms	60
52	10.255.255.7	0ms	60
53	10.255.255.7	0ms	60
54	10.255.255.7	0ms	60
55	10.255.255.7	0ms	60
56	10.255.255.7	0ms	60
57	10.255.255.7	0ms	60
58	10.255.255.7	0ms	60

59 of 59 pac... 0% packet l... Min: 0 ... Avg: 0 ... Max: 0 ms





Wrong LDP interfaces at R3

- LDP:
 - *ether2*
 - ether3
 - ether4
- OSPF
 - ether3
 - ether4
 - ether5

MPLS

LDP Interface | LDP Neighbor | Accept Filter | Advertise Filter | Forwarding Table | MPLS Interface | Local Bindings | Remote Bindings

MPLS Settings | LDP Settings

Interface	Hello Interval	Hold Time	Transport Address	Accept Dy...
<i>ether2</i>	00:00:05	00:00:15		yes
ether3	00:00:05	00:00:15		yes
ether4	00:00:05	00:00:15		yes

3 items

OSPF

Interfaces | Instances | Networks | Areas | Area Ranges | Virtual Links | Neighbors | NBMA Neighbors | Sham Links | LSA | Routes | AS Border Routers

Interface	Cost	Priority	Authentic...	Authenticatio...	Network Type	Instance	Area	Neig...	State
ether3	10	1	none	*****	point to point	default	backbone	1	point to point
ether4	100	1	none	*****	point to point	default	area 10	1	point to point
ether5	10	1	none	*****	point to point	default	area 10	1	point to point
DP loopback	10	1	none	*****	broadcast	default	backbone	0	passive

4 items out of 3

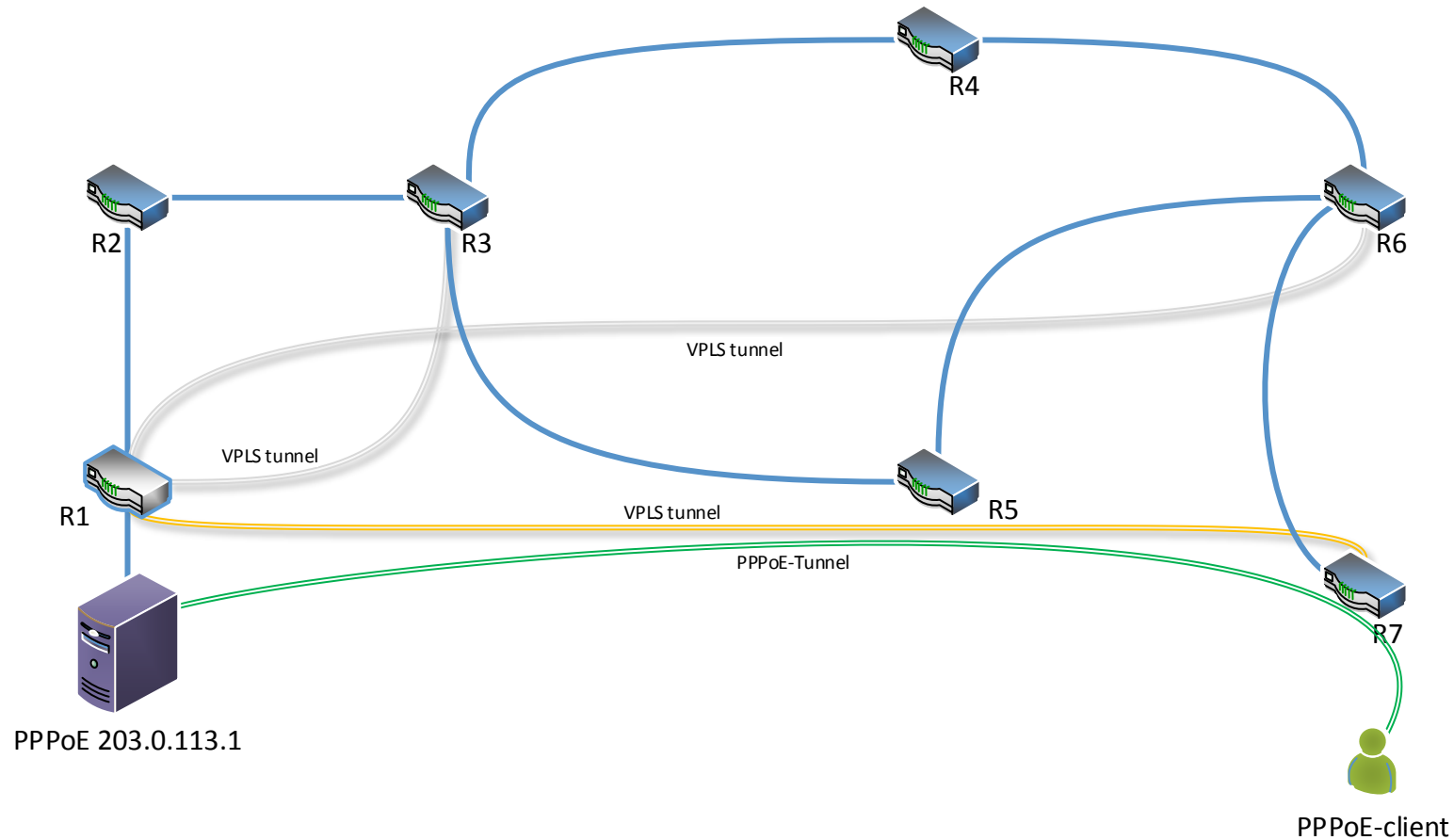


Examine setup



Monitor a PPPoE session

Bandwidth-test: PPPoE client to PPPoE server (download)

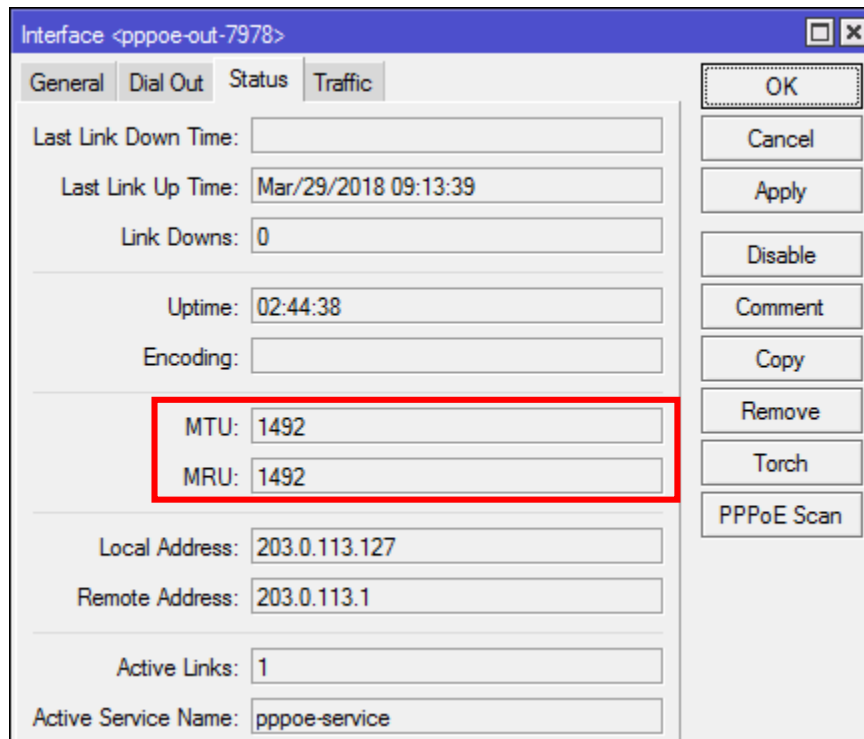




Monitor a PPPoE session

Bandwidth-test: PPPoE client to PPPoE server (download)

MTU PPPoE Client: 1492 → Bandwidth-test with 1492



Interface <pppoe-out-7978>

General | Dial Out | Status | Traffic

Last Link Down Time:

Last Link Up Time: Mar/29/2018 09:13:39

Link Downs: 0

Uptime: 02:44:38

Encoding:

MTU: 1492

MRU: 1492

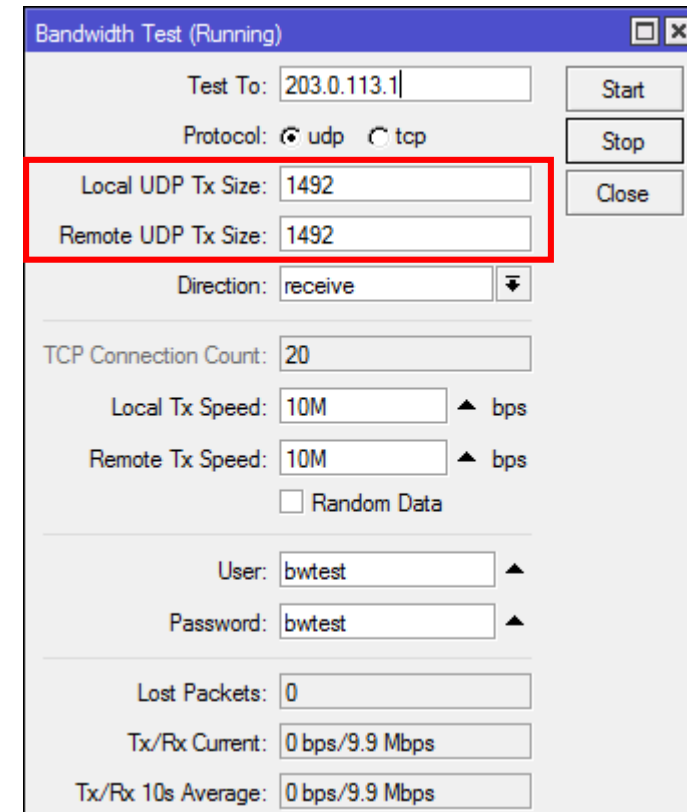
Local Address: 203.0.113.127

Remote Address: 203.0.113.1

Active Links: 1

Active Service Name: pppoe-service

Buttons: OK, Cancel, Apply, Disable, Comment, Copy, Remove, Torch, PPPoE Scan



Bandwidth Test (Running)

Test To: 203.0.113.1

Protocol: udp tcp

Local UDP Tx Size: 1492

Remote UDP Tx Size: 1492

Direction: receive

TCP Connection Count: 20

Local Tx Speed: 10M bps

Remote Tx Speed: 10M bps

Random Data

User: bwtest

Password: bwtest

Lost Packets: 0

Tx/Rx Current: 0 bps/9.9 Mbps

Tx/Rx 10s Average: 0 bps/9.9 Mbps

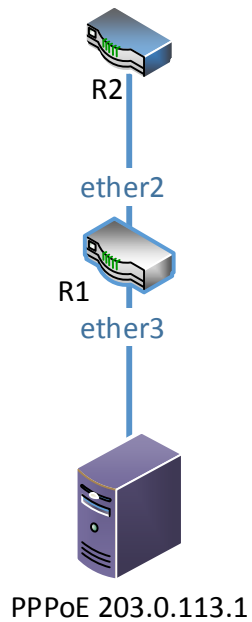
Buttons: Start, Stop, Close



Monitor a PPPoE session

On R1

- Interface to R2: 1697 p/s
- Interface to PPPoE: 846 p/s



Bandwidth Test (Running)

Test To: 203.0.113.1

Protocol: udp tcp

Local UDP Tx Size: 1492

Remote UDP Tx Size: 1492

Direction: receive

TCP Connection Count: 20

Local Tx Speed: 10M bps

Remote Tx Speed: 10M bps

Random Data

Start

Stop

Close

Interface List

Interface	Name	MTU	Actual MTU	L2 MTU	Tx	Rx	Tx Packet (p/s)	Rx Packet (p/s)
R	ether1	1500	1500	1598	119.7 kbps	20.4 kbps	13	34
::: To R2								
R	ether2	1500	1500	1598	10.7 Mbps	6.2 kbps	1697	10
::: To PPPoE network								
RS	ether3	1500	1500	1598	1856 bps	10.2 Mbps	3	846
	ether4	1500	1500	1598	0 bps	0 bps	0	0
	ether5	1500	1500	1598	0 bps	0 bps	0	0

5 items out of 15



Fragmentation

Packet fragmentation?

Benefits of VPLS vs. EoIP

- VPLS: No fragmentation (*if done right*)

Interface	Name	MTU	Actual MTU	L2 MTU	Tx	Rx	Tx Packet (p/s)	Rx Packet (p/s)
R	ether1	1500	1500	1598	119.7 kbps	20.4 kbps	13	34
::: To R2								
R	ether2	1500	1500	1598	10.7 Mbps	6.2 kbps	1 697	10
::: To PPPoE network								
RS	ether3	1500	1500	1598	1856 bps	10.2 Mbps	3	846
	ether4	1500	1500	1598	0 bps	0 bps	0	0
	ether5	1500	1500	1598	0 bps	0 bps	0	0

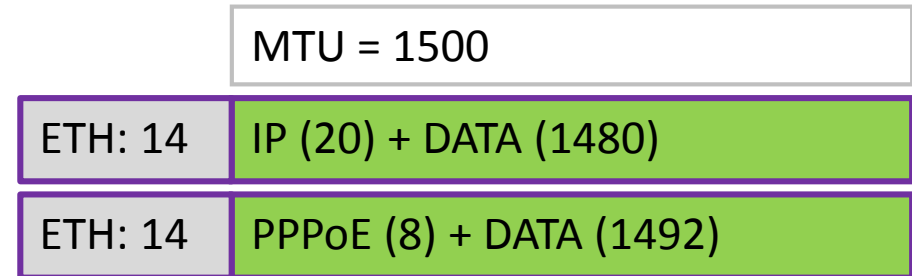
5 items out of 15



Packet sizes

Original frame

- L3 Size = 1500
- Full Frame Size = 1514

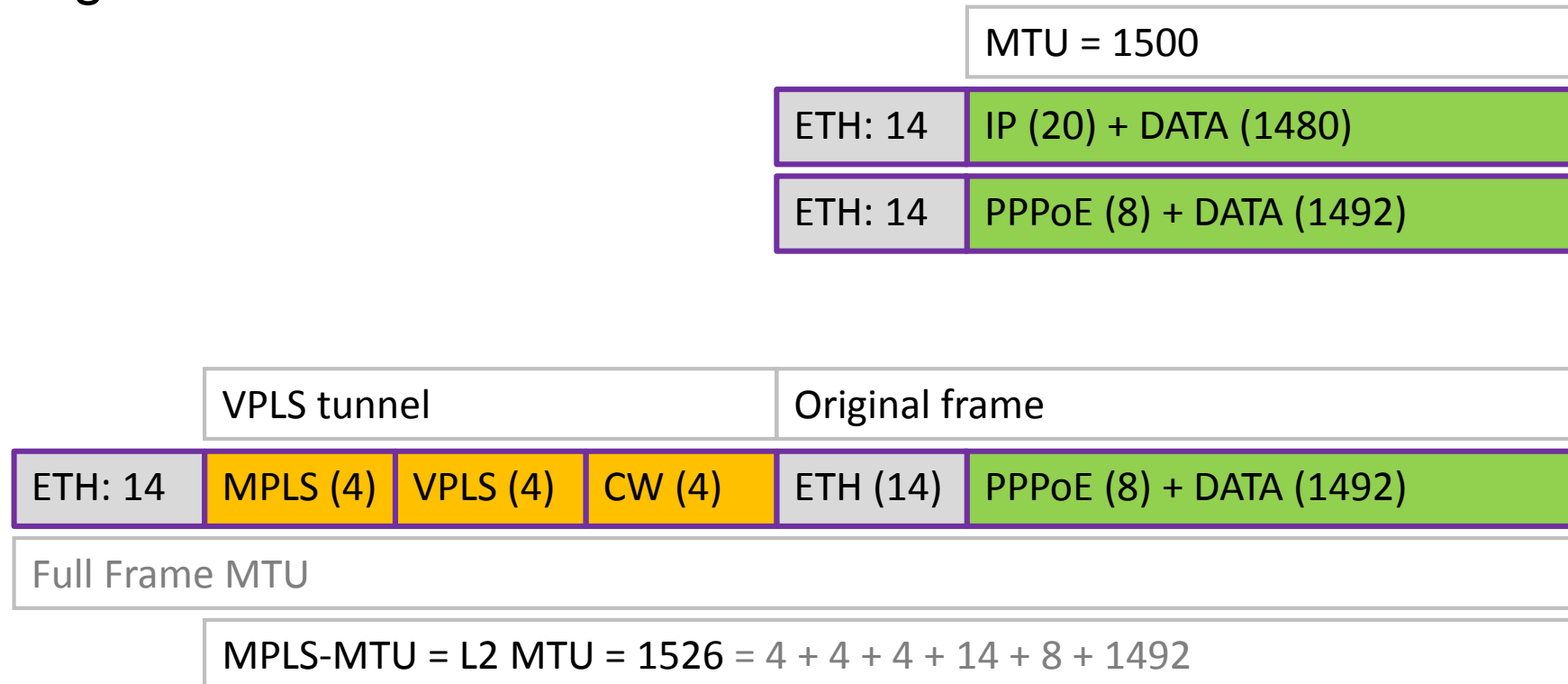




Packet sizes

Insertion of 1500 bytes (MTU) packet into VPLS tunnel:

No fragmentation



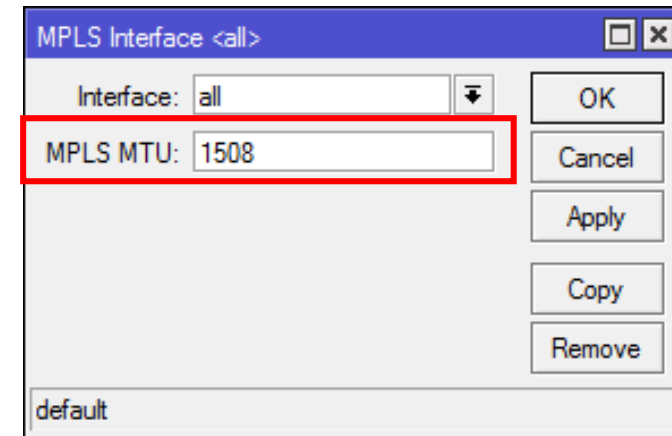


Packet sizes

VPLS packet is *fragmented* because:

Resulting MPLS-MTU: 1526

Interface MPLS MTU: 1508 (default)



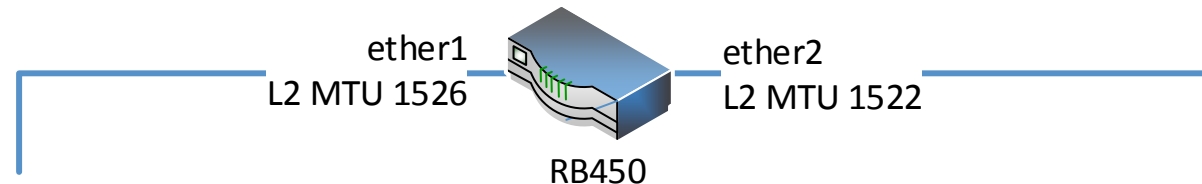
MPLS-MTU = L2 MTU = 1526



Increase interface MPLS MTU

If hardware capable: Increase interface MPLS MTU

- L2 MTU (see Maximum Transmission Unit on RouterBoards)
 - RB433, RB450, RB493: ether1: 1526, ether2-last: 1522
 - RB433GL, RB450G, RB493G: all interfaces: 1520
 - ...
- Switches, media converters, ...





MPLS MTU set to 1526

MPLS Interface MTU: 1526 →
Corresponding packet counters

- PPPoE client
- Interface to backbone
- Interface to PPPoE server

Interface <pppoe-out-7978>

General | Dial Out | Status | Traffic

Tx/Rx Rate: 10.0 Mbps / 936 bps

Tx/Rx Packet Rate: 840 p/s / 2 p/s

FP Tx/Rx Rate: 0 bps / 936 bps

FP Tx/Rx Packet Rate: 0 p/s / 2 p/s

Tx/Rx Bytes: 329.4 MiB / 20.8 GiB

Tx/Rx Packets: 312 667 / 15 228 138

Tx/Rx Drops: 0 / 0

Tx/Rx Errors: 0 / 0

Status: connected

Interface List

Interface	Name	MTU	Actual MTU	L2 MTU	Tx	Rx	Tx Packet (p/s)	Rx Packet (p/s)
R	ether1	1500	1500	1598	123.4 kbps	35.9 kbps	14	53
	::: To R2							
R	ether2	1500	1500	1598	7.1 kbps	10.4 Mbps	11	853
	::: To PPPoE network							
RS	ether3	1500	1500	1598	10.2 Mbps	1344 bps	849	2
	ether4	1500	1500	1598	0 bps	0 bps	0	0
	ether5	1500	1500	1598	0 bps	0 bps	0	0

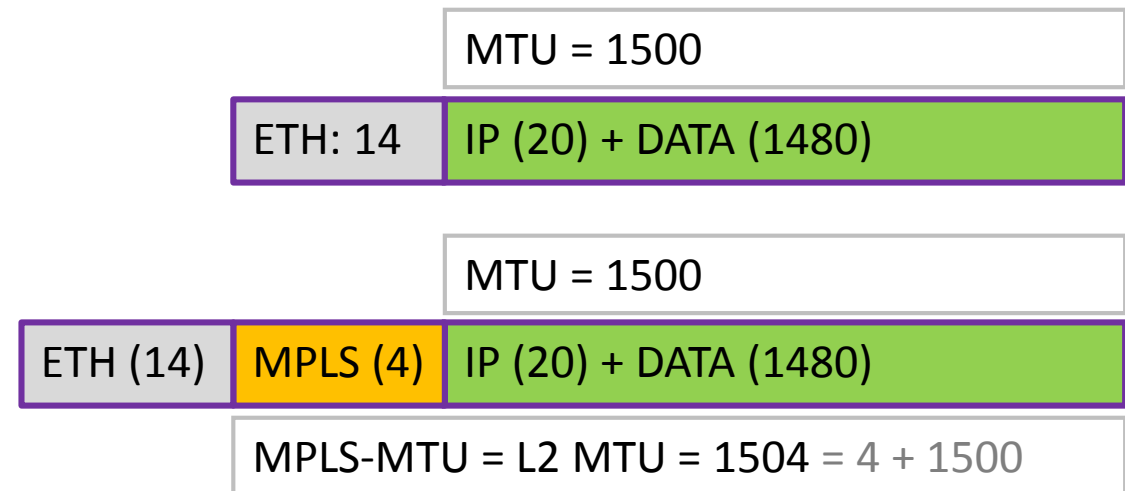
5 items out of 15



Why 1508?

1508 is enough for

- MPLS for packet forwarding (1 MPLS label)

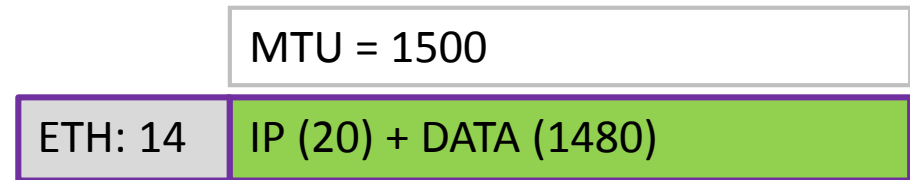




Why 1508?

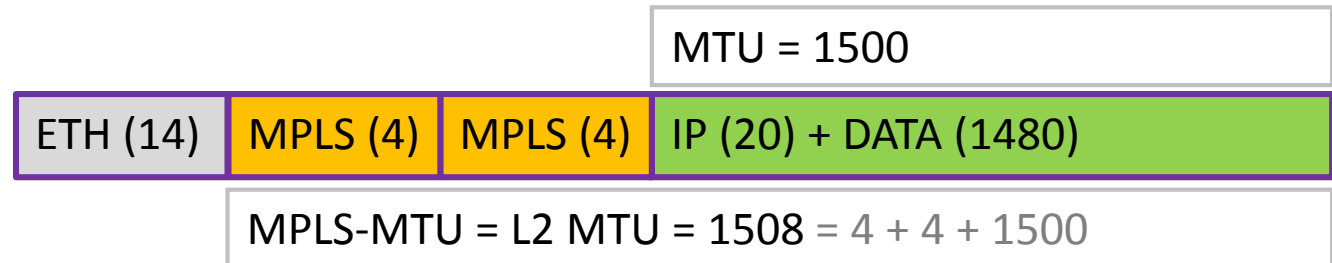
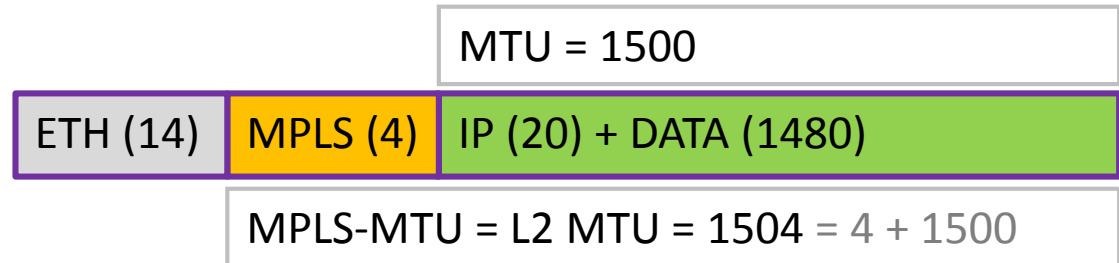
1508 is enough for

- MPLS for packet forwarding (1 MPLS label)
- Targeted LDP (2 MPLS labels)



Default 1526

- Too large (?)

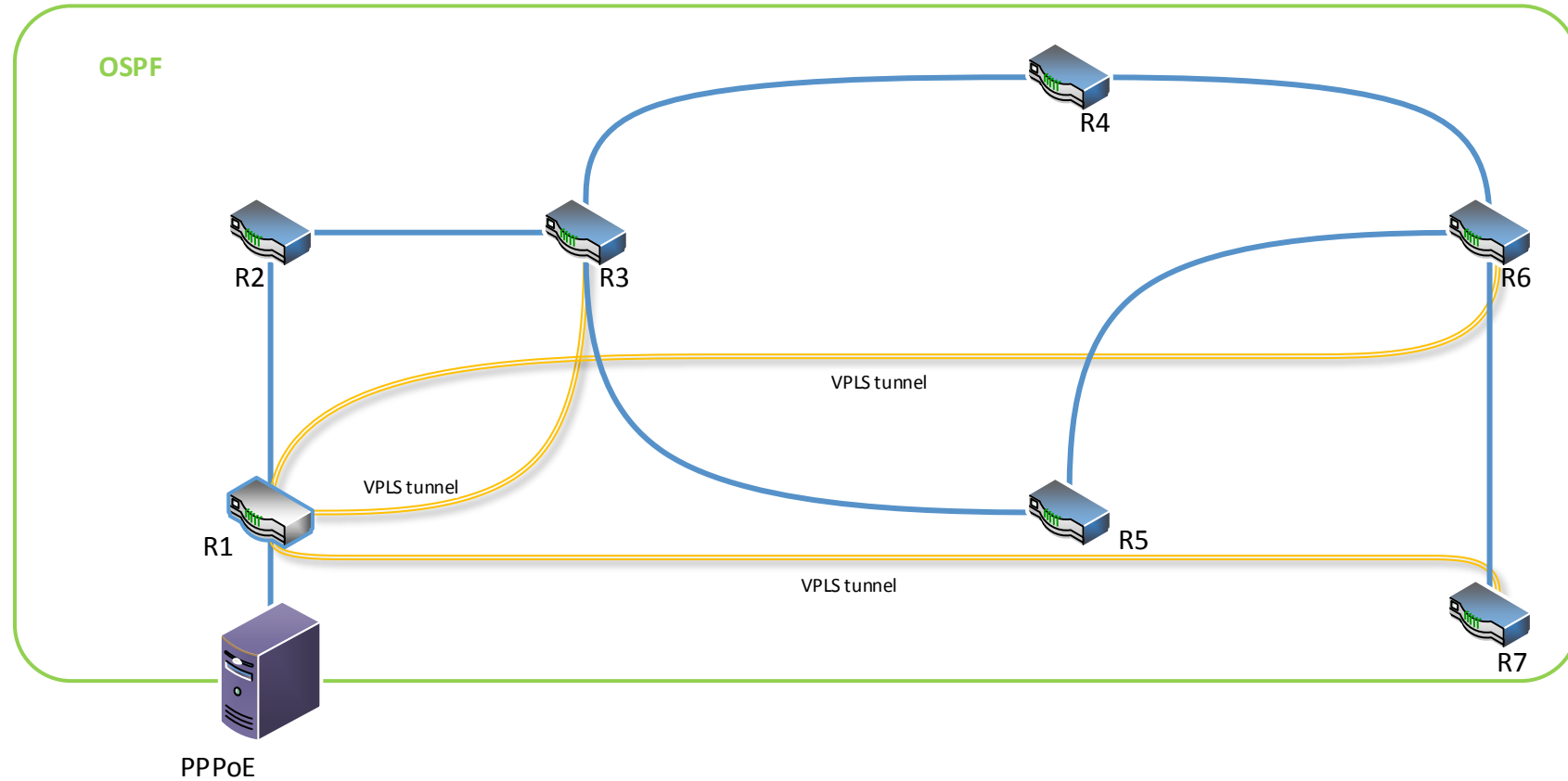




Network improvements



Current network





Redundancy

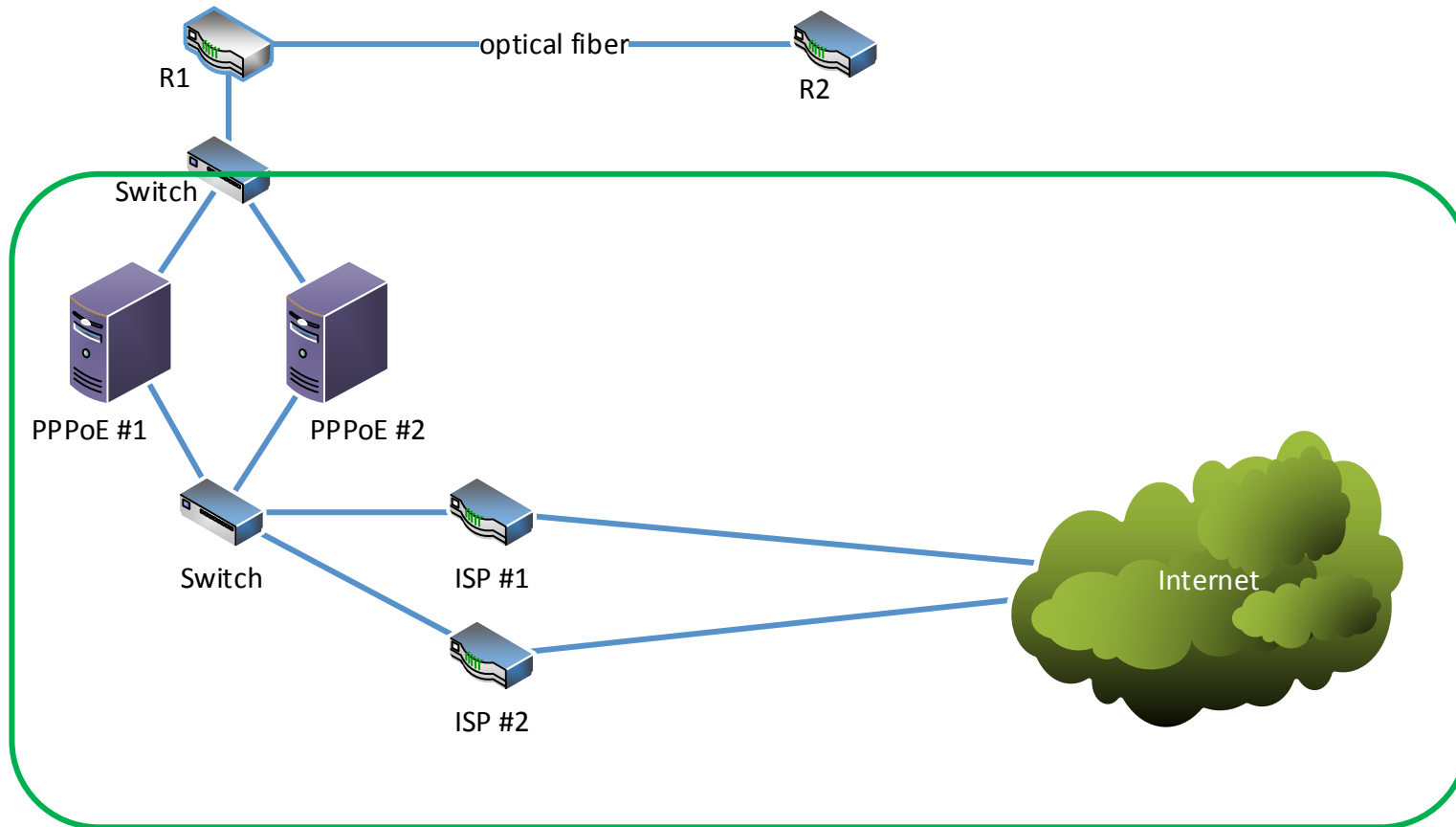
Redundancy:

- Type / coverage depends on
 - setup
 - needs
 - customer / network
- No claim for completeness
- Examples

Redundancy can become complex. Complexity can result in issues.



Redundancy at main site

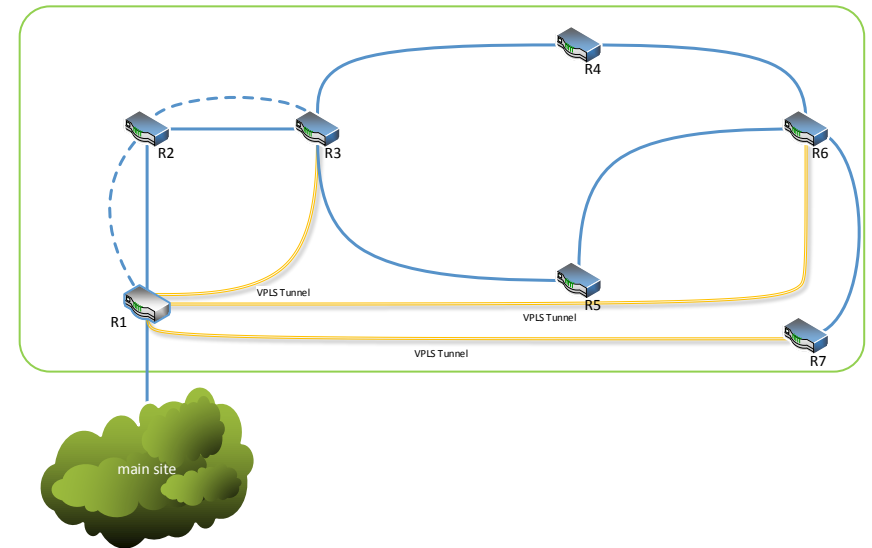


Green frame: See presentation of Patrik Schaub
(Access all FMS Internetservice presentations: [click](#))



Redundancy at backbone

- Additional link / ip subnet between R1/R2 and R2/R3
- 2nd link is backup – same as on R3
 - OSPF interfaces:
High(er) cost for backup link
- Don't forget to add LDP interface

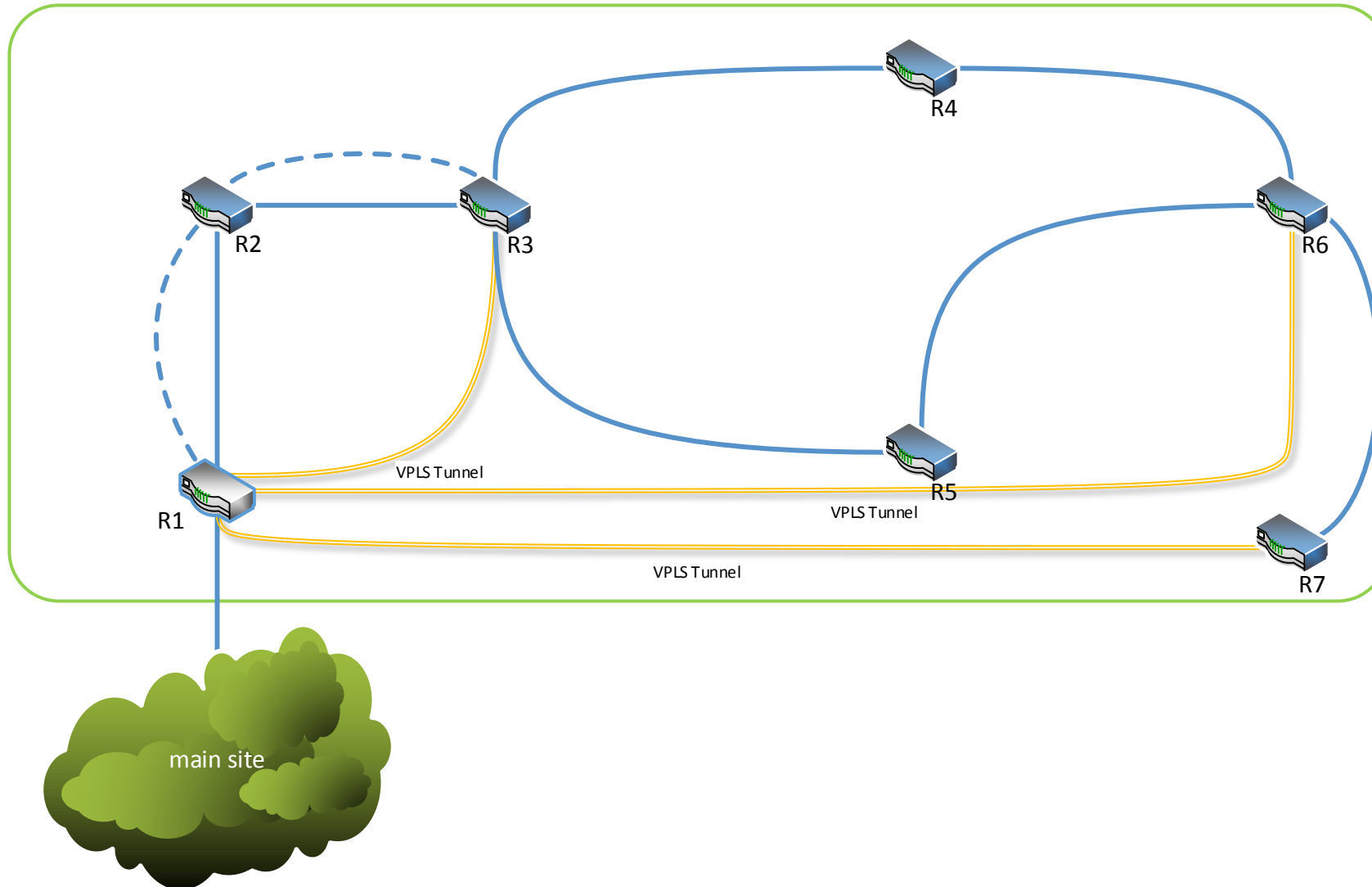


Interface	Cost	Network Type	Instance	Area	Neighbors	State
::: From R1 to R2, main link						
ether2	10	point to point	default	area5	1	point to point
::: From R1 to R2, backup link						
ether5	100	point to point	default	area5	1	point to point
::: Loopback interface, passive						
loopback	10	broadcast	default	area5	0	passive

Interface	Hello Interval	Hold Time	Transport Address	Accept Dy...
ether2	00:00:05	00:00:15		yes
ether5	00:00:05	00:00:15		yes



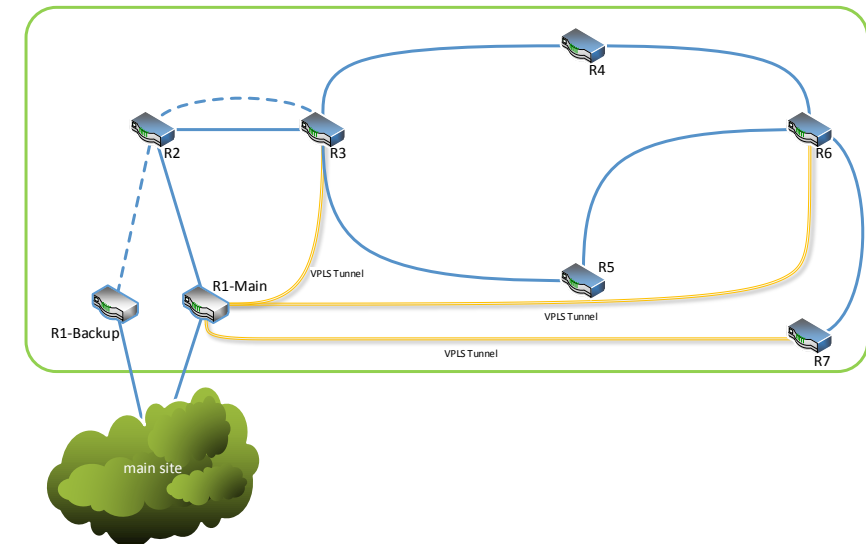
Redundancy at backbone





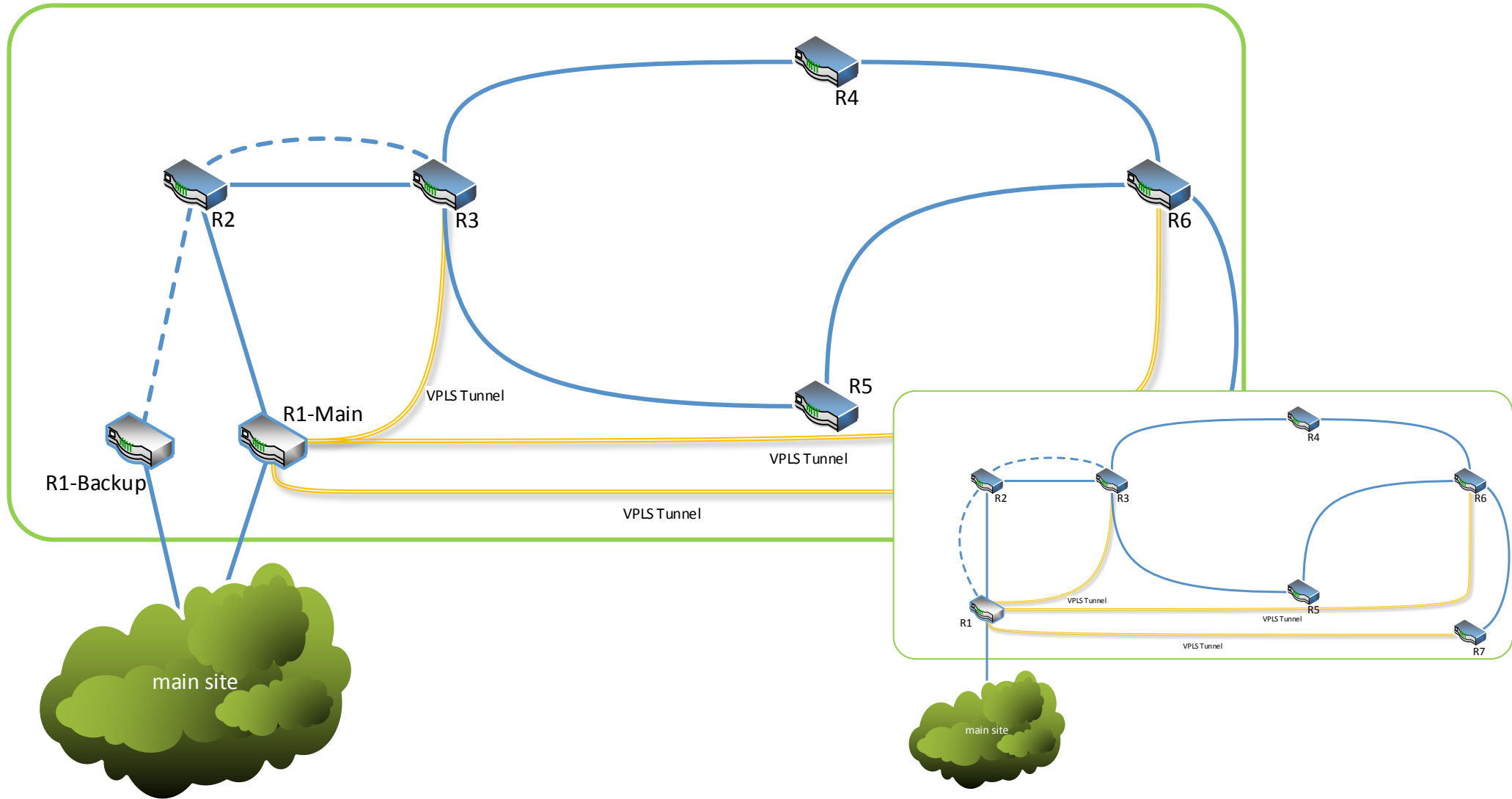
Redundancy for R1

- Clone R1:
 - R1-Main (10.255.255.11)
 - R1-Backup (10.255.255.12)
- Main link connected to R1-Main
- Backup link connected to R1-Backup
- VPLS go to R1-Main (10.255.255.1)



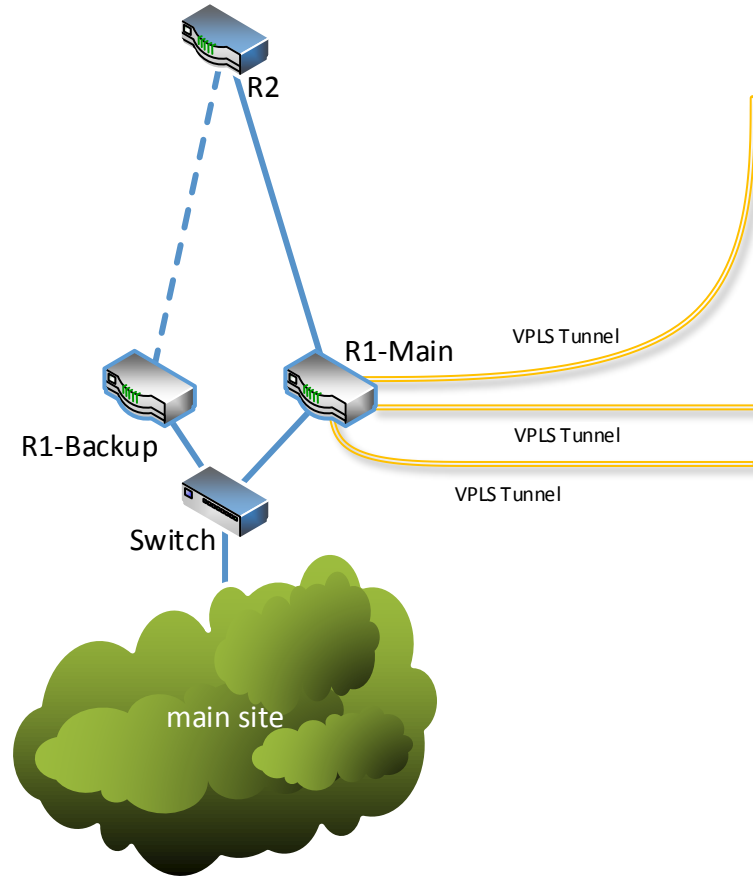


Redundancy for R1





Redundancy for R1



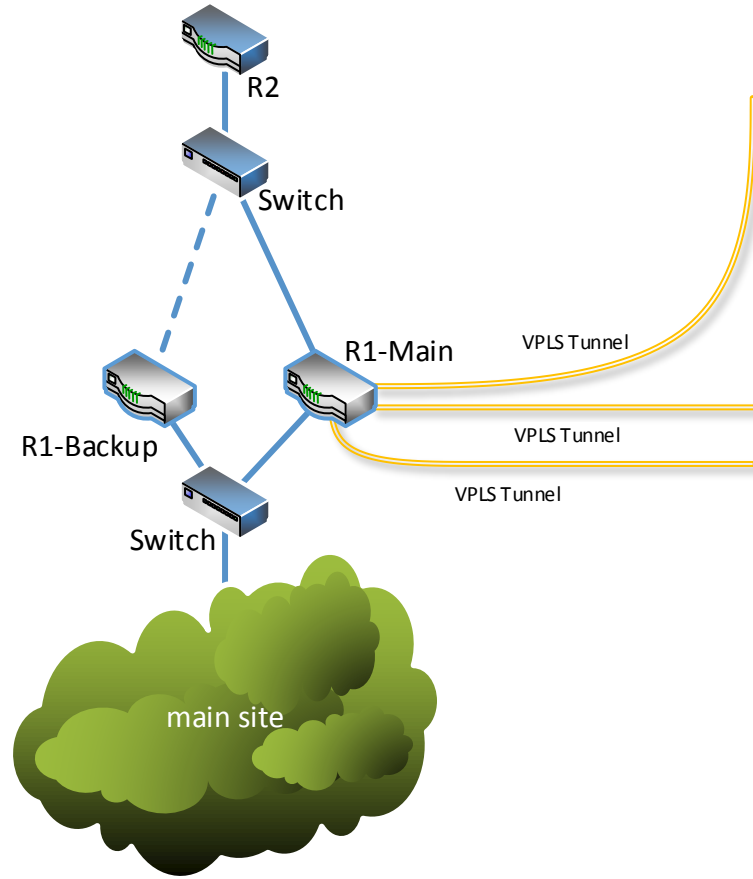
Who is R1-Main / R1-Backup?

Who is 10.255.255.1?

- No VRRP between Main / Backup on Interface to R2 (different L3 networks)



Redundancy for R1



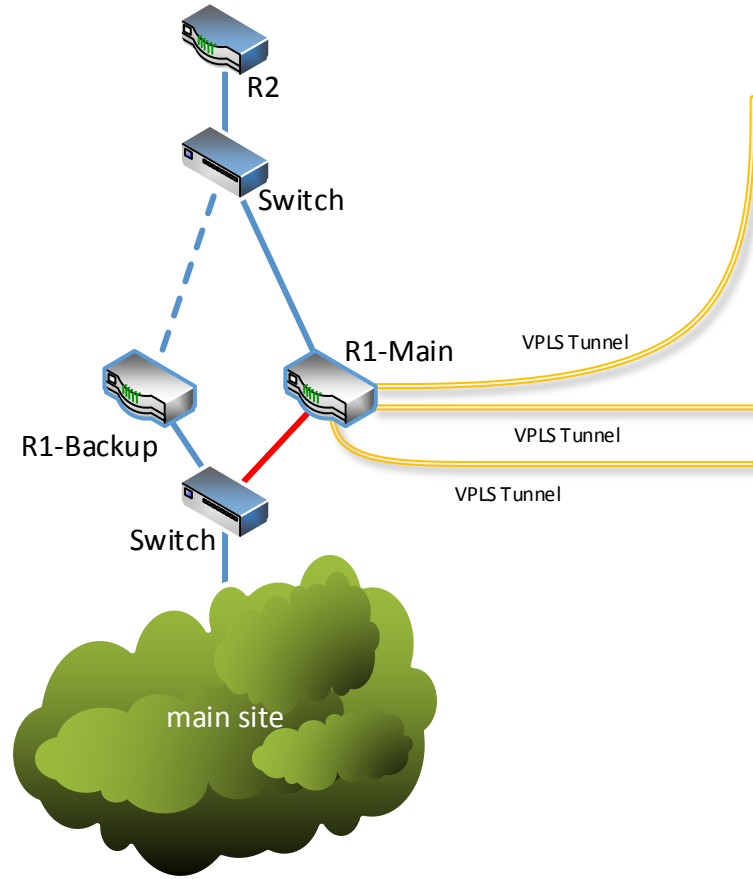
Who is R1-Main / R1-Backup?

Who is 10.255.255.1?

- Same L2 for R1-Main, R1-Backup and R2
 - VRRP on R2 side
 - Backup path: Decision by RSTP



Redundancy for R1



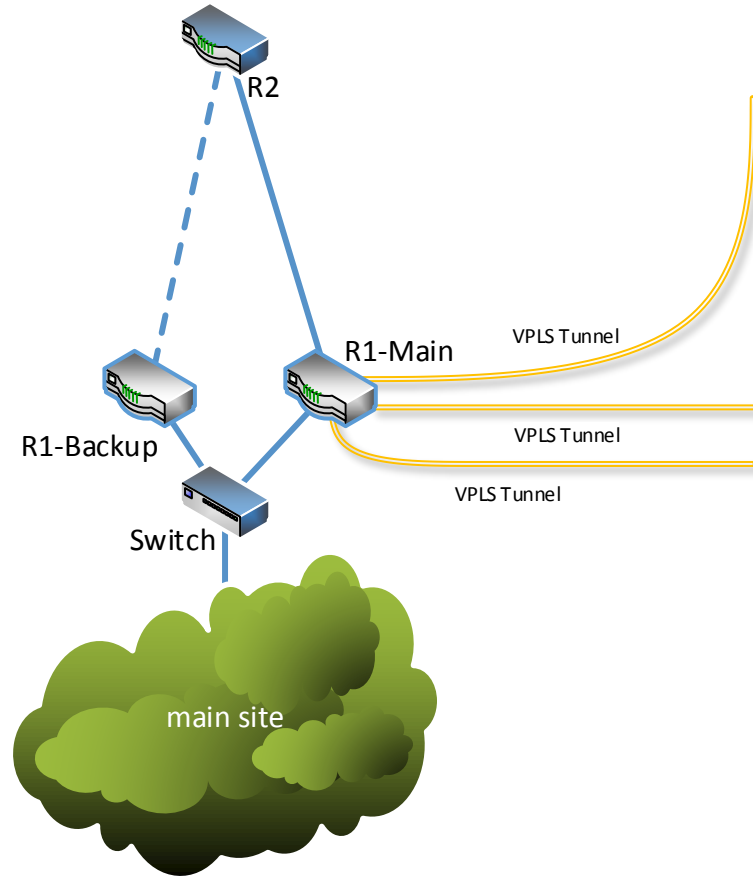
Who is R1-Main / R1-Backup?

Who is 10.255.255.1?

- Same L2 for R1-Main, R1-Backup and R2
 - VRRP on R2 side
 - Backup path: Decision by RSTP
- Failure on link to main site
 - VRRP is fine
 - Clients offline



Redundancy for R1



Who is R1-Main / R1-Backup?

Who is 10.255.255.1?

- R1-Main and R1-Backup:
Connected to main site switch
- VRRP on this side
 - Management VLAN?



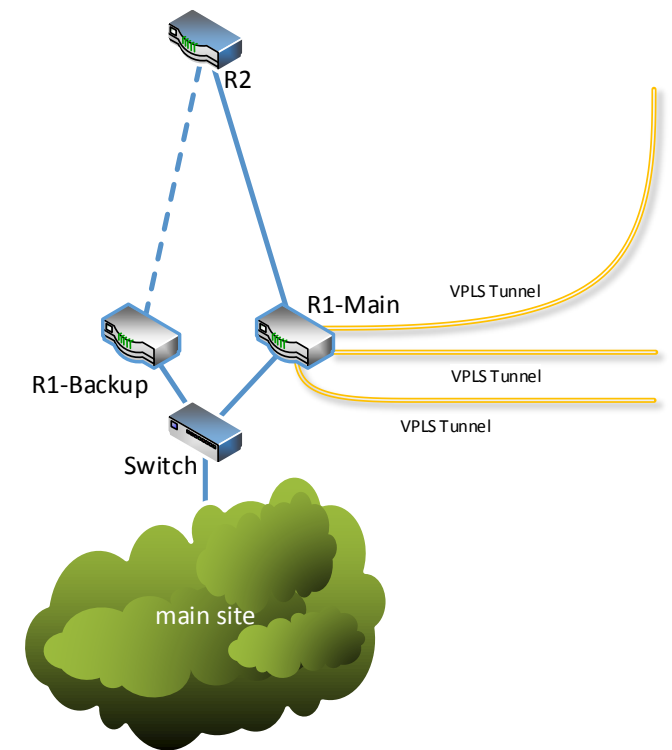
Redundancy for R1

VRRP and MPLS on R1-Main

- IP 10.255.255.1/32 on VRRP interface
- LSR ID = Transport address 10.255.255.1
- 10.255.255.11/32 on loopback, for OSPF

VRRP and MPLS on R1-Backup

- IP 10.255.255.1/32 on VRRP interface
- LSR ID = Transport address 10.255.255.1
- 10.255.255.12/32 on loopback, for OSPF





Let's ~~break~~ test things

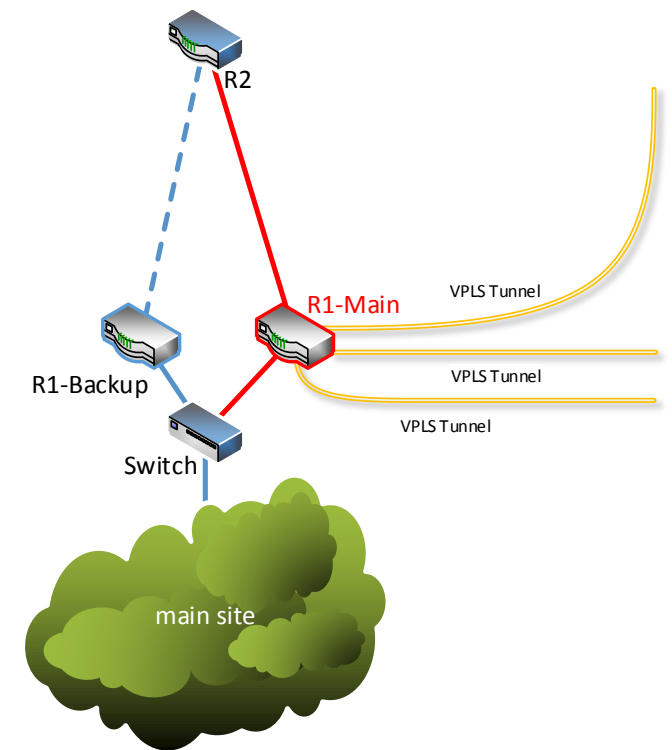
Failure of R1-Main or failure of link to main site

Expected behaviour

- 10.255.255.1 on R1-Backup
- VPLS tunnels to R1-Backup → up
- PPPoE clients reconnecting

Observed behaviour

- Everything fine (~~stop testing!~~)



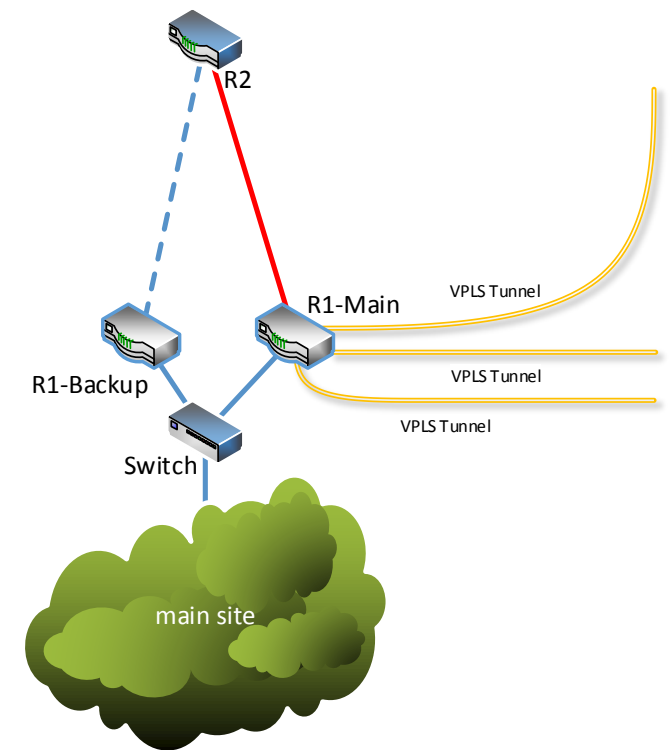


Let's break test things

Failure of link R1-Main to R2

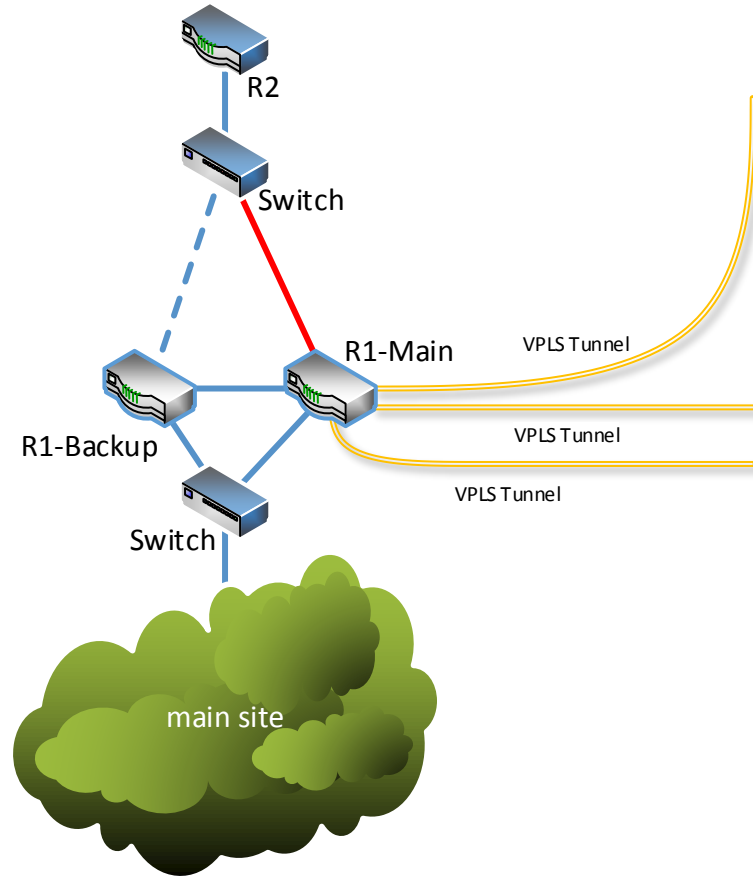
Expected behaviour

- 10.255.255.1 on R1-Main
- R1-Main VPLS master
- R2: *No route* to 10.255.255.1 (OSPF)
- Clients offline





Let's break test things



OSPF and LDP on crosslink

Expected behaviour

- 10.255.255.1 on R1-Main
- R2: route to 10.255.255.1
- VPLS ok & clients online

Observed behaviour

- Clients offline



Let's break test things

Tests from R2:

- Route via R1-Backup
- Ping to 10.255.255.1 ok
- Traceroute ok

	Dst. Address	Gateway	Distance
DAo	10.255.255.1	10.14.12.12 reachable ether5	110
DAC	10.255.255.2	loopback reachable	0
DAo	10.255.255.3	10.14.2.3 reachable ether3	110
DAo	10.255.255.4	10.14.2.3 reachable ether3	110
DAo	10.255.255.5	10.14.2.3 reachable ether3	110
DAo	10.255.255.6	10.14.2.3 reachable ether3	110
DAo	10.255.255.7	10.14.2.3 reachable ether3	110
DAo	10.255.255.11	10.14.12.12 reachable ether5	110
DAo	10.255.255.12	10.14.12.12 reachable ether5	110

```
Terminal
[admin@R02] > /ping 10.255.255.1 count=5
  SEQ HOST                SIZE TTL TIME  STATUS
  ---
  0 10.255.255.1           56  63 0ms
  1 10.255.255.1           56  63 0ms
  2 10.255.255.1           56  63 0ms
  3 10.255.255.1           56  63 0ms
  4 10.255.255.1           56  63 0ms
sent=5 received=5 packet-loss=0% min-rtt=0ms avg-rtt=0ms max-rtt=0ms

[admin@R02] > /tool traceroute 10.255.255.1 src-address=10.255.255.2
# ADDRESS                LOSS SENT    LAST    AVG    BEST    WORST STD-DEV STATUS
1 10.14.12.12             0%    7    0.3ms  0.3    0.3    0.3    0    <MPLS:L=57,E=0>
2 10.255.255.1           0%    7    0.2ms  0.2    0.2    0.3    0
-- [Q quit|D dump|C-z pause]
```




Let's break test things

Tests from R7:

- Ping to 10.255.255.1 ok
- Traceroute... ?

```
Terminal
[admin@S:R07] > /ping 10.255.255.1 count=5
  SEQ HOST                SIZE TTL TIME  STATUS
  --- ---                -
  0 10.255.255.1          56  59 0ms
  1 10.255.255.1          56  59 0ms
  2 10.255.255.1          56  59 0ms
  3 10.255.255.1          56  59 0ms
  4 10.255.255.1          56  59 0ms
  sent=5 received=5 packet-loss=0% min-rtt=0ms avg-rtt=0ms max-rtt=0ms

[admin@S:R07] > /tool traceroute 10.255.255.1 src-address=10.255.255.7
# ADDRESS                LOSS SENT  LAST  AVG  BEST  WORST STD-DEV STATUS
1 10.14.5.6                0%  14  0.5ms 0.5  0.5  0.6    0 <MPLS:L=51,E=0>
2 10.14.41.4               0%  14  0.4ms 0.4  0.4  0.5    0 <MPLS:L=61,E=0>
3 10.14.31.3               0%  14  0.4ms 0.4  0.3  0.7    0.1 <MPLS:L=53,E=0>
4 10.14.2.2                0%  14  0.4ms 0.3  0.3  0.4    0 <MPLS:L=50,E=0>
5 10.14.12.12              0%  14  0.3ms 0.3  0.3  0.3    0
6 10.255.255.1             0%  14  0.4ms 0.3  0.3  0.4    0
- [Q quit|D dump|C-z pause]
```



Let's break test things

Routing between R1-Backup and R1-Main ok

MPLS/LDP broken on R1-Backup

- No forwarding Table

The screenshot shows a network configuration window titled 'MPLS'. It has several tabs: 'LDP Interface', 'LDP Neighbor', 'Accept Filter', 'Advertise Filter', 'Forwarding Table', 'MPLS Interface', 'Local Bindings', and 'Remote Bindings'. The 'Forwarding Table' tab is selected. Below the tabs is a search bar with a magnifying glass icon and a 'Find' button. A table displays the forwarding table contents. The table has columns for 'In Label', 'Out Labels', 'Interface', 'Nexthop', 'Destination', 'Bytes', 'Packets', 'Hw. Bytes', and 'Hw.Pack...'. A single entry is visible: 'expl-null' in the 'In Label' column, with '0 B' in 'Bytes', '0' in 'Packets', and '0 B' in 'Hw. Bytes'. The status bar at the bottom indicates '1 item'.

In Label	Out Labels	Interface	Nexthop	Destination	Bytes	Packets	Hw. Bytes	Hw.Pack...
expl-null					0 B	0	0 B	0

Routing is not enough for VPLS!



Let's ~~break~~ test things

Simple reason:

- LSR ID and Transport Address 10.255.255.1 is used on R1-Backup *and* R1-Main(!)
- IP 10.255.255.1 is active only on R1-Main (VRRP master)
- Duplicate ID (and transport address): Good idea? (No.)

A screenshot of a Windows-style dialog box titled "LDP Settings". The dialog has a blue title bar with standard window controls (minimize, maximize, close). The main area is light gray and contains several settings:

- A dropdown menu at the top is set to "Enabled" with a checkmark.
- A text box labeled "LSR ID:" contains the value "10.255.255.1".
- A text box labeled "Transport Address:" contains the value "10.255.255.1".
- A text box labeled "Path Vector Limit:" contains the value "255".
- A text box labeled "Hop Limit:" contains the value "255".
- Three checkboxes are located at the bottom:
 - Loop Detect
 - Use Explicit Null
 - Distribute For Default Route

On the right side of the dialog, there are three buttons: "OK", "Cancel", and "Apply".



Let's fix things

(One possible) Solution:

- On VRRP Master:

Set LSR ID and Transport Address to 10.255.255.1

- On VRRP Backup:

Set LSR ID and Transport Address to router unique address

(available on loopback)

Result: Working MPLS between routers

(OSPF was using unique address as Router ID.)



Let's fix things

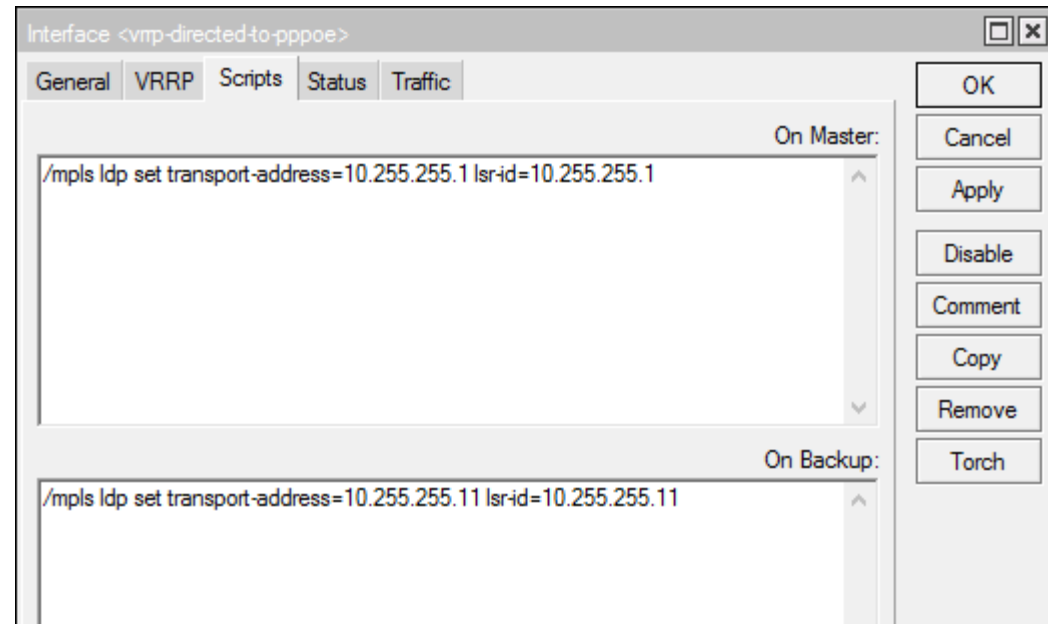
```
/interface vrrp
add interface=ether3 name=vrrp-directed-to-pppoe \
on-backup="/mpls ldp set transport-address=10.255.255.11 lsr-id=10.255.255.11" \
on-master="/mpls ldp set transport-address=10.255.255.1 lsr-id=10.255.255.1" \
preemption-mode=no vrid=5
```

R1-Main: 10.255.255.11

R1-Backup: 10.255.255.12

Note: Change of LSR ID

Service affecting





Traffic improvement

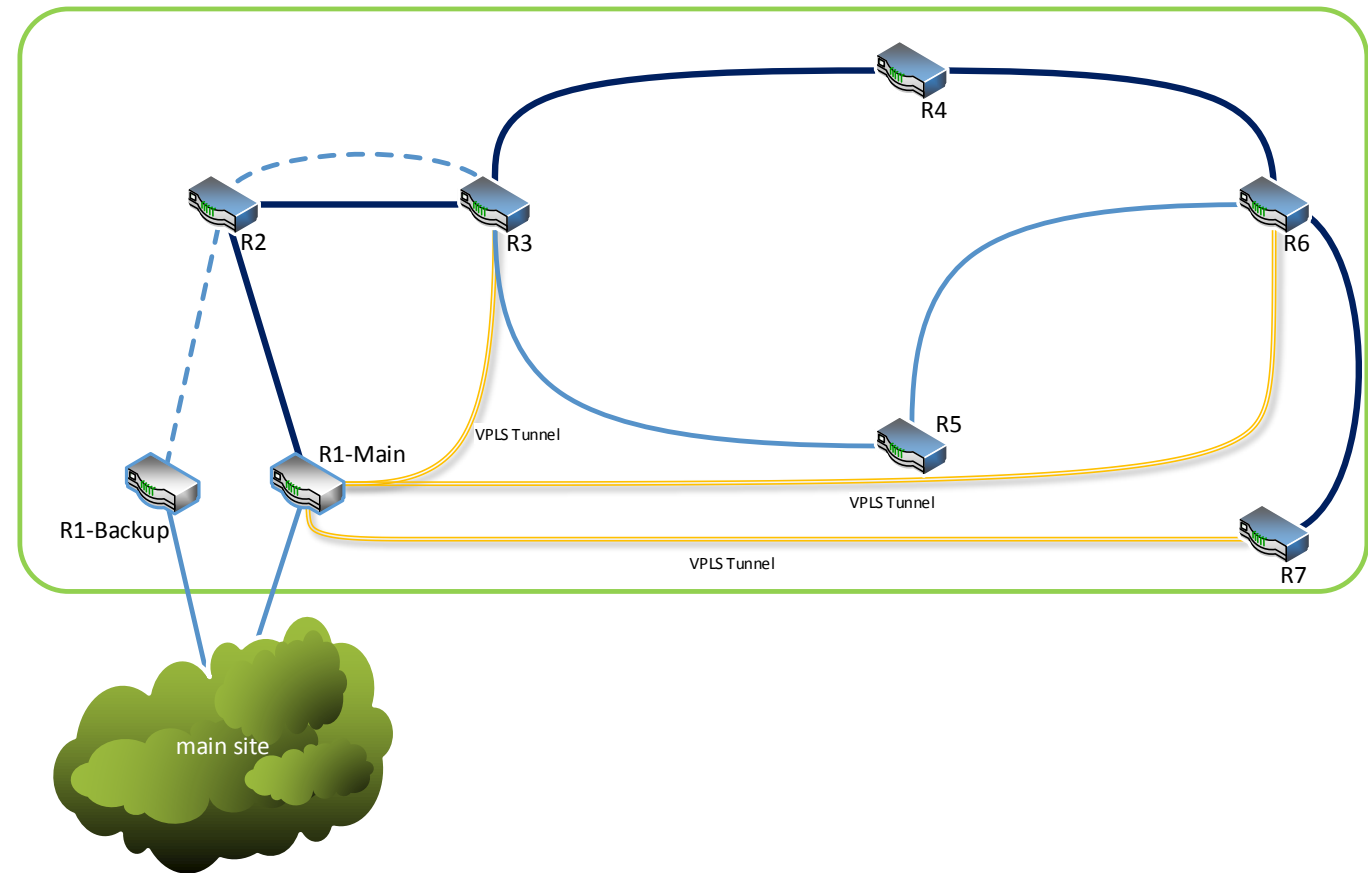


Use backup link

- Traffic from R7 to R1 through R4

But:

- OSPF goes through R5
- MPLS goes through R5
- VPLS goes through R5





Traffic engineering (TE) tunnel

- Enable TE support on all involved interfaces

For example on R3:

```
/mpls traffic-eng interface
```

```
add interface=ether3
```

```
add interface=ether4
```

```
add interface=ether5
```

(Compare with MPLS interfaces)

The screenshot shows a configuration window titled "TE Interface <ether3>". The interface is set to "ether3". The configuration parameters are as follows:

Interface:	ether3	OK
Bandwidth:	0 bps	Cancel
K Factor:	3	Apply
Resource Class:	0 Hex	Disable
Refresh Time:	30.000	Copy
<input type="checkbox"/> Use UDP		Remove
Blockade K Factor:	3	
TE Metric:	1	
IGP Flood Period:	00:03:00	
Up Flood Thresholds:	15,30,45,60,75,80,85,90,95,97,98,99,100	
Down Flood Thresholds:	15,30,45,60,75,80,85,90,95,97,98,99,100	
Remaining Bw.:	0 bps	

enabled



Traffic engineering (TE) tunnel

Use TE tunnel.

Here:

- No need for OSPF adjustments / single OSPF area
- No need for bandwidth reservation / definition
- No need for Constrained Shortest Path First (CSPF)



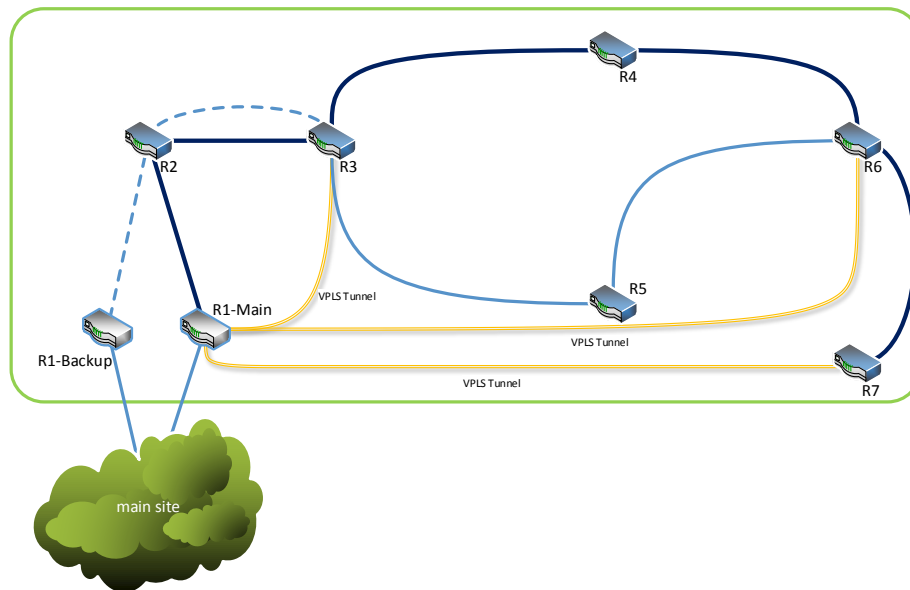
Traffic engineering (TE) tunnel

- Configure primary and secondary tunnel path (R1, R7)

```
/mpls traffic-eng tunnel-path
```

```
add name=tunnel-path-via-r4 use-cspf=no hops=10.255.255.4:loose
```

```
add name=dynamic-path use-cspf=no
```



Tunnel Path <dynamic-path>

Name:

Use CSPF

Setup Priority:

Holding Priority:

Record Route:

Affinity Include All:

Affinity Include Any:

Affinity Exclude:

Reoptimize Interval:

Hops:

enabled

Tunnel Path <tunnel-path-via-r4>

Name:

Use CSPF

Setup Priority:

Holding Priority:

Record Route:

Affinity Include All:

Affinity Include Any:

Affinity Exclude:

Reoptimize Interval:

Hops: :

enabled



Traffic engineering (TE) tunnel

- Create TE Tunnel (R1, R7)

```
/interface traffic-eng add \  
name=traffic-eng-to-r7 \  
from-address=10.255.255.1 \  
to-address=10.255.255.7 \  
primary-path=tunnel-path-via-r4 \  
secondary-paths=dynamic-path
```

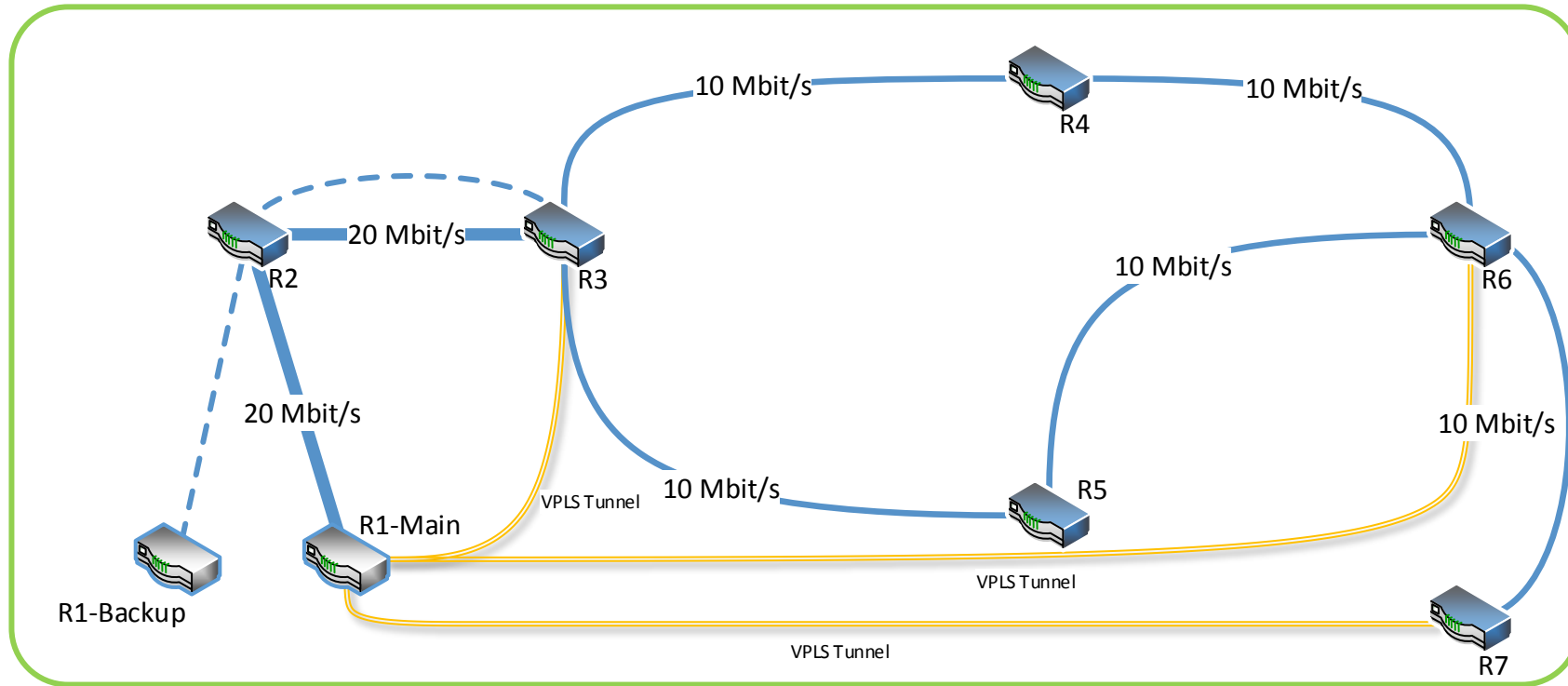
The screenshot shows the configuration window for an interface named 'traffic-eng-to-r7'. The window has tabs for 'General', 'TE', 'Bandwidth', 'Status', and 'Traffic'. The 'TE' tab is active. The configuration fields are as follows:

- Name: traffic-eng-to-r7
- Type: Traffic Eng Interface
- MTU: 1500
- Actual MTU: 1500
- L2 MTU: 65535
- From Address: 10.255.255.1
- To Address: 10.255.255.7
- Bandwidth: 0
- Primary Path: tunnel-path-via-r4
- Secondary Paths: dynamic-path

At the bottom of the window, there are three status indicators: 'enabled', 'running', and 'slave'. On the right side, there are buttons for 'OK', 'Cancel', 'Apply', 'Disable', 'Comment', 'Copy', 'Remove', and 'Torch'.



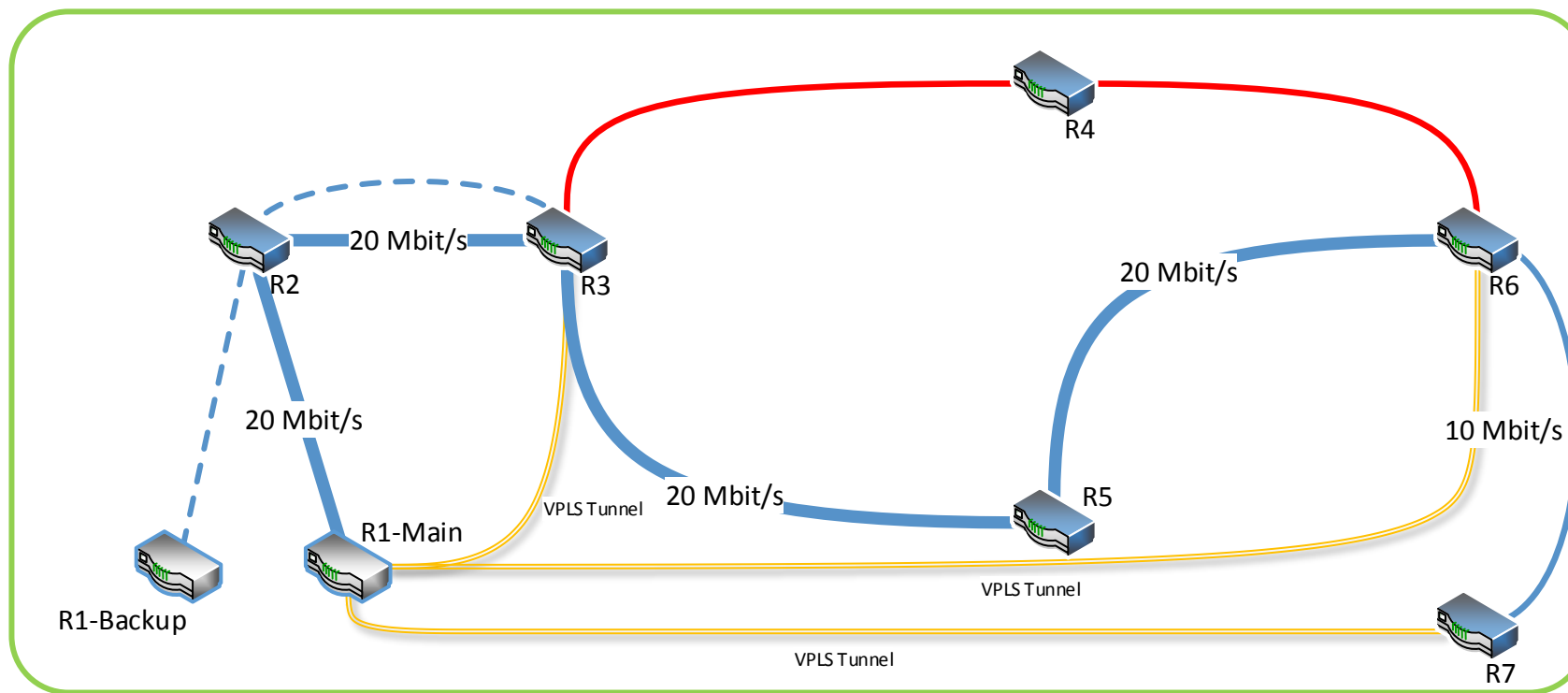
Result



10 Mbit/s to PPPoE client at R6 and R7



Result



10 Mbit/s to PPPoE client at R6 and R7

Failure of R4: Traffic through R5 (same for R5)



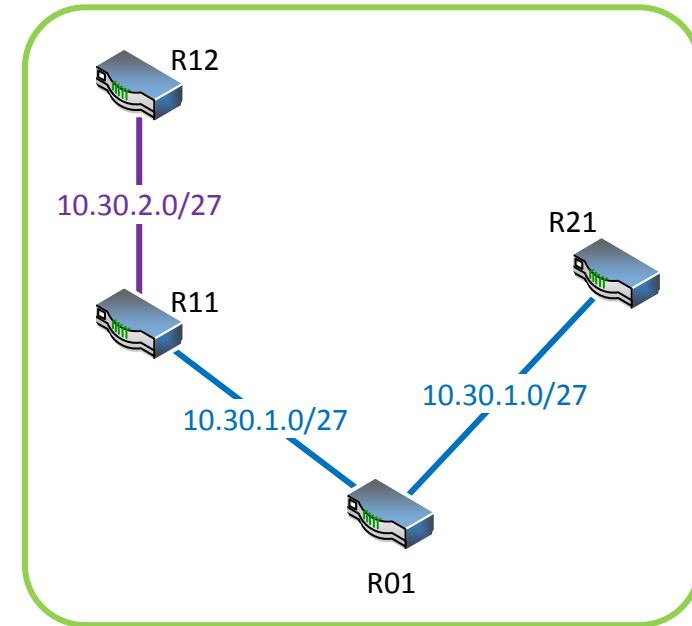
OSPF issue



OSPF setup (simplified)

R01, R11 and R21 on same subnet

- Bridge on R01
- Same horizon value
- R01 OSPF neighbors: R11, R21



#	Interface	Bridge	Horizon	Priority (h...)	Path Cost	Role	Root Path Cost
0	ether5	bridge	5	80	10	designated port	
1	ether4	bridge	5	80	10	designated port	

Instance	Router ID	Address	Interface	State Changes	State
default	10.10.10.21	10.30.1.21	bridge	11	Full
default	10.10.10.11	10.30.1.11	bridge	11	Full



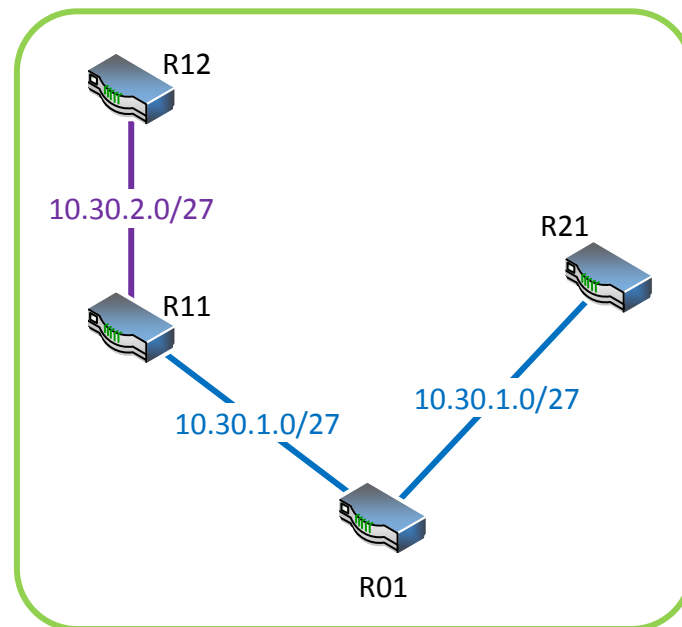
OSPF setup (simplified)

Expected behaviour on R21

- OSPF neighbour (only) R01
- Route to 10.30.2.0/27

Observed behaviour

- As expected



Instance	Router ID	Address	Interface	DR Address	State Changes	State
default	10.10.10.1	10.30.1.1	ether5	10.30.1.1	5	Full

1 item

	Dst. Address	Gateway	Distance
DAC	10.10.10.21	loopback reachable	0
DAC	10.30.1.0/27	ether5 reachable	0
DAo	10.30.2.0/27	10.30.1.11 reachable ether5	110

3 items



OSPF setup (simplified)

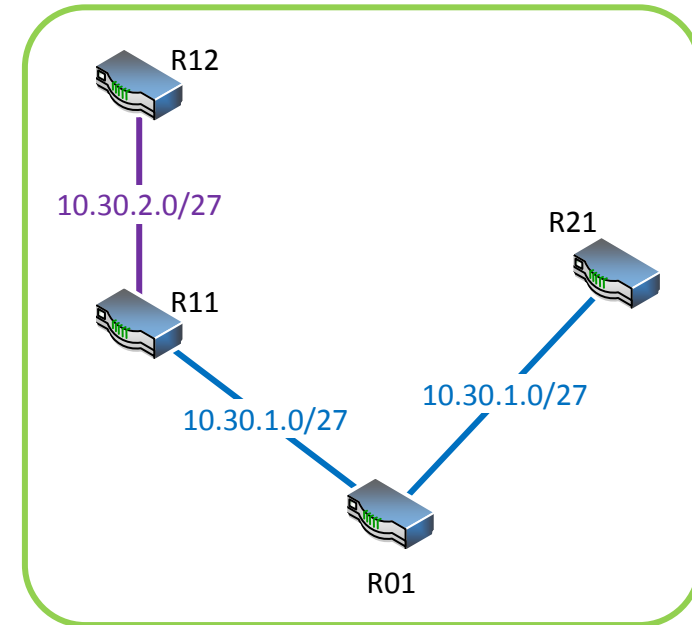
Reboot R01. No config change.

Expected behaviour on R21

- OSPF neighbour (only) R01
- Route to 10.30.2.0/27

Observed behaviour

- 10.30.2.0/27 missing





Debug R21

Debug R21

- OSPF state to R01 full
- 10.30.2.0/27 missing

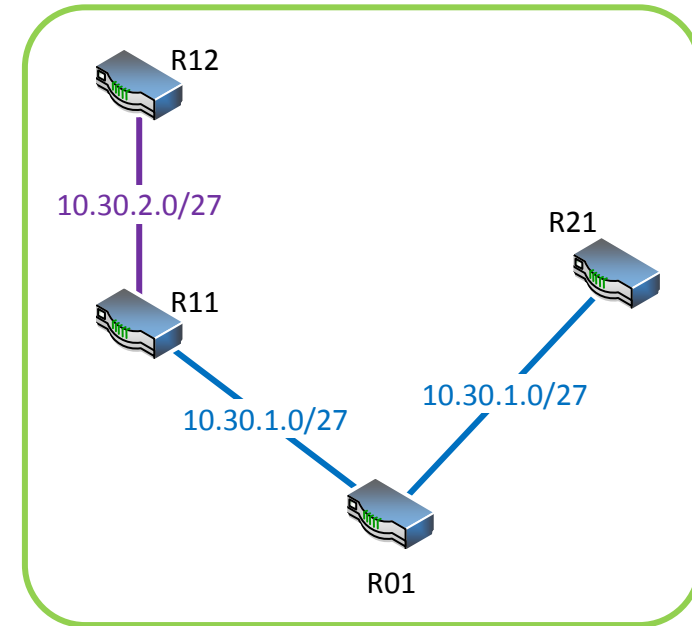
Route List

Routes Nexthops Rules VRF

+ - ✓ ✗ [Find] all

	Dst. Address	Gateway	Distance
DAC	▶ 10.10.10.21	loopback reachable	0
DAC	▶ 10.30.1.0/27	ether5 reachable	0

2 items



OSPF

Virtual Links Neighbors NBMA Neighbors Sham Links LSA Routes AS Border Routers Area Border Routers ...

[Find]

Instance	Router ID	Address	Interface	DR Address	State Changes	State
default	10.10.10.1	10.30.1.1	ether5	10.30.1.21	5	Full

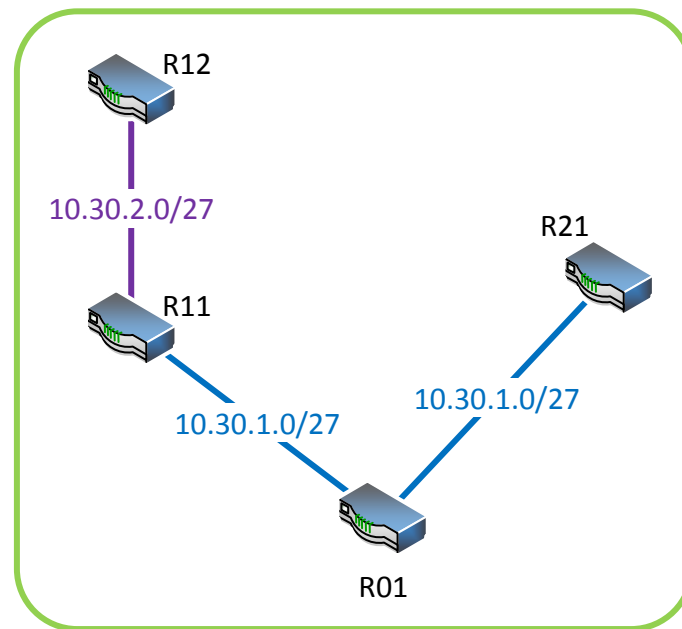
1 item



Debug R01

Debug R01

- OSPF state to R11 & R21 full
- 10.30.2.0/27 missing



Route List			
Routes			
Nexthops			
Rules			
VRF			
+	-	✓	✗
🔍	Find	all	▼
	Dst. Address	Gateway	Distance
DAC	▶ 10.10.10.1	loopback reachable	0
DAC	▶ 10.30.1.0/27	bridge reachable	0
2 items			

OSPF					
Virtual Links					
Neighbors					
NBMA Neighbors					
Sham Links					
LSA					
Routes					
AS Border Routers					
🔍	Find				
Instance	Router ID	Address	Interface	State Changes	State
🌈 default	10.10.10.11	10.30.1.11	bridge	6	Full
🌈 default	10.10.10.21	10.30.1.21	bridge	11	Full
2 items					

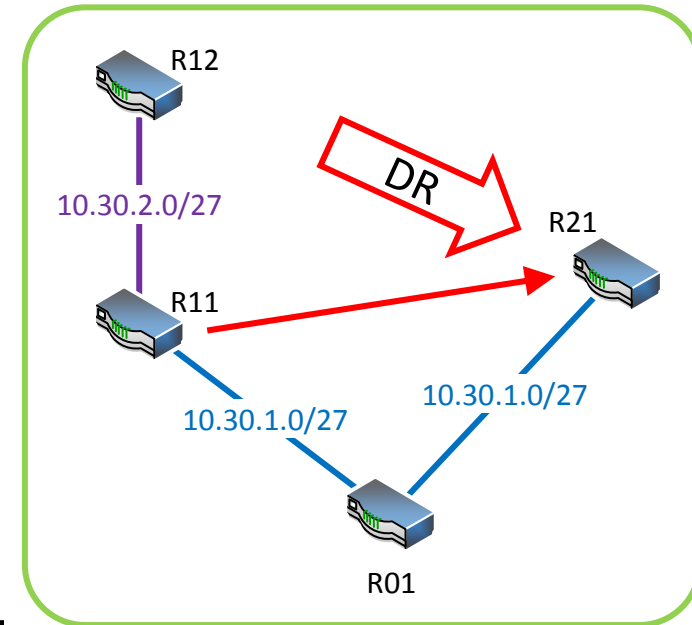


OSPF Designated Router

OSPF with network type Broadcast will elect Designated Router (DR).

Who is DR? R21 is DR!

- R11 tries to update R21 - not allowed
 - by bridge horizon
 - or wireless default forward
 - or bridge filter
 - ...

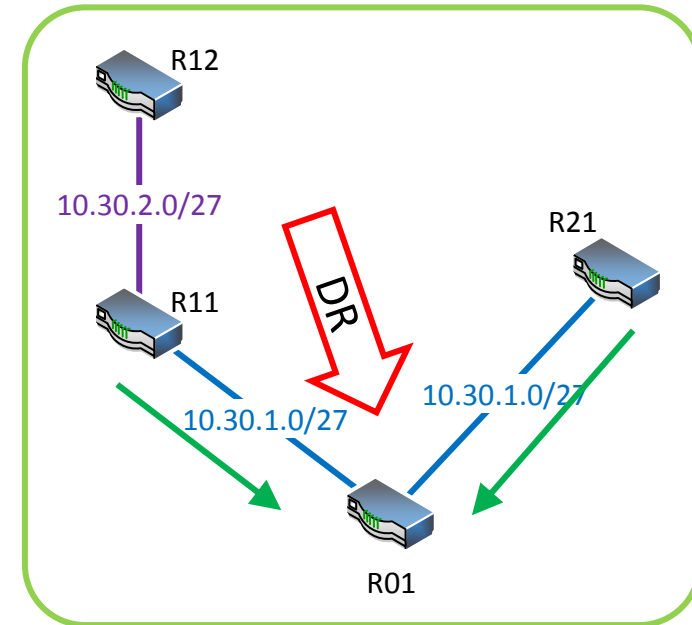




Possibilities

Possible solutions

- Force R01 to be DR
- Use network type ptmp





Thank you



FMS Internetservice GmbH

Phone: +49 761 2926500

Web: www.fmsweb.de

Shop: www.mikrotik-shop.de

Email: sales@fmsweb.de

Twitter: https://twitter.com/fmsweb_de