



MikroTik RouterOS v7:
Layer 3 Deep Dive

Overview of transition to ROSv7 and
updates to routing protocols

Quick Notes Before We Get Started

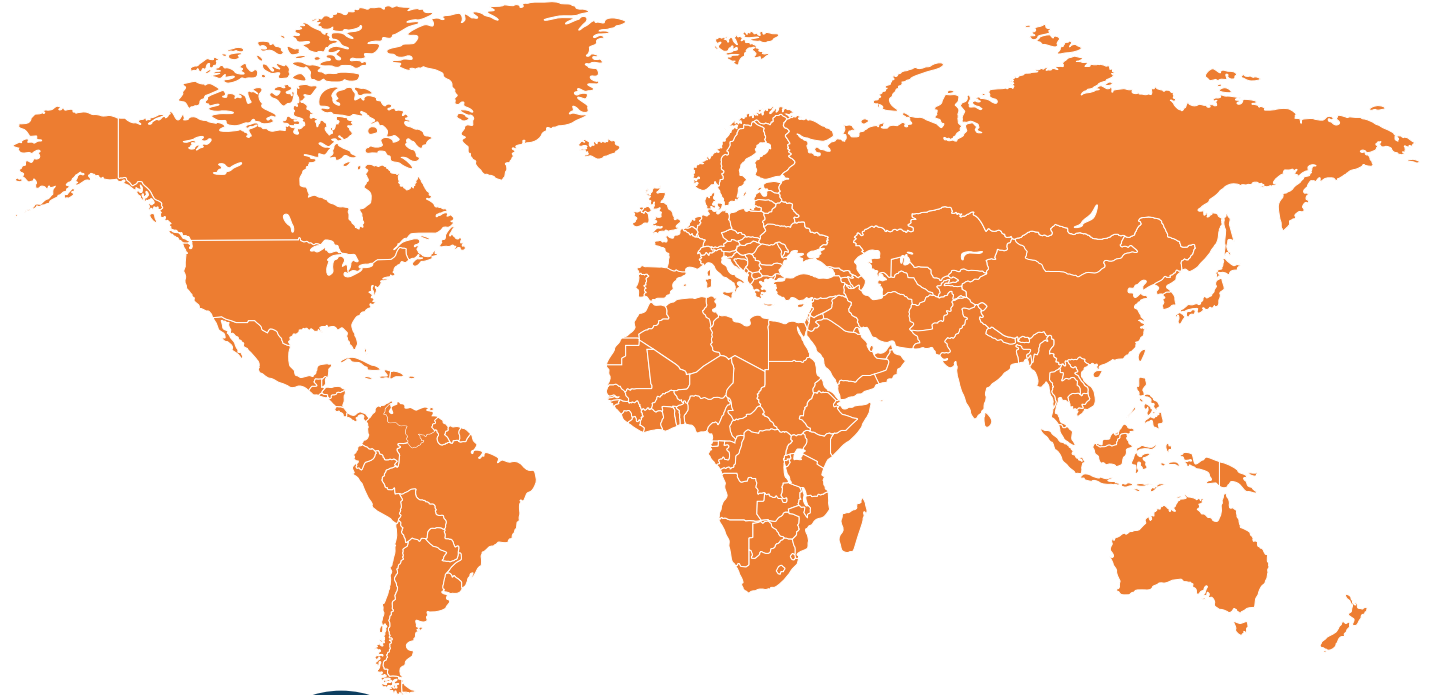
- Ask Questions!
 - In Zoom's Webinar client, you should see a Q&A button to pose questions. If we are able to address them, I will interrupt Kevin and we're happy to take a slight detour.
 - If we are unable to address them live, they influence the content and lab work we do in the future, **so please ask anyway!**
- This will be a series of webinars on Router OS 7. Today's topic is Layer 3 focused, if your particular Router OS 7 interest isn't covered today, it will likely be covered in depth in a future webinar.
- A recording and copy of the slides will be sent to registrants next week. We will post the full webinar and slides publically a short time after that.

IP ArchiTechs

The leading independent
Network consulting group

Unlike VARs, we are completely
focused on design and protocols
to meet business goals. Not
hardware sales.

The largest MikroTik consulting
firm globally.



Offices in the US,
Europe and South America



100+ Active Projects
1000s of clients across 6 continents

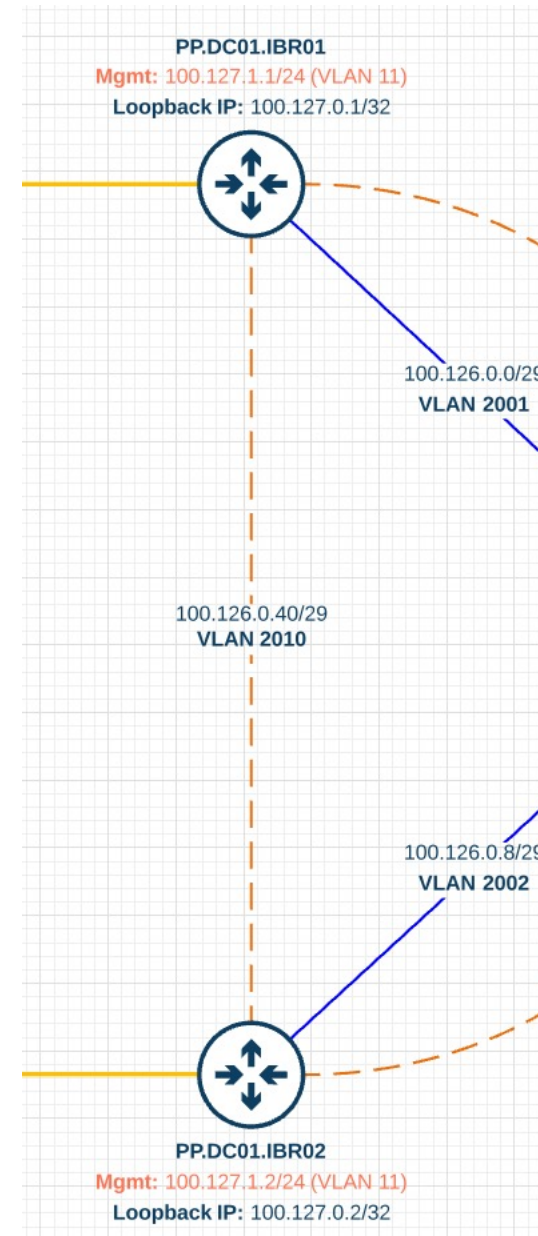
IPA | What we do?

Here are some typical engagement contexts our clients come to us with.

- Network Discovery/Audit – Consultative review of observations and findings
- Strategy consultation on network growth cycles. (Bridged -> Routed -> Basic Dynamic Routing -> and beyond)
- Systems/Software Integration and Automation
- Network migration planning, simulation, execution. We can take the driver seat and get things done that you otherwise don't have time for.
- Break/Fix Troubleshooting and Incident Response

We'll work with you to identify the best path forward, considering business goals. We can help strategize, plan, configure, and implement.

The operational tempo of our team has us uniquely trained to plan and lead executions of network migrations at any size.



2001:db8::/32	Active
• 2001:db8:10::/44	Active
• • 2001:db8:10::/56	Reserved

```
interfaces
  lo0 {
    unit 0 {
      family inet {
        address 127.0.0.1/8
        address 10.255.4.1/24
        primary;
      }
    }
    family iso {
      address 47.0005.8000.0000.0000.0000.0000.0000
    }
    family inet6 {
      address abcd::10:2
      primary;
    }
  }
}
```




MikroTik RouterOS v7:
Layer 3 Deep Dive

Overview of transition to ROSv7 and
updates to routing protocols

ROSv7 – Routing and L3 forwarding

- 01** ROSv7 basics, lab setup and `/routing/route/`
- 02** BGP and OSPF
- 03** L3 hardware offload



MikroTik RouterOS v7:
Layer 3 Deep Dive

ROSv7 basics and lab setup

Section 1 | ROSv7 basics

01

- | ROSv7 overview
- | Lab setup
- | /routing/route
- | What about RIP?

S1.1 | ROSv7 overview

- **Linux Kernel Upgrade**
 - Router OS v6 = 3.3.5 (May 2012)
 - Router OS v7 = 5.6.3 (April 2020)
- **8 Years Of Progress**
 - Major work on IPv6
 - Hardware Offloading
 - Network Namespaces
 - Netfilter
 - Wireguard (Great example of leveraging "free" features of the kernel.)
 - Packet Scheduler and Queue Disciplines
 - VxLAN and other modern protocol improvements (IPv6 Segment Routing for example)
 - 100G improvements throughout Kernel 4.x.

Countless other improvements across the entire networking stack over the busiest era in Linux's ~30 year history from 1000s of contributors.

How much MikroTik leverages underlying functionality of the Linux kernel vs how much is implemented by MikroTik themselves is a bit of speculation.

8 years ago, it might've made sense to write custom implementations which is more onerous to implement new functionality, protocols, and RFCs.

We hope with this major kernel upgrade, MikroTik is able to leverage the underlying kernel more and expend its resources on other items.



S1.1 | ROSv7 overview

Apparent Focus On Granular Process/CPU Scheduler

```
[zuul@rtr-edge-01.v7.ipa.dev] /routing/stats> step/print
Flags: R - RUNNING
Columns: NAME, CONTEXT, PROCID, ORDER, RUNS, TARGETS, MAX-TIME, CUR-TIME, STATE, SCHED
# NAME CONTEXT PROCID ORDER RUNS TAR
23 publish mark MARK_DSTS mpls 0 0
24 publish mark MARK_DSTS mpls 0 0
25 publish mark MARK_DSTS mpls 0 0
26 publish mark MARK_DSTS 26 0 0
```

Projecting based on new menus, tools, and information exposed in Router OS 7, MikroTik seems to be taking a increased focused on CPU, process, and scheduler management. This may also explain the seeming convergence to ARM based chipsets on newer devices.

```
BGP Output 200:1ce:127::3
11 BGP Input 100.127.1.3 512.0KiB 512.0KiB 0 0 0 1 27 4552 1 2
BGP Output 100.127.1.3
12 Global memory 256.0KiB global 0 0
[zuul@rtr-edge-01.v7.ipa.dev] /routing/stats> process/print
Columns: TASKS, PRIVATE-MEM-BLOCKS, SHARED-MEM-BLOCKS, PSS, RSS, VMS, RETIRED, ID, PID,
USY, CUR-CALC, MAX-CALC
# TASKS PRIVATE- SHARED-ME PSS R V R ID PID R P
```

S1.1 | ROS v7 overview

New features

New Kernel is implemented in RouterOSv7, that lead to performance changes due to route cache, as well some tasks might require higher CPU and RAM usage for different processes.

- completely new NTP client and server implementation
- merged individual packages, only bundle, and a few extra packages remain
- new Command Line Interface (CLI) style (RouterOS v6 commands are still supported)
- support for Let's Encrypt certificate generation
- support for REST API
- support for UEFI boot mode on x86
- CHR FastPath support for "vmxnet3" and "virtio-net" drivers
- support for "Cake" and "FQ_Codel" type queues
- support for IPv6 NAT
- support for Layer 3 hardware acceleration on all CRS3xx devices
- support for MBIM driver with basic functionality support for all modems with MBIM mode
- support for MLAG on CRS3xx devices
- support for VRRP grouping and connection tracking data synchronization between nodes
- support for Virtual eXtensible Local Area Network (VXLAN)
- support for L2TPv3
- support for OpenVPN UDP transport protocol
- support for WireGuard
- support for hardware offloaded VLAN filtering on RTL8367 (RB4011, RB100AHx4) and MT7621 (hEX, hEX S, RBM33G) switches
- support for ZeroTier on ARM and ARM64 devices
- completely new alternative wireless package "wifiiwave2" with 802.11ac Wave2, WPA3, and 802.11w management frame protection support (requires ARM CPU and 256MB RAM)
- support for hardware offloaded VLAN filtering on RTL8367 (RB4011, RB100AHx4) and MT7621 (hEX, hEX S, RBM33G) switches

S1.2 | Lab setup - why hardware vs. CHR?



- **Ability to test hardware specific features**

- Hardware specific features are now common in MikroTik CRS and CCR2K platforms
- Bridging and L2 – while possible to test some aspects on a software bridge in a CHR, most of these features require an ASIC to fully test HW offload and interop with other L2 devices
- L3 hardware offload – Can only be tested on an actual router

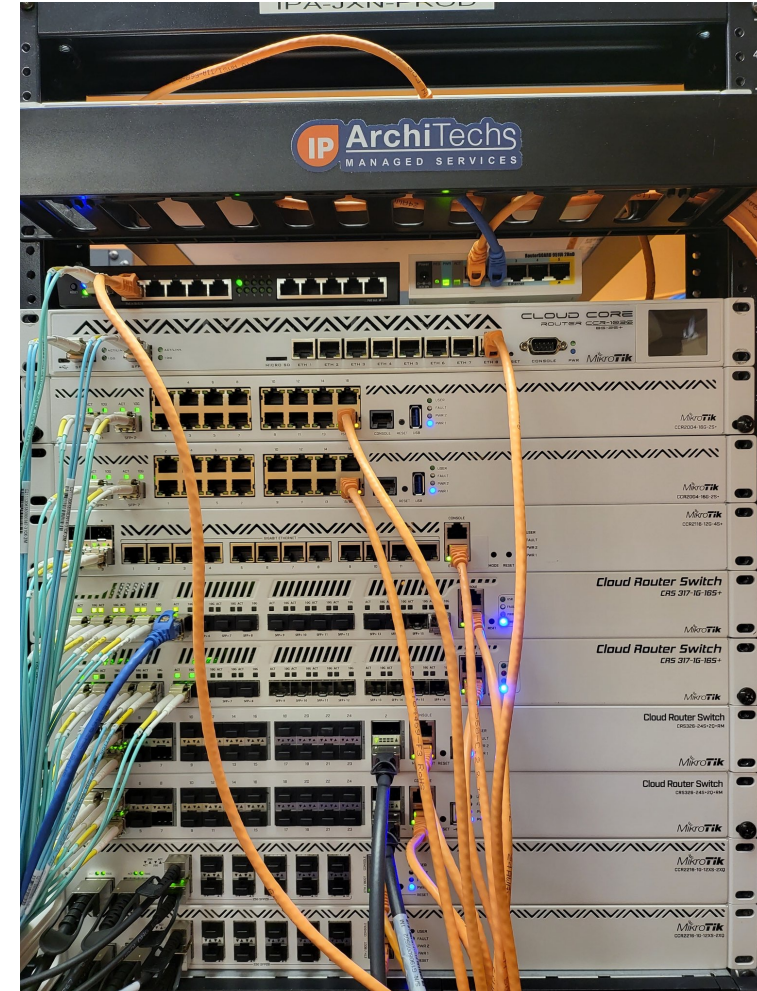
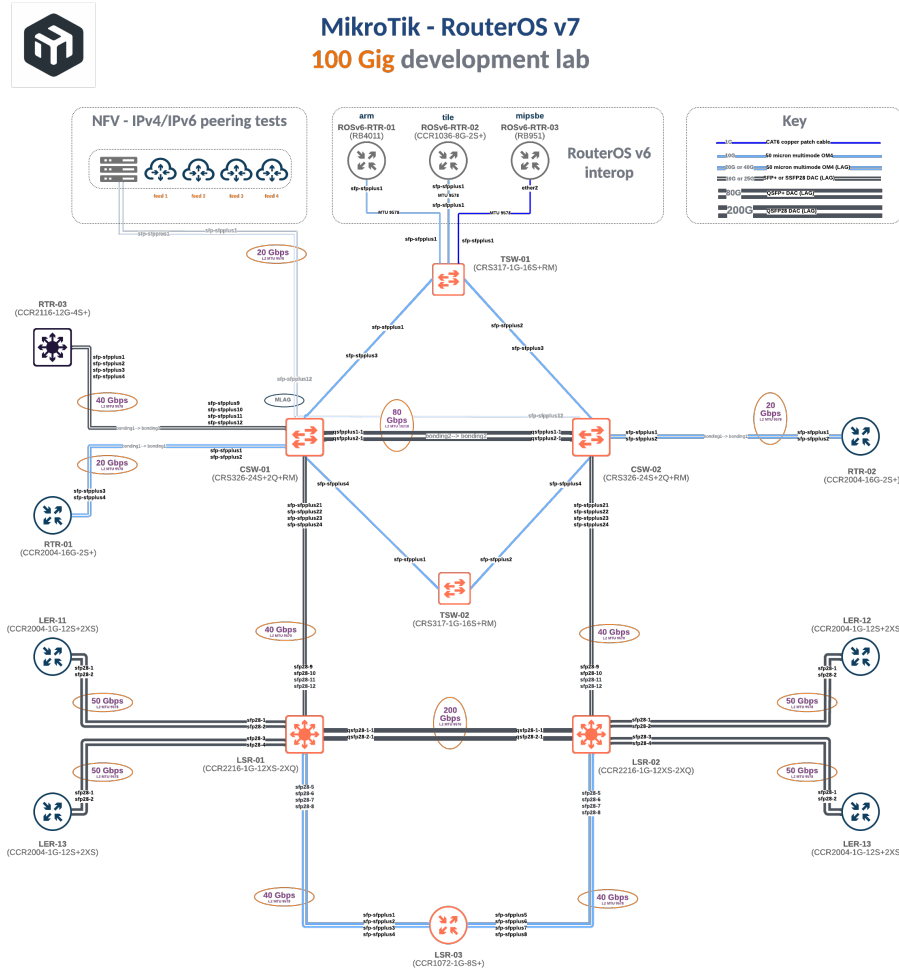
- **Stability**

- Because we are in a transition between ROSv6 and ROSv7, using hardware likely to go in prod helps to validate design and protocol combinations

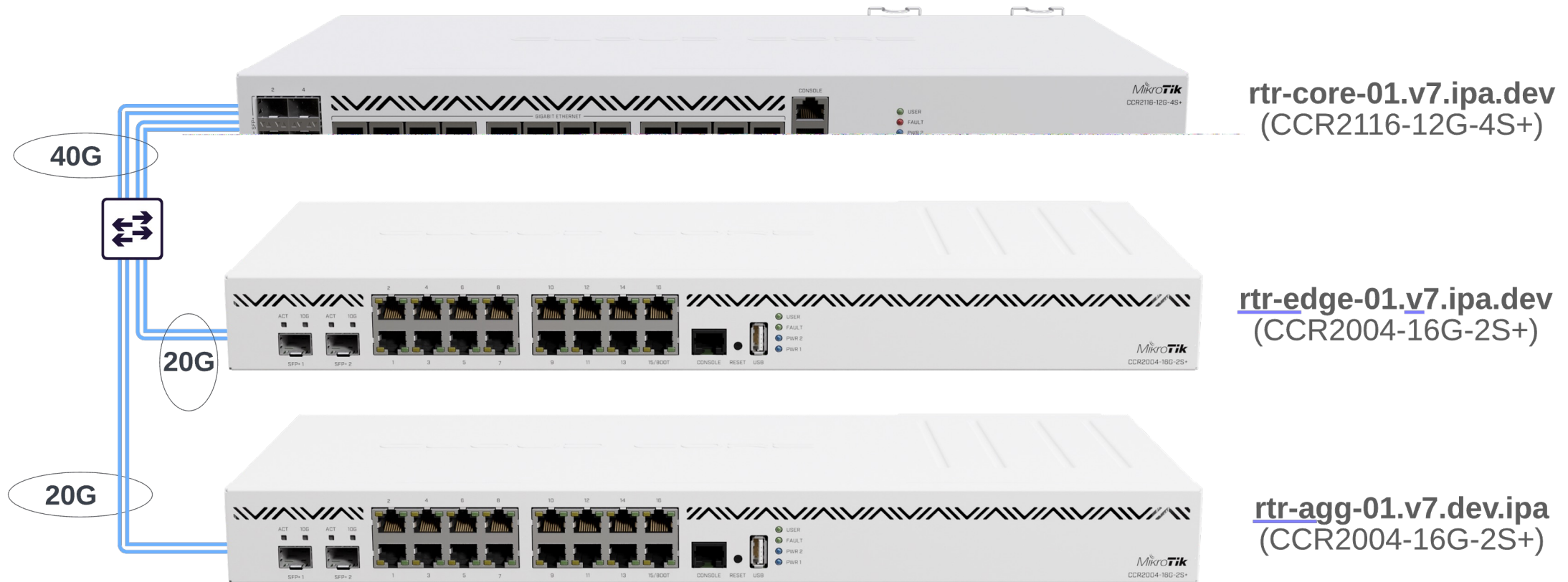
- **Bug hunting**

- While ROSv7 stability has improved significantly since early 2022, there are still bugs to be found and labbing common use cases helps us to find those and report them to MikroTik.

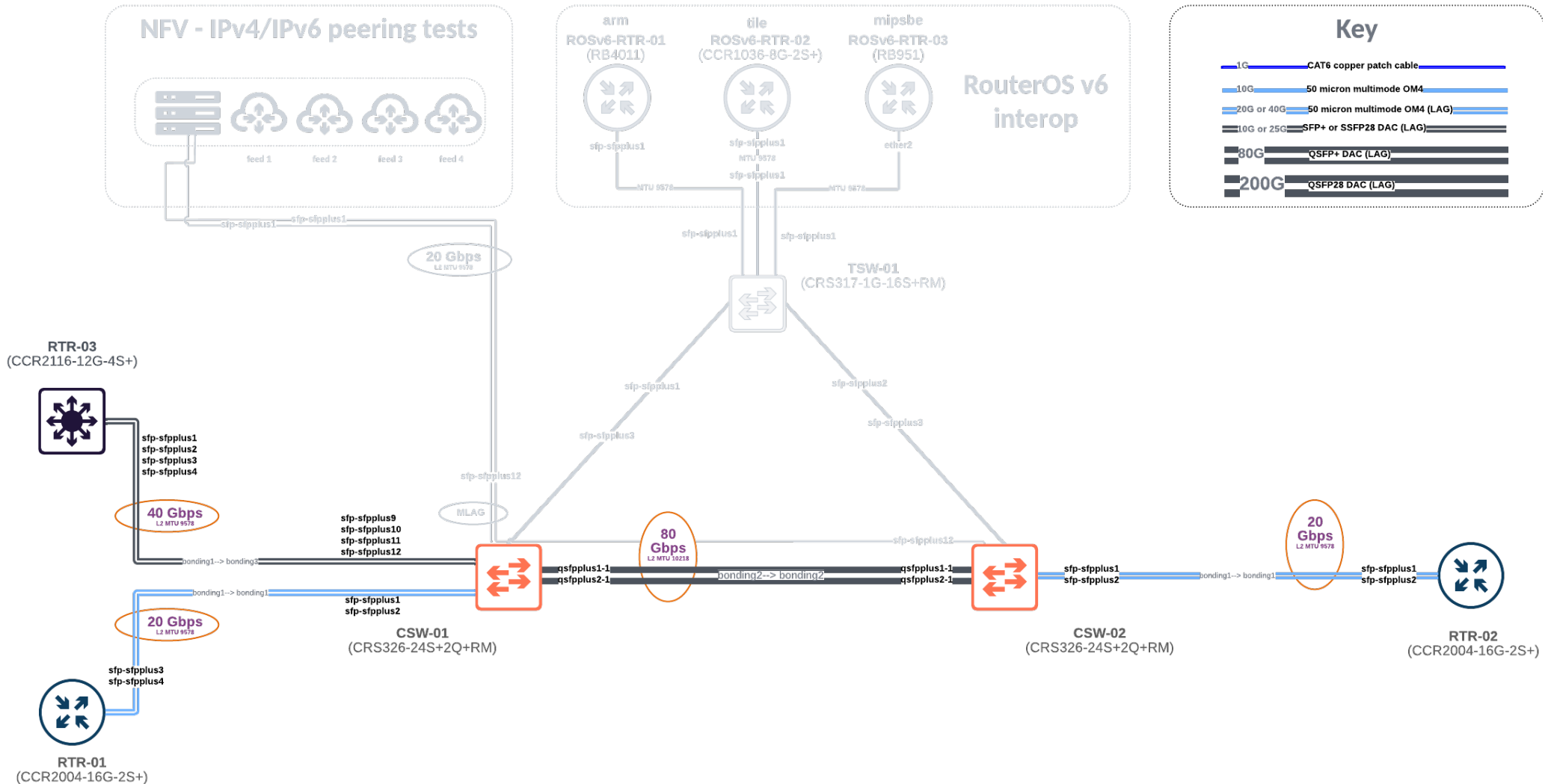
S1.2 | Lab setup – 100G development lab



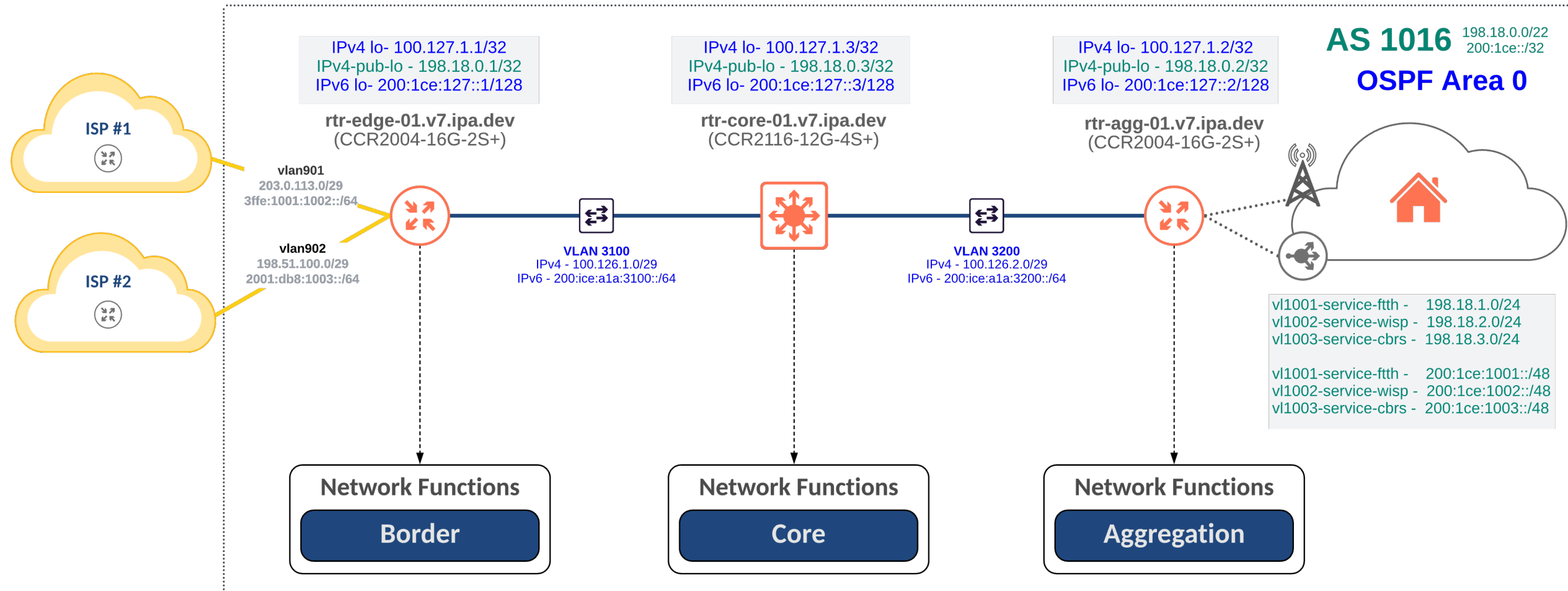
S1.2 | Lab setup – physical



S1.2 | Lab setup – specific device focus



S1.2 | Logical lab overview



S1.3 | /routing/route - overview

What is /routing/route ?

- A read-only table that lists routes from all the address families as well as all filtered routes with all possible route attributes.
- **/ip/route** and **/ipv6/route** still exist but have less information overall compared to **/routing/route**

```
[zuul@rtr-edge-01.v7.ipa.dev] > routing/route/print
Flags: U, A - ACTIVE; c, s, b, o, d, y - COPY; H - HW-OFFLOADED; +, B - BLACKHOLE
Columns: DST-ADDRESS, GATEWAY, AFI, DISTANCE, SCOPE, TARGET-SCOPE, IMMEDIATE-GW
DST-ADDRESS      GATEWAY          AFI  DISTANCE  SCOPE  TARGET-SCOPE  IMMEDIATE-GW
As + 0.0.0.0/0     198.51.100.1     ip4   1         30     10          198.51.100.1%vlan902
As + 0.0.0.0/0     203.0.113.1      ip4   1         30     10          203.0.113.1%vlan901
Ac 100.126.1.0/29   vlan3100         ip4   0         10     10          vlan3100
Ao 100.126.1.8/29  100.126.1.2%vlan3100 ip4   110      20     10          100.126.1.2%vlan3100
Ac 100.127.1.1/32  lo-ipv4          ip4   0         10     10          lo-ipv4
Ao 100.127.1.2/32  100.126.1.2%vlan3100 ip4   110      20     10          100.126.1.2%vlan3100
Ao 100.127.1.3/32  100.126.1.2%vlan3100 ip4   110      20     10          100.126.1.2%vlan3100
Ac 172.16.77.0/24  vlan777          ip4   0         10     10          vlan777
Ac 172.16.88.0/24  vlan888          ip4   0         10     10          vlan888
b 198.18.0.1/32    100.127.1.3      ip4   200      40     30          100.126.1.2%vlan3100
Ac 198.18.0.1/32  lo-ipv4-pub      ip4   0         10     10          lo-ipv4-pub
Ab 198.18.1.0/24   100.127.1.2      ip4   200      40     30          100.126.1.2%vlan3100
Ab 198.18.2.0/24   100.127.1.2      ip4   200      40     30          100.126.1.2%vlan3100
Ac 198.51.100.0/29 vlan902           ip4   0         10     10          vlan902
Ac 203.0.113.0/29 vlan901           ip4   0         10     10          vlan901
Ad 0.0.0.0/0       10.255.44.1@vrf-mgmt ip4   1         30     10          10.255.44.1%ether15
Ac 10.255.44.0/22 ether15@vrf-mgmt ip4   0         10     10          ether15
UoH 200:1ce:127::1/128 lo-ipv6          ip6   110      20     10          lo-ipv6
Ac 200:1ce:127::1/128 lo-ipv6          ip6   0         10     10          lo-ipv6
Ao 200:1ce:127::2/128 fe80::de2c:6eff:fe7a:d0e6%vlan3100 ip6   110      20     10          fe80::de2c:6eff:fe7a:d0e6%vlan3100
Ao 200:1ce:127::3/128 fe80::de2c:6eff:fe7a:d0e6%vlan3100 ip6   110      20     10          fe80::de2c:6eff:fe7a:d0e6%vlan3100
UoH 200:1ce:a1a:3100::/64 vlan3100         ip6   110      20     10          vlan3100
Ac 200:1ce:a1a:3100::/64 vlan3100         ip6   0         10     10          vlan3100
Ac 200:1ce:a1a:3100::/64 fe80::de2c:6eff:fe7a:d0e6%vlan3100 ip6   110      20     10          fe80::de2c:6eff:fe7a:d0e6%vlan3100
```

S1.3 | /routing/route - validation

Provides very detailed information when used with “where” command

- `routing/route/print detail where bgp && afi=ip4`
- Shows BGP communities, local pref, peer learned from, etc
- `Debug.fwp-ptr` attribute allows for debugging of the routing table by MikroTik

```
zuul@rtr-edge-01.v7.ipa.dev] > routing/route/print detail where bgp && afi=ip4
lags: X - disabled, F - filtered, U - unreachable, A - active;
- connect, s - static, r - rip, b - bgp, o - ospf, d - dhcp, v - vpn, m - modem, a - ldp-address, l - ldp-mapping, y - copy;
- hw-offloaded; + - ecmp, B - blackhole
b   afi=ip4 contribution=best-candidate dst-address=198.18.0.1/32 routing-table=main gateway=100.127.1.3
    immediate-gw=100.126.1.2%vlan3100 distance=200 scope=40 target-scope=30 belongs-to="bgp-IP-100.127.1.3"
    bgp.peer-cache-id=*B000003 .large-communities=1016:1016:4 .local-pref=100 .atomic-aggregate=yes .origin=igp
    debug.fwp-ptr=0x20342D20

Ab  afi=ip4 contribution=active dst-address=198.18.1.0/24 routing-table=main gateway=100.127.1.2
    immediate-gw=100.126.1.2%vlan3100 distance=200 scope=40 target-scope=30 belongs-to="bgp-IP-100.127.1.3"
    bgp.peer-cache-id=*B000003 .local-pref=100 .atomic-aggregate=yes .origin=igp
    debug.fwp-ptr=0x20342D80

Ab  afi=ip4 contribution=active dst-address=198.18.2.0/24 routing-table=main gateway=100.127.1.2
    immediate-gw=100.126.1.2%vlan3100 distance=200 scope=40 target-scope=30 belongs-to="bgp-IP-100.127.1.3"
    bgp.peer-cache-id=*B000003 .local-pref=100 .atomic-aggregate=yes .origin=igp
    debug.fwp-ptr=0x20342D80
zuul@rtr-edge-01.v7.ipa.dev] > █
```


S1.3 | /routing/route -

Can be used to view IPv4 and IPv6 simultaneously

- `routing/route/print detail where ospf && active`
- Shows afi **ip4** and **ip6**

```
[zuul@rtr-edge-01.v7.ipa.dev] > routing/route/print detail where ospf && active
Flags: X - disabled, F - filtered, U - unreachable, A - active;
C - connect, S - static, R - rip, B - bgp, O - ospf, D - dhcp, V - vpn, M - modem, A - ldp-address, I - ldp-mapping, Y - copy;
H - hw-offloaded; + - ecmp, B - blackhole
Ao  afi=ip4 contribution=active dst-address=100.126.1.8/29 routing-table=main gateway=100.126.1.2%vlan3100
    immediate-gw=100.126.1.2%vlan3100 distance=110 scope=20 target-scope=10 belongs-to="ospf-instance-v2-ipv4"
    ospf.metric=2 .type=intra
    debug.fwp-ptr=0x20342CC0

Ao  afi=ip4 contribution=active dst-address=100.127.1.2/32 routing-table=main gateway=100.126.1.2%vlan3100
    immediate-gw=100.126.1.2%vlan3100 distance=110 scope=20 target-scope=10 belongs-to="ospf-instance-v2-ipv4"
    ospf.metric=3 .type=intra
    debug.fwp-ptr=0x20342CC0

Ao  afi=ip4 contribution=active dst-address=100.127.1.3/32 routing-table=main gateway=100.126.1.2%vlan3100
    immediate-gw=100.126.1.2%vlan3100 distance=110 scope=20 target-scope=10 belongs-to="ospf-instance-v2-ipv4"
    ospf.metric=2 .type=intra
    debug.fwp-ptr=0x20342CC0

Ao  afi=ip6 contribution=active dst-address=200:1ce:127::2/128 routing-table=main gateway=fe80::de2c:6eff:fe7a:d0e6%vlan3100
    immediate-gw=fe80::de2c:6eff:fe7a:d0e6%vlan3100 distance=110 scope=20 target-scope=10 belongs-to="ospf-instance-v3-ipv6"
    ospf.metric=2 .type=intra
    debug.fwp-ptr=0x20342C60
```

S1.4 | What about RIP?



- **Not used often**
 - Most ISP, DC and Enterprise networks moved away from RIP a while ago
 - Used in very specific roles like SatCom, Cable ISPs and Tunnel headends
- **Out of scope**
 - Since OSPF and BGP are the most popular, this presentation will focus on those protocols
- **You've been RIP Rolled ;)**



MikroTik RouterOS v7:
Layer 3 Deep Dive

BGP & OSPF for IPv4/IPv6

Section 2 | Routing Protocols

02

- | OSPF - What's new
- | OSPF - Overview
- | BGP – What's new
- | BGP – Overview

S2.1 | OSPF - What's new for v7?

Feature	v7.1	v7.2	v7.3	v7.6	v7.7
OSPF					
Convert OSPF config from v6 to v7 after upgrade	Known conversion problems: <ul style="list-style-type: none">• NBMA neighbors place in backbone• ospf-v2 networks + interface may have issues• dynamic interfaces may have issues• MPLS PE CE features are not converted				
OSPF neighbors in NSSA Area					
OSPF in broadcast network					
OSPF with routing filters					
OSPF Virtual Link					
OSPF input filtering					
HMAC-SHA auth RFC5709	N/A			Initial support	
BGP and OSPF SNMP monitoring	N/A				

Protocol overview for v7

- Use the feature matrix from help.mikrotik.com under **Pages / RouterOS / Routing** to track progress
- Config conversion typically works without issue now
- Roadmap includes SNMP monitoring for OSPF

S2.1 | OSPF - What's new for v7?

Single OSPF configuration for v2/v3

- Instances are built for v2 (IPv4) or v3 (IPv6)

```
[zuul@rtr-core-01.v7.ipa.dev] > routing/ospf/export
# sep/22/2022 00:48:01 by RouterOS 7.6beta6
# software id = HBGP-U3V8
#
# model = CCR2116-12G-4S+
# serial number = F19A0F2DF6C6
/routing ospf instance
add disabled=no name=ospf-instance-v3-ipv6 router-id=rid-main version=3
add disabled=no name=ospf-instance-v2-ipv4 router-id=rid-main
/routing ospf area
add disabled=no instance=ospf-instance-v3-ipv6 name=ospf-area-0-ipv6
add disabled=no instance=ospf-instance-v2-ipv4 name=ospf-area-0-ipv4
/routing ospf interface-template
add area=ospf-area-0-ipv6 disabled=no interfaces=lo-ipv6 networks=200:1ce:127::3/128 passive
add area=ospf-area-0-ipv4 disabled=no interfaces=vlan3100 type=ptp
add area=ospf-area-0-ipv4 disabled=no interfaces=vlan3200 type=ptp
add area=ospf-area-0-ipv4 disabled=no interfaces=lo-ipv4
add area=ospf-area-0-ipv6 disabled=no interfaces=vlan3100 type=ptp
add area=ospf-area-0-ipv6 disabled=no interfaces=vlan3200 type=ptp
[zuul@rtr-core-01.v7.ipa.dev]
```

S2.1 | OSPF - What's new for v7?

routing/id can be used to define a system wide router id

- Dynamic IDs are built for each vrf using the highest configured IPv4 address
- Works for BGP, RIP and LDP as well

```
[zuul@rtr-core-01.v7.ipa.dev] > routing/id/export
# sep/22/2022 00:40:42 by RouterOS 7.6beta6
# software id = HBGP-U3V8
#
# model = CCR2116-12G-4S+
# serial number = F19A0F2DF6C6
/routing id
add id=100.127.1.3 name=rid-main select-dynamic-id=only-vrf select-from-vrf=main
[zuul@rtr-core-01.v7.ipa.dev] > routing/id/print
Flags: D, I - INACTIVE
Columns: NAME, ID, DYNAMIC-ID, SELECT-DYNAMIC-ID, SELECT-FROM-VRF
#   NAME      ID          DYNAMIC-ID  SELECT-DYNAMIC-ID  SELECT-FROM-VRF
0 D main      100.127.1.3 only-vrf        main
1 D vrf-mgmt
2   rid-main  100.127.1.3 only-vrf        vrf-mgmt
                                only-vrf        main
[zuul@rtr-core-01.v7.ipa.dev] > █
```


S2.1 | OSPF - What's new for v7?

interface-template command combines ROSv6 interface and network menus

- OSPF network statements have been moved to the interface level and aren't always required – listing an interface will attempt to form a neighbor adjacency on all prefixes that are configured for the interface
- IPv6 loopbacks require the network statement for loopbacks to advertise properly
- **Behavior change:** Passive interface default for all interfaces will advertise all prefixes (ROSv6 required a network statement)

```
[zuul@rtr-core-01.v7.ipa.dev] > routing/ospf/interface-template/export
# sep/22/2022 18:04:49 by RouterOS 7.6beta6
# software id = HBGP-U3V8
#
# model = CCR2116-12G-4S+
# serial number = F19A0F2DF6C6
/routing ospf interface-template
add area=ospf-area-0-ipv6 disabled=no interfaces=lo-ipv6 networks=200:1ce:127::3/128 passive
add area=ospf-area-0-ipv4 disabled=no interfaces=vlan3100 type=ptp
add area=ospf-area-0-ipv4 disabled=no interfaces=vlan3200 type=ptp
add area=ospf-area-0-ipv4 disabled=no interfaces=lo-ipv4
add area=ospf-area-0-ipv6 disabled=no interfaces=vlan3100 type=ptp
add area=ospf-area-0-ipv6 disabled=no interfaces=vlan3200 type=ptp
[zuul@rtr-core-01.v7.ipa.dev] >
```

S2.1 | OSPF - What's new for v7?

OSPF network type: ptp-unnumbered

- Instances are built for v2 (IPv4) or v3 (IPv6)

type (*broadcast* | *nbma* | *ptp* | *ptmp* | *ptp-unnumbered* | *virtual-link*;
Default: **broadcast**)

the OSPF network type on this interface. Note that if interface configuration does not exist, the default network type is 'point-to-point' on PtP interfaces and 'broadcast' on all other interfaces.

- broadcast - network type suitable for Ethernet and other multicast capable link layers. Elects designated router
- nbma - Non-Broadcast Multiple Access. Protocol packets are sent to each neighbor's unicast address. Requires manual configuration of neighbors. Elects designated router
- ptp - suitable for networks that consist only of two nodes. Do not elect designated router
- ptmp - Point-to-Multipoint. Easier to configure than NBMA because it requires no manual configuration of neighbor. Do not elect a designated router. This is the most robust network type and as such suitable for wireless networks, if 'broadcast' mode does not work well enough for them
- ptp-unnumbered - works the same as ptp, except that the remote neighbor does not have an associated IP address to a specific PTP interface. For example, in case an IP unnumbered is used on Cisco devices.
- virtual-link - for virtual link setups.

S2.2 | OSPF – ROSv6 vs. ROSv7 overview

OSPF Configuration

OSPFv3 and OSPFv2 are now merged into one single menu `/routing ospf`. At the time of writing this article, there are no default instances and areas.

To start both OSPFv2 and OSPF v3 instances, first, you need to create an instance for each and then add an area to the instance.

```
/routing ospf instance
add name=v2inst version=2 router-id=1.2.3.4
add name=v3inst version=3 router-id=1.2.3.4
/routing ospf area
add name=backbone_v2 area-id=0.0.0.0 instance=v2inst
add name=backbone_v3 area-id=0.0.0.0 instance=v3inst
```

At this point, you are ready to start OSPF on the network interface. In the case of IPv6, you add either interface on which you want to run OSPF (the same as ROSv6) or the IPv6 network. In the second case, OSPF will automatically detect the interface. Here are some interface configuration examples:

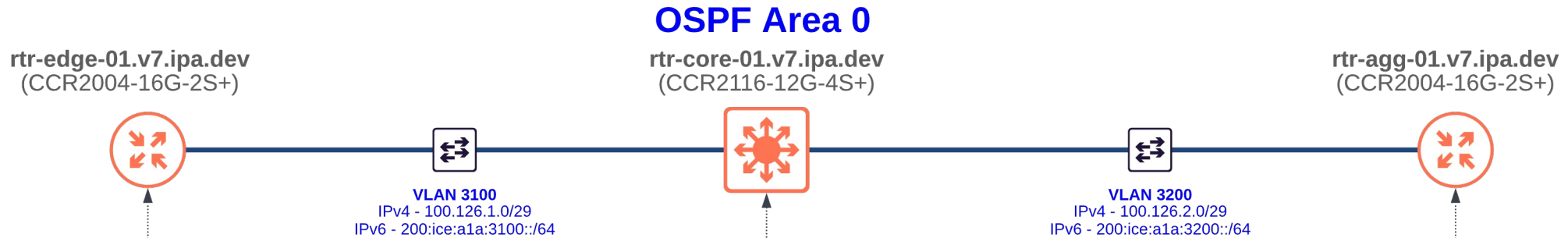
```
/routing ospf interface-template
add network=192.168.0.0/24 area=backbone_v2
add network=2001:db8::/64 area=backbone_v3
add network=ether1 area=backbone_v3
```

ROSv7 uses templates to match the interface against the template and apply configuration from the matched template. OSPF menus `interface` and `neighbor` contains read-only entries purely for status monitoring.

Example of config comparison

- Use *Moving from ROSv6 to ROSv7 with examples* at help.mikrotik.com under **Pages / RouterOS / Routing** for more detail
- Interoperability with ROSv6 is relatively stable as of 7.5
- **Known issues:** disable BFD on all interfaces in ROSv6 (enabled dynamically by default) to avoid route flapping

S2.2 | OSPF – config overview



```
/routing ospf instance
add disabled=no name=ospf-instance-v2-ipv4 router-id=rid-main \
  routing-table=main
add disabled=no name=ospf-instance-v3-ipv6 router-id=rid-main \
  routing-table=main version=3
/routing ospf area
add disabled=no instance=ospf-instance-v2-ipv4 name=\
  ospf-area-0-ipv4
add disabled=no instance=ospf-instance-v3-ipv6 name=\
  ospf-area-0-ipv6
/routing ospf interface-template
add area=ospf-area-0-ipv4 disabled=no interfaces=vlan3100 \
  type=ptp
add area=ospf-area-0-ipv4 disabled=no interfaces=lo-ipv4
add area=ospf-area-0-ipv6 disabled=no interfaces=vlan3100 \
  type=ptp
add area=ospf-area-0-ipv6 disabled=no interfaces=lo-ipv6 \
  networks=200:1ce:127::1/128 passive
[zuul@rtr-edge-01.v7.ipa.dev] >
```

```
/routing ospf instance
add disabled=no name=ospf-instance-v3-ipv6 router-id=rid-main \
  version=3
add disabled=no name=ospf-instance-v2-ipv4 router-id=rid-main
/routing ospf area
add disabled=no instance=ospf-instance-v3-ipv6 name=\
  ospf-area-0-ipv6
add disabled=no instance=ospf-instance-v2-ipv4 name=\
  ospf-area-0-ipv4
/routing ospf interface-template
add area=ospf-area-0-ipv6 disabled=no interfaces=lo-ipv6 \
  networks=200:1ce:127::3/128 passive
add area=ospf-area-0-ipv4 disabled=no interfaces=vlan3100 \
  type=ptp
add area=ospf-area-0-ipv4 disabled=no interfaces=vlan3200 \
  type=ptp
add area=ospf-area-0-ipv4 disabled=no interfaces=lo-ipv4
add area=ospf-area-0-ipv6 disabled=no interfaces=vlan3100 \
  type=ptp
add area=ospf-area-0-ipv6 disabled=no interfaces=vlan3200 \
  type=ptp
[zuul@rtr-core-01.v7.ipa.dev] >
```

```
/routing ospf instance
add disabled=no name=ospf-instance-v2-ipv4 router-id=rid-main
add disabled=no name=ospf-instance-v3-ipv6 router-id=rid-main \
  version=3
/routing ospf area
add disabled=no instance=ospf-instance-v2-ipv4 name=\
  ospf-area-0-ipv4
add disabled=no instance=ospf-instance-v3-ipv6 name=\
  ospf-area-0-ipv6
/routing ospf interface-template
add area=ospf-area-0-ipv4 disabled=no interfaces=vlan3200 \
  type=ptp
add area=ospf-area-0-ipv4 disabled=no interfaces=lo-ipv4
add area=ospf-area-0-ipv6 disabled=no interfaces=lo-ipv6 \
  networks=200:1ce:127::2/128 passive
add area=ospf-area-0-ipv6 disabled=no interfaces=vlan3200 \
  type=ptp
[zuul@rtr-agg-01.v7.ipa.dev] >
```

S2.2 | OSPF – neighbor/interface validation

```
[zuul@rtr-agg-01.v7.ipa.dev] > routing/ospf/neighbor/print detail
Flags: V - virtual; D - dynamic
0 D instance=ospf-instance-v2-ipv4 area=ospf-area-0-ipv4 address=100.126.1.10 router-id=100.127.1.3 state="Full"
  state-changes=6 adjacency=1d1h13m10s timeout=31s

1 D instance=ospf-instance-v3-ipv6 area=ospf-area-0-ipv6 address=fe80::de2c:6eff:fe7a:d0e6%vlan3200
  router-id=100.127.1.3 state="Full" state-changes=6 adjacency=1d1h13m10s timeout=31s
[zuul@rtr-agg-01.v7.ipa.dev] > routing/ospf/interface/print detail
Flags: D - dynamic
0 D address=100.126.1.9%vlan3200 area=ospf-area-0-ipv4 state=ptp network-type=ptp cost=1 retransmit-interval=5s
  transmit-delay=1s hello-interval=10s dead-interval=40s

1 D address=100.127.1.2%lo-ipv4 area=ospf-area-0-ipv4 state=dr network-type=broadcast cost=1 priority=128
  retransmit-interval=5s transmit-delay=1s hello-interval=10s dead-interval=40s

2 D address=fe80::1458:daff:fe81:387e%lo-ipv6 area=ospf-area-0-ipv6 state=passive network-type=broadcast cost=1
  priority=128 retransmit-interval=5s transmit-delay=1s hello-interval=10s dead-interval=40s

3 D address=fe80::de2c:6eff:fe8a:990f%vlan3200 area=ospf-area-0-ipv6 state=ptp network-type=ptp cost=1
  retransmit-interval=5s transmit-delay=1s hello-interval=10s dead-interval=40s
[zuul@rtr-agg-01.v7.ipa.dev] >
```

S2.2 | OSPF – OSPF active validation

New colored output to help troubleshoot configuration issues

- If there are dependencies OSPF requires that are disabled, the output will show in a print command

```
[zuul@rtr-edge-01.v7.ipa.dev] > routing/ospf/interface-template/print
Flags: X - disabled, I - inactive
0 I ;;; ospf area not active
  area=ospf-area-0-ipv4 interfaces=vlan3100 instance-id=0 type=ptp retransmit-interval=5s transmit-delay=1s hello-interval=10s dead-interval=40s priority=128 cost=1
1 I ;;; ospf area not active
  area=ospf-area-0-ipv4 interfaces=lo-ipv4 instance-id=0 type=broadcast retransmit-interval=5s transmit-delay=1s hello-interval=10s dead-interval=40s priority=128 cost=1
2 I ;;; ospf area not active
  area=ospf-area-0-ipv6 interfaces=vlan3100 instance-id=0 type=ptp retransmit-interval=5s transmit-delay=1s hello-interval=10s dead-interval=40s priority=128 cost=1
3 I ;;; ospf area not active
  area=ospf-area-0-ipv6 interfaces=lo-ipv6 instance-id=0 type=broadcast retransmit-interval=5s transmit-delay=1s hello-interval=10s dead-interval=40s priority=128 cost=1 passive
```


S2.2 | OSPF – Route validation

Three different places to look at OSPF routes

- /ip/route/print
- /ipv6/route/print
- /routing/route/print

```
[zuul@rtr-edge-01.v7.ipa.dev] > ip/route/print where ospf
Flags: D - DYNAMIC; A - ACTIVE; o, y - COPY
Columns: DST-ADDRESS, GATEWAY, DISTANCE
DST-ADDRESS    GATEWAY          DISTANCE
DAo 100.126.1.8/29 100.126.1.2%vlan3100 110
DAo 100.127.1.2/32 100.126.1.2%vlan3100 110
DAo 100.127.1.3/32 100.126.1.2%vlan3100 110
[zuul@rtr-edge-01.v7.ipa.dev] > ipv6/route/print where ospf
Flags: D - DYNAMIC; I, A - ACTIVE; o, y - COPY; H - HW-OFFLOADED
Columns: DST-ADDRESS, GATEWAY, DISTANCE
DST-ADDRESS    GATEWAY          DISTANCE
DIOH 200:1ce:127::1/128 lo-ipv6          110
DAo 200:1ce:127::2/128 fe80::de2c:6eff:fe7a:d0e6%vlan3100 110
DAo 200:1ce:127::3/128 fe80::de2c:6eff:fe7a:d0e6%vlan3100 110
DIOH 200:1ce:a1a:3100::/64 vlan3100          110
DAo 200:1ce:a1a:3200::/64 fe80::de2c:6eff:fe7a:d0e6%vlan3100 110
DAo 200:1ce:a1a:3200::3/128 fe80::de2c:6eff:fe7a:d0e6%vlan3100 110
[zuul@rtr-edge-01.v7.ipa.dev] > routing/route/print where ospf
Flags: U, A - ACTIVE; o, y - COPY; H - HW-OFFLOADED
Columns: DST-ADDRESS, GATEWAY, AFI, DISTANCE, SCOPE, TARGET-SCOPE, IMMEDIATE-GW
DST-ADDRESS    GATEWAY          AFI  DISTANCE  SCOPE  TARGET-SCOPE  IMMEDIATE-GW
Ao 100.126.1.8/29 100.126.1.2%vlan3100 ip4   110      20      10 100.126.1.2%vlan3100
Ao 100.127.1.2/32 100.126.1.2%vlan3100 ip4   110      20      10 100.126.1.2%vlan3100
Ao 100.127.1.3/32 100.126.1.2%vlan3100 ip4   110      20      10 100.126.1.2%vlan3100
UoH 200:1ce:127::1/128 lo-ipv6          ip6   110      20      10
Ao 200:1ce:127::2/128 fe80::de2c:6eff:fe7a:d0e6%vlan3100 ip6   110      20      10 fe80::de2c:6eff:fe7a:d0e6%vlan3100
Ao 200:1ce:127::3/128 fe80::de2c:6eff:fe7a:d0e6%vlan3100 ip6   110      20      10 fe80::de2c:6eff:fe7a:d0e6%vlan3100
UoH 200:1ce:a1a:3100::/64 vlan3100          ip6   110      20      10
Ao 200:1ce:a1a:3200::/64 fe80::de2c:6eff:fe7a:d0e6%vlan3100 ip6   110      20      10 fe80::de2c:6eff:fe7a:d0e6%vlan3100
Ao 200:1ce:a1a:3200::3/128 fe80::de2c:6eff:fe7a:d0e6%vlan3100 ip6   110      20      10 fe80::de2c:6eff:fe7a:d0e6%vlan3100
[zuul@rtr-edge-01.v7.ipa.dev] >
```


S2.2 | OSPF – LSA detail

```
[zuul@rtr-agg-01.v7.ipa.dev] > routing/ospf/lsa/print detail where instance=ospf-instance-v3-ipv6
Flags: S - self-originated, F - flushing, W - wraparound; D - dynamic
3  D instance=ospf-instance-v3-ipv6 area=ospf-area-0-ipv6 type="router" originator=100.127.1.1 id=0.0.0.2
   sequence=0x80000033 age=1344 checksum=0xA461 body=
     options=V6|E|R
     type=p2p iface-id=9 neighbor-iface-if=15 neighbor-router-id=100.127.1.3 metric=1
4  SD instance=ospf-instance-v3-ipv6 area=ospf-area-0-ipv6 type="router" originator=100.127.1.2 id=0.0.0.1
   sequence=0x80000033 age=1363 checksum=0x4CBF body=
     options=V6|E|R
     type=p2p iface-id=5 neighbor-iface-if=13 neighbor-router-id=100.127.1.3 metric=1
5  D instance=ospf-instance-v3-ipv6 area=ospf-area-0-ipv6 type="router" originator=100.127.1.3 id=0.0.0.0
   sequence=0x800000E0 age=1337 checksum=0xA9A5 body=
     options=V6|E|R
     type=p2p iface-id=15 neighbor-iface-if=9 neighbor-router-id=100.127.1.1 metric=1
     type=p2p iface-id=13 neighbor-iface-if=5 neighbor-router-id=100.127.1.2 metric=1
6  D instance=ospf-instance-v3-ipv6 area=ospf-area-0-ipv6 type="intra-area-prefix" originator=100.127.1.1 id=0.0.0.0
   sequence=0x800000E0 age=747 checksum=0xB41A body=
     ref-type=router
     ref-id=0.0.0.0
     ref-router-id=100.127.1.1
     prefix=200:1ce:127::1
     prefix=200:1ce:127::1 options=LA
     prefix=200:1ce:a1a:3100::/64
7  SD instance=ospf-instance-v3-ipv6 area=ospf-area-0-ipv6 type="intra-area-prefix" originator=100.127.1.2 id=0.0.0.0
   sequence=0x800000DE age=118 checksum=0xAE93 body=
     ref-type=router
     ref-id=0.0.0.0
     ref-router-id=100.127.1.2
     prefix=200:1ce:127::2
     prefix=200:1ce:127::2 options=LA
```

S2.3 | BGP - What's new for v7 – part 1

Feature	v7.1	v7.2	v7.3	v7.6	v7.7
BGP					
Convert BGP config from v6 to v7 after upgrade					
BGP Templates and dynamic peers					
BGP connect listen on a network					
BGP guess remote.as					
Show from which peer route received	OK (/routing/route/print detail --> belongs-to)				
BGP Address Families					
BGP input.accept-*					
eBGP nexthop self					
Input Filter					
Output Filter					
BGP Local address auto selection					
BGP route reflect					
BGP route server					
BGP Roles https://datatracker.ietf.org/doc/draft-ietf-idr-bgp-open-policy/?include_text=1	rfc roles not fully implemented				
BGP session uptime in "established" state					
BGP session last established time					

Protocol overview for v7

- Use the feature matrix from help.mikrotik.com under **Pages / RouterOS / Routing** to track progress
- Config conversion typically works without issue now
- BGP Roles enhance peering safety by establishing default filtering behavior
- BGP Route Server capability for use in IXes
- Roadmap includes SNMP monitoring for BGP

S2.3 | BGP - What's new for v7 – part 2

Feature	v7.1	v7.2	v7.3	v7.6	v7.7
BGP Flow Spec	Flow spec attributes are forwarded				
BGP Selection					
BGP Selection (Multipath)	N/A				
BGP Confederation					
BGP Aggregation	N/A				
BGP ORF	N/A				
Discard prefix RTBH RFC 6666	N/A				
AS-wide Unique BGP Identifier RFC 6286	N/A				
Exported PDU PCAP saver					
Exported PDU PCAP loader					
BGP Advertisement monitoring		Initial implementation by dumping to pcap		Initial implementation of /routing/stats/adverts	
BGP Prefix limit			Initial support		

Protocol overview for v7

- Use the feature matrix from help.mikrotik.com under **Pages / RouterOS / Routing** to track progress
- RTBH and Flow Spec basic support has been added
- BGP Advertisement commands are still a work in progress and planned to exist under /routing/stats/adverts

(Command does not exist in ROSv7.6beta6)

S2.3 | BGP – ROSv6 vs. ROSv7

BGP Configuration

There is a complete redesign of the BGP configuration compared to ROSv6. The first biggest difference is that there is no more **instance** and **peer** configuration menus. Instead, we have **connection**, **template** and **session** menus.

The reason for such a structure is to strictly split parameters that are responsible for connection and parameters that are BGP protocol specific.

Let's start with the Template. It contains all BGP protocol-related configuration options. It can be used as a template for dynamic peers and apply a similar config to a group of peers. Note that this is not the same as peer groups on Cisco devices, where the group is more than just a common configuration.

By default, there is a default template that requires you to set your own AS.

```
/routing/bgp/template set default as=65533
```

⚠ Starting from v7.1beta4 template parameters are exposed in the "connection" configuration. This means that the template is not mandatory anymore, allowing for an easier basic BGP connection setup, similar to what it was in ROSv6.

Most of the parameters are similar to ROSv6 except that some are grouped in the output and input section making the config more readable and easier to understand whether the option is applied on input or output. If you are familiar with CapsMan then the syntax is the same, for example, to specify the output selection chain you set `output.filter-chain=myBgpChain`.

You can even inherit template parameters from another template, for example:

```
/routing/bgp/template
add name=myAsTemplate as=65500 output.filter-chain=myAsFilter
set default template=myAsTemplate
```

Example of config comparison

- Use *Moving from ROSv6 to ROSv7 with examples* at help.mikrotik.com under **Pages / RouterOS / Routing** for more detail
- Move to templating makes peering scale easier
- **Known issues:** iBGP over IPv6 works in ROSv7 but has limitations in ROSv6 and requires nexthop self as a workaround – interop iBGP over IPv6 between ROSv6 & v7 carefully to avoid loops

S2.3 | BGP - What's new for v7?

BGP roles (RFC 9234) prevent inadvertent route leaks with default filtering

- BGP roles are mandatory in ROSv7
- Doesn't replace the need for filtering but creates default filtering behavior to avoid major route leak issues

local - a group of parameters associated with the local side of the connection		
	.address (IPv4/6; Default: ::)	Local connection address.
	.port (integer [0..65535]; Default: 179)	Local connection port.
	.role (ebgp ebgp-customer ebgp-peer ebgp-provider ebgp-rs ebgp-rs-client ibgp ibgp-rr ibgp-rr-client; Default:)	BGP role, in most common scenarios it should be set to iBGP or eBGP. More information on BGP roles can be found in the corresponding RFC draft https://datatracker.ietf.org/doc/draft-ietf-idr-bgp-open-policy/?include_text=1

```
[zuu]@rtr-agg-01.v7.ipa.dev] > routing/bgp/connection/print
Flags: D - dynamic, X - disabled, I - inactive
0  name="peer-ipv4-rtr-core-01"
   remote.address=100.127.1.3 .as=1016
   local.address=100.127.1.2 .role=ibgp-rr-client
   routing-table=main router-id=100.127.1.2 templates=as1016-v4 as=1016 address-families=ip
   output.filter-chain=as1016-ipv4-services-no-default .network=add-list-as1016-ipv4-services

1  name="peer-ipv6-rtr-core-01"
   remote.address=200:1ce:127::3 .as=1016
   local.address=200:1ce:127::2 .role=ibgp-rr-client
   routing-table=main router-id=100.127.1.2 templates=as1016-v6 as=1016 address-families=ipv6
```

S2.3 | BGP - What's new for v7?

BGP peering affinity allows for configuration of session balancing across CPU cores

```
[zuul@rtr-edge-01.v7.ipa.dev] > routing/bgp/connection/print
Flags: D - dynamic, X - disabled, I - inactive
0  name="peer-ipv4-rtr-core-01"
   remote.address=100.127.1.3 .as=1016
   local.address=100.127.1.1 .role=ibgp-rr-client
   routing-table=main router-id=100.127.1.1 templates=as1016-v4 as=1016 nexthop-choice=force-self
   address-families=ip
   output.filter-chain=as1016-ipv4-internal-permit-default .network=add-list-as1016-ipv4-internal
   .default-originate=always
   input.affinity=afi

1  name="peer-ipv6-rtr-core-01"
   remote.address=200:1ce:127::3 .as=1016
   local.address=200:1ce:127::1 .role=ibgp-rr-client
   routing-table=main router-id=100.127.1.1 templates=as1016-v6 as=1016 nexthop-choice=force-self
   address-families=ipv6
   output.filter-chain=as1016-ipv6-internal-permit-default .network=add-list-as1016-ipv6-internal
   .default-originate=always
   input.affinity=afi

[zuul@rtr-edge-01.v7.ipa.dev] > routing/stats/process/print
Columns: TASKS, PRIVATE-MEM-BLOCKS, SHARED-MEM-BLOCKS, PSS, RSS, VMS, ID, PID, RPID, PROCESS-TIME, KERNEL-TIME, CUR-BUSY, MAX-BUSY
# TASKS          PRIVATE-M  SHARED-ME  P  R  V  ID      PID  R  PROCESS-  KERNEL-T  CUR  MAX-B
0 routing tables  768.0KiB   1792.0KiB  0  0  0  main    102  0  10s820ms  10s880ms  0ms  20ms
  rib
1 fib            0          0          0  0  0  fib      119  1  1s180ms  1s710ms   310ms
2 ospf           768.0KiB   256.0KiB   0  0  0  ospf     123  1  18s430ms  10s90ms   10ms
3 pimsn          0          0          0  0  0  pim      124  1  1s400ms  1s140ms   30ms
4 fantasy        0          0          0  0  0  fantasy  126  1  1s60ms   1s200ms   20ms
5 configuration and reporting  0          512.0KiB   0  0  0  static   127  1  1s460ms  1s120ms   10ms
6 ldp            0          512.0KiB   0  0  0  mpls     125  1  4s490ms  2s640ms   40ms
  Copy
7 rip            0          0          0  0  0  rip      122  1  1s280ms  1s200ms   10ms
8 routing policy configuration  512.0KiB   512.0KiB   0  0  0  policy   120  1  1s350ms  1s300ms   10ms
9 BGP service     256.0KiB   0          0  0  0  bgp      121  1  2s360ms  3s650ms   10ms
10 BGP Input 200:1ce:127::3  512.0KiB   512.0KiB   0  0  0  26       4551  1
   BGP Output 200:1ce:127::3
11 BGP Input 100.127.1.3    512.0KiB   512.0KiB   0  0  0  27       4552  1
   BGP Output 100.127.1.3
12 Global memory          256.0KiB          global  0  0
```

S2.3 | BGP - What's new for v7?

BGP large communities

```
[zuul@rtr-core-01.v7.ipa.dev] > routing/route/print detail where bgp.large-communities
Flags: X - disabled, F - filtered, U - unreachable, A - active;
c - connect, s - static, r - rip, b - bgp, o - ospf, d - dhcp, v - vpn, m - modem, a - ldp-address, l - ldp-mapping, y -
copy;
H - hw-offloaded; + - ecmp, B - blackhole
AbH  afi=ip4 contribution=active dst-address=198.18.0.1/32 routing-table=main gateway=100.127.1.1 immediate-gw=100.126.
1.1%vlan3100
      distance=200 scope=40 target-scope=30 belongs-to="bgp-IP-100.127.1.1"
      bgp.peer-cache-id=*B000002 .large-communities=1016:1016:4 .local-pref=100 .atomic-aggregate=no .origin=igp
      debug.fwp-ptr=0x20282960

bH   afi=ip6 contribution=best-candidate dst-address=200:1ce:127::1/128 routing-table=main gateway=200:1ce:127::1
immediate-gw=fe80::de2c:6eff:fe8a:982d%vlan3100 distance=200 scope=40 target-scope=30 belongs-to="bgp-IP6-200:1ce
:127::1"
      bgp.peer-cache-id=*B000001 .large-communities=1016:1016:6 .local-pref=100 .atomic-aggregate=no .origin=igp
      debug.fwp-ptr=0x202828A0
[zuul@rtr-core-01.v7.ipa.dev] > routing/filter/community-large-list/print
```

```
/routing filter community-large-list
add communities=1016:1016:4 list=bgp-community-large-1016-1016-4
add communities=1016:1016:6 list=bgp-community-large-1016-1016-6
/routing filter rule
add chain=as1016-ipv4-internal-permit-default rule="if (dst in 198.18.0.0/22 && dst-len > 22) { set bgp-large-communit\
ies bgp-community-large-1016-1016-4; accept; }"
add chain=as1016-ipv4-internal-permit-default rule="if (dst== 0.0.0.0/0) { accept; }"
add chain=as1016-ipv4-internal-permit-default rule="if (dst in 0.0.0.0/0) { reject;}"
add chain=as1016-ipv6-internal-permit-default rule=\
"if (dst in 200:1ce::/32 && dst-len > 32) { set bgp-large-communities bgp-community-large-1016-1016-6; accept; }"
add chain=as1016-ipv6-internal-permit-default rule="if (dst==::/0) { accept; }"
add chain=as1016-ipv6-internal-permit-default rule="if (dst in ::/0) { reject;}"
[zuul@rtr-core-01.v7.ipa.dev] >
```


S2.3 | BGP - What's new for v7?

iBGP recursive routing over IPv6 is now functional

- Linux kernel limitations prevented this in ROSv6
- Track recursive routing using **gateway** and **immediate-gw**

```
[zuul@rtr-agg-01.v7.ipa.dev] > routing/route/print detail where afi=ip6 && bgp
Flags: X - disabled, F - filtered, U - unreachable, A - active;
c - connect, s - static, r - rip, b - bgp, o - ospf, d - dhcp, v - vpn, m - modem, a - ldp-address, l -
copy;
H - hw-offloaded; + - ecmp, B - blackhole
Ab  afi=ip6 contribution=active dst-address=::/0 routing-table=main gateway=200:1ce:127::1
    immediate-gw=fe80::de2c:6eff:fe7a:d0e6%vlan3200 distance=200 scope=40 target-scope=30
    belongs-to="bgp-IP6-200:1ce:127::3"
    bgp.peer-cache-id=*B000002 .local-pref=100 .atomic-aggregate=yes .origin=igp
    debug.fwp-ptr=0x202C2660
[zuul@rtr-agg-01.v7.ipa.dev] > tool/traceroute 3ffe:1001:1002::2
Columns: ADDRESS, LOSS, SENT, LAST, AVG, BEST, WORST, STD-DEV
#  ADDRESS      LOSS  SENT  LAST  AVG  BEST  WORST  STD-DEV
1  200:1ce:a1a:3200::3  0%    8    0.1ms  0    0    0.1    0
2  3ffe:1001:1002::2   0%    8     0ms   0    0     0     0
```

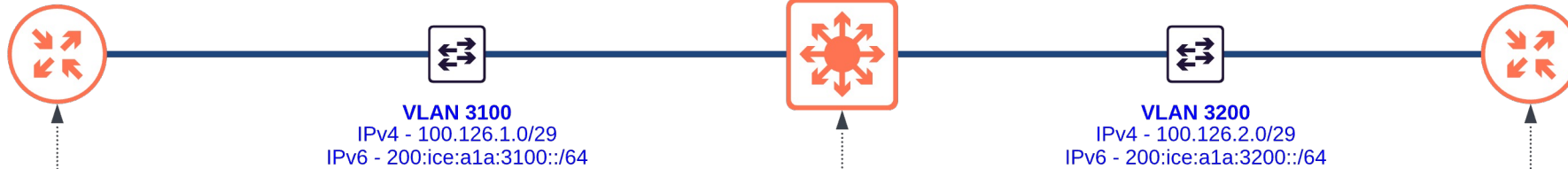

S2.4 | BGP – config overview

AS 1016 198.18.0.0/22
200:1ce::/32

rtr-edge-01.v7.ipa.dev
(CCR2004-16G-2S+)

rtr-core-01.v7.ipa.dev
(CCR2116-12G-4S+)

rtr-agg-01.v7.ipa.dev
(CCR2004-16G-2S+)



```
/routing bgp template
add address-families=ip as=1016 name=as1016-v4 router-id=100.127.1.1
add address-families=ipv6 as=1016 name=as1016-v6 router-id=100.127.1.1
/routing id
add id=100.127.1.1 name=rid-main select-dynamic-id-only-vrf select-from-vrf=main
/routing ospf instance
add disabled=no name=ospf-instance-v2-ipv4 router-id=rid-main routing-table=main
add disabled=no name=ospf-instance-v3-ipv6 router-id=rid-main routing-table=main \
version=3
/routing ospf area
add disabled=no instance=ospf-instance-v2-ipv4 name=ospf-area-0-ipv4
add disabled=no instance=ospf-instance-v3-ipv6 name=ospf-area-0-ipv6
/routing bgp connection
add disabled=no local.address=100.127.1.1 .role=ibgp-rr-client name=\
peer-ipv4-rtr-core-01 nexthop-choice=force-self output.default-originate=always \
.filter-chain=as1016-ipv4-internal-permit-default .network=\
add-list-as1016-ipv4-internal remote.address=100.127.1.3 .as=1016 templates=\
as1016-v4
add local.address=200:1ce:127::1 .role=ibgp-rr-client name=peer-ipv6-rtr-core-01 \
nexthop-choice=force-self output.default-originate=always .filter-chain=\
as1016-ipv6-internal-permit-default .network=add-list-as1016-ipv6-internal \
remote.address=200:1ce:127::3 .as=1016 templates=as1016-v6
/routing filter community-large-list
add communities=1016:1016:4 list=bgp-community-large-1016-1016-4
add communities=1016:1016:6 list=bgp-community-large-1016-1016-6
/routing filter rule
add chain=as1016-ipv4-internal-permit-default rule="if (dst in 198.18.0.0/22 && dst-len > 22) { set bgp-large-communities bgp-community-large-1016-1016-4; accept; }"
add chain=as1016-ipv4-internal-permit-default rule="if (dst== 0.0.0.0/0) { accept; }"
add chain=as1016-ipv4-internal-permit-default rule=\
"if (dst in 0.0.0.0/0) { reject; }"
add chain=as1016-ipv6-internal-permit-default rule="if (dst in 200:1ce::/32 && dst-len > 32) { set bgp-large-communities bgp-community-large-1016-1016-6; accept; }"
add chain=as1016-ipv6-internal-permit-default rule="if (dst==::/0) { accept; }"
add chain=as1016-ipv6-internal-permit-default rule="if (dst in ::/0) { reject; }"
```

```
/routing bgp template
add address-families=ip as=1016 cluster-id=100.127.1.3 name=as1016-v4-rr router-id=\
100.127.1.3
add address-families=ipv6 as=1016 cluster-id=100.127.1.3 name=as1016-v6-rr router-id=\
100.127.1.3
/routing id
add id=100.127.1.3 name=rid-main select-dynamic-id-only-vrf select-from-vrf=main
/routing ospf instance
add disabled=no name=ospf-instance-v3-ipv6 router-id=rid-main version=3
add disabled=no name=ospf-instance-v2-ipv4 router-id=rid-main
/routing ospf area
add disabled=no instance=ospf-instance-v3-ipv6 name=ospf-area-0-ipv6
add disabled=no instance=ospf-instance-v2-ipv4 name=ospf-area-0-ipv4
/routing bgp connection
add cluster-id=100.127.1.3 local.address=100.127.1.3 .role=ibgp-rr name=\
peer-ipv4-rtr-edge-01 remote.address=100.127.1.1 .as=1016 templates=as1016-v4-rr
add cluster-id=100.127.1.3 input.filter=as1016-ipv4-services-no-default \
local.address=100.127.1.3 .role=ibgp-rr name=peer-ipv4-rtr-agg-01 \
output.default-originate=if-installed remote.address=100.127.1.2 .as=1016 \
templates=as1016-v4-rr
add local.address=200:1ce:127::3 .role=ibgp-rr name=peer-ipv6-rtr-edge-01 \
remote.address=200:1ce:127::1 .as=1016 templates=as1016-v6-rr
add local.address=200:1ce:127::3 .role=ibgp-rr name=peer-ipv6-rtr-agg-01 \
output.default-originate=if-installed remote.address=200:1ce:127::2 .as=1016 \
templates=as1016-v6-rr
/routing filter rule
add chain=as1016-ipv4-services-no-default rule=\
"if (dst in 198.18.1.0/24 && dst-len >= 24) { accept; }"
add chain=as1016-ipv4-services-no-default rule=\
"if (dst in 198.18.2.0/24 && dst-len >= 24) { accept; }"
add chain=as1016-ipv4-services-no-default rule=\
"if (dst in 198.18.3.0/24 && dst-len >= 24) { accept; }"
add chain=as1016-ipv4-services-no-default rule="if (dst in 0.0.0.0/0) { reject; }"
```

```
/routing bgp template
add address-families=ip as=1016 name=as1016-v4 router-id=100.127.1.2
add address-families=ipv6 as=1016 name=as1016-v6 router-id=100.127.1.2
/routing id
add id=100.127.1.2 name=rid-main
add id=100.127.1.2 name=rid-main select-dynamic-id-only-vrf select-from-vrf=main
/routing ospf instance
add disabled=no name=ospf-instance-v2-ipv4 router-id=rid-main
add disabled=no name=ospf-instance-v3-ipv6 router-id=rid-main version=3
/routing ospf area
add disabled=no instance=ospf-instance-v2-ipv4 name=ospf-area-0-ipv4
add disabled=no instance=ospf-instance-v3-ipv6 name=ospf-area-0-ipv6
/routing bgp connection
add local.address=100.127.1.2 .role=ibgp-rr-client name=peer-ipv4-rtr-core-01 \
output.filter-chain=as1016-ipv4-services-no-default .network=\
add-list-as1016-ipv4-services remote.address=100.127.1.3 .as=1016 templates=\
as1016-v4
add local.address=200:1ce:127::2 .role=ibgp-rr-client name=peer-ipv6-rtr-core-01 \
remote.address=200:1ce:127::3 .as=1016 templates=as1016-v6
/routing filter community-large-list
add communities=1016:1016:4 list=bgp-community-large-1016-1016-4
/routing filter rule
add chain=as1016-ipv4-services-no-default rule=\
"if (dst in 198.18.1.0/24 && dst-len >= 24) { accept; }"
add chain=as1016-ipv4-services-no-default rule=\
"if (dst in 198.18.2.0/24 && dst-len >= 24) { accept; }"
add chain=as1016-ipv4-services-no-default rule=\
"if (dst in 198.18.3.0/24 && dst-len >= 24) { accept; }"
add chain=as1016-ipv4-services-no-default rule="if (dst in 0.0.0.0/0) { reject; }"
```

S2.4 | BGP – route validation

Three different places to look at BGP routes

- `/ip/route/print`
- `/ipv6/route/print`
- `/routing/route/print`

```
[zuul@rtr-edge-01.v7.ipa.dev] > ip/route/print where bgp
Flags: D - DYNAMIC; A - ACTIVE; b, y - COPY
Columns: DST-ADDRESS, GATEWAY, DISTANCE
  DST-ADDRESS  GATEWAY  DISTANCE
D b 198.18.0.1/32 100.127.1.3      200
DA b 198.18.1.0/24 100.127.1.2      200
DA b 198.18.2.0/24 100.127.1.2      200
[zuul@rtr-edge-01.v7.ipa.dev] > ipv6/route/print where bgp
Flags: D - DYNAMIC; A - ACTIVE; b, y - COPY
Columns: DST-ADDRESS, GATEWAY, DISTANCE
  DST-ADDRESS  GATEWAY  DISTANCE
DA b 200:1ce:1001::/48 200:1ce:127::2    200
DA b 200:1ce:1002::/48 200:1ce:127::2    200
DA b 200:1ce:1003::/48 200:1ce:127::2    200
[zuul@rtr-edge-01.v7.ipa.dev] > routing/route/print where bgp
Flags: A - ACTIVE; b, y - COPY
Columns: DST-ADDRESS, GATEWAY, AFI, DISTANCE, SCOPE, TARGET-SCOPE, IMMEDIATE-GW
  DST-ADDRESS  GATEWAY  AFI  DISTANCE  SCOPE  TARGET-SCOPE  IMMEDIATE-GW
  b 198.18.0.1/32 100.127.1.3  ip4    200     40          30 100.126.1.2%vlan3100
Ab 198.18.1.0/24 100.127.1.2  ip4    200     40          30 100.126.1.2%vlan3100
Ab 198.18.2.0/24 100.127.1.2  ip4    200     40          30 100.126.1.2%vlan3100
Ab 200:1ce:1001::/48 200:1ce:127::2 ip6    200     40          30 fe80::de2c:6eff:fe7a:d0e6%vlan3100
Ab 200:1ce:1002::/48 200:1ce:127::2 ip6    200     40          30 fe80::de2c:6eff:fe7a:d0e6%vlan3100
Ab 200:1ce:1003::/48 200:1ce:127::2 ip6    200     40          30 fe80::de2c:6eff:fe7a:d0e6%vlan3100
```

S2.4 | BGP – peer validation

```
[zuul@rtr-edge-01.v7.ipa.dev] > routing/bgp/session/print detail
Flags: E - established
0 E name="peer-ipv6-rtr-core-01-1"
  remote.address=200:1ce:127::3 .as=1016 .messages=1595 .bytes=30373 .eor=""
  local.role=ibgp-rr-client .address=200:1ce:127::1 .as=1016 .id=100.127.1.1 .capabilities=mp,rr,gr,as4 .afi=ipv6
  .messages=1596 .bytes=30446 .eor=""
  output.procid=24 .filter-chain=as1016-ipv6-internal-permit-default .network=add-list-as1016-ipv6-internal
  .default-originate=always .last-notification=ffffffffffffffffffffffffffffffff0015030400
  input.procid=24 ibgp
  nexthop-choice=force-self multihop=yes hold-time=3m

1 E name="peer-ipv4-rtr-core-01-1"
  remote.address=100.127.1.3 .as=1016 .messages=1597 .bytes=30442 .eor=""
  local.role=ibgp-rr-client .address=100.127.1.1 .as=1016 .id=100.127.1.1 .capabilities=mp,rr,gr,as4
  .messages=1596 .bytes=30398 .eor=""
  output.procid=25 .filter-chain=as1016-ipv4-internal-permit-default .network=add-list-as1016-ipv4-internal
  .default-originate=always .last-notification=ffffffffffffffffffffffffffffffff0015030400
  input.procid=25 ibgp
  nexthop-choice=force-self multihop=yes hold-time=3m
[zuul@rtr-edge-01.v7.ipa.dev] > █
```

S2.4 | BGP – received routes validation

```
[zuul@rtr-edge-01.v7.ipa.dev] > routing/route/print where belongs-to="bgp-IP6-200:1ce:127::3"
Flags: A - ACTIVE; b, y - COPY
Columns: DST-ADDRESS, GATEWAY, AFI, DISTANCE, SCOPE, TARGET-SCOPE, IMMEDIATE-GW
  DST-ADDRESS      GATEWAY      AFI  DISTANCE  SCOPE  TARGET-SCOPE  IMMEDIATE-GW
Ab 200:1ce:1001::/48 200:1ce:127::2 ip6      200      40      30  fe80::de2c:6eff:fe7a:d0e6%vlan3100
Ab 200:1ce:1002::/48 200:1ce:127::2 ip6      200      40      30  fe80::de2c:6eff:fe7a:d0e6%vlan3100
Ab 200:1ce:1003::/48 200:1ce:127::2 ip6      200      40      30  fe80::de2c:6eff:fe7a:d0e6%vlan3100
[zuul@rtr-edge-01.v7.ipa.dev] > routing/route/print detail where belongs-to="bgp-IP6-200:1ce:127::3"
Flags: X - disabled, F - filtered, U - unreachable, A - active;
c - connect, s - static, r - rip, b - bgp, o - ospf, d - dhcp, v - vpn, m - modem, a - ldp-address, l - ldp-mapp
copy;
H - hw-offloaded; + - ecmp, B - blackhole
Ab  afi=ip6 contribution=active dst-address=200:1ce:1001::/48 routing-table=main gateway=200:1ce:127::2
    immediate-gw=fe80::de2c:6eff:fe7a:d0e6%vlan3100 distance=200 scope=40 target-scope=30
    belongs-to="bgp-IP6-200:1ce:127::3"
    bgp.peer-cache-id=*B000001 .local-pref=100 .atomic-aggregate=yes .origin=igp
    debug.fwp-ptr=0x20342D20

Ab  afi=ip6 contribution=active dst-address=200:1ce:1002::/48 routing-table=main gateway=200:1ce:127::2
    immediate-gw=fe80::de2c:6eff:fe7a:d0e6%vlan3100 distance=200 scope=40 target-scope=30
    belongs-to="bgp-IP6-200:1ce:127::3"
    bgp.peer-cache-id=*B000001 .local-pref=100 .atomic-aggregate=yes .origin=igp
    debug.fwp-ptr=0x20342D20

Ab  afi=ip6 contribution=active dst-address=200:1ce:1003::/48 routing-table=main gateway=200:1ce:127::2
    immediate-gw=fe80::de2c:6eff:fe7a:d0e6%vlan3100 distance=200 scope=40 target-scope=30
    belongs-to="bgp-IP6-200:1ce:127::3"
    bgp.peer-cache-id=*B000001 .local-pref=100 .atomic-aggregate=yes .origin=igp
    debug.fwp-ptr=0x20342D20
```


S2.4 | BGP – advertised routes validation

```
[zuul@sw-core-01.jan1.us.ipa] > routing/stats/pcap/print detail where file=edge-02-bgp-v4.pcap
file="edge-02-bgp-v4.pcap" timestamp=2724w6d15h19m53s170ms771us600ns src=100.126.32.65 dst=100.126.32.66 protocol="bgp"
data=ffffffffffffffffffffffffffffffff00670200000028400101005002000602010000fdf4400304647e2041400600d020000c0000fdf4000027090000000418647e0018647f0018647e0a18647f0a180aff0a18647f6018
647d00130aff40130aff60130aff80
bgp.type=2
bgp.update.nextthop=100.126.32.65
.nlri=100.126.0.0/24,100.127.0.0/24,100.126.10.0/24,100.127.10.0/24,10.255.10.0/24,100.127.96.0/24,100.125.0.0/24,10.255.64.0/19,10.255.96.0/19,10.255.128.0/19 .origin=0
.as-path=sequence 65012 .large-communities=65012:9993:4 .atomic-aggregate=yes

file="edge-02-bgp-v4.pcap" timestamp=2724w6d15h19m53s170ms823us440ns src=100.126.32.65 dst=100.126.32.66 protocol="bgp"
data=ffffffffffffffffffffffffffffffff00400200000028400101005002000602010000fdf4400304647e2041400600d020000c0000fdf4000000000000000400
bgp.type=2
bgp.update.nextthop=100.126.32.65 .nlri=0.0.0.0/0 .origin=0 .as-path=sequence 65012 .large-communities=65012:0:4 .atomic-aggregate=yes

file="edge-02-bgp-v4.pcap" timestamp=2724w6d15h19m53s170ms881us720ns src=100.126.32.65 dst=100.126.32.66 protocol="bgp"
data=ffffffffffffffffffffffffffffffff00620200000015400101005002000602010000fdf4400304647e2041180aff22180aff23180aff20180aff21180aff2b18c0a82020647f20031d647e20501d647e20581d647e2000
1d647e20401d647e2048
bgp.type=2
bgp.update.nextthop=100.126.32.65 .nlri=10.255.34.0/24,10.255.35.0/24,10.255.32.0/24,10.255.33.0/24,10.255.43.0/24,192.168.32.0/24,100.127.32.3,100.126.32.80/29,100.126.32.88/29,
100.126.32.0/29,100.126.32.64/29,100.126.32.72/29
.origin=0 .as-path=sequence 65012
[zuul@sw-core-01.jan1.us.ipa] >
```



MikroTik RouterOS v7:
Layer 3 Deep Dive

L3 hw offload - using Marvell Prestera ASICs
for wirespeed forwarding

Section 3 | L3 hw/fw/nat offload

03

- | Requirements for hw offload
- | l3 hw-offload for IPv4 and IPv6
- | Requirements for nat/fw offload
- | l3 fw/nat offload for IPv4

S3.1 | l3 hw offload – requirements

- **Design considerations and requirements**

- All L3 configuration must be on the bridge with VLAN interface tied to the bridge
- LACP Bonding is permitted but subinterfaces of bonds or standalone interfaces will break l3 hw offload
- Only one bridge with member ports will be hw offloaded
- MPLS is not yet supported as a P router (it was previously in 6.x code beta)
- Some devices reboot when l3 hw offload is enabled – from our testing, it's generally a good idea to reboot the router/switch when making ASIC changes

```
[zuul@rtr-core-01.v7.ipa.dev] > interface/ethernet/switch/print detail
Flags: I - invalid
0  name="switch1" type=Marvell-98DX3255 mirror-source=none mirror-target=none mirror-egress-target=none
   l3-hw-offloading=yes
[zuul@rtr-core-01.v7.ipa.dev] > interface/ethernet/switch/l3hw-settings/print
      fasttrack-hw: yes
          ipv6-hw: yes
  icmp-reply-on-error: yes
hw-supports-fasttrack: yes
```

S3.1 | L3 hw offload – requirements

- L3 hw offload for IPv4 and IPv6
 - 'H' flag denotes forwarding in hardware for both AFIs

```
[zuul@rtr-core-01.v7.ipa.dev] > routing/route/print where hw-offloaded && afi=ip4
Flags: A - ACTIVE; c, b, o, y - COPY; H - HW-OFFLOADED
Columns: DST-ADDRESS, GATEWAY, AFI, DISTANCE, SCOPE, TARGET-SCOPE, IMMEDIATE-GW
DST-ADDRESS    GATEWAY          AFI  DISTANCE  SCOPE  TARGET-SCOPE  IMMEDIATE-GW
AbH 0.0.0.0/0    100.127.1.1      ip4    200      40      30  100.126.1.1%vlan3100
AcH 100.126.1.0/29  vlan3100        ip4     0      10           vlan3100
AcH 100.126.1.8/29  vlan3200        ip4     0      10           vlan3200
AoH 100.127.1.1/32  100.126.1.1%vlan3100 ip4    110     20      10  100.126.1.1%vlan3100
AoH 100.127.1.2/32  100.126.1.9%vlan3200 ip4    110     20      10  100.126.1.9%vlan3200
AbH 198.18.0.1/32  100.127.1.1      ip4    200     40      30  100.126.1.1%vlan3100
AbH 198.18.1.0/24  100.127.1.2      ip4    200     40      30  100.126.1.9%vlan3200
AbH 198.18.2.0/24  100.127.1.2      ip4    200     40      30  100.126.1.9%vlan3200
[zuul@rtr-core-01.v7.ipa.dev] > routing/route/print where hw-offloaded && afi=ip6
Flags: U, A - ACTIVE; c, b, o, y - COPY; H - HW-OFFLOADED
Columns: DST-ADDRESS, GATEWAY, AFI, DISTANCE, SCOPE, TARGET-SCOPE
DST-ADDRESS    GATEWAY          AFI  DISTANCE  SCOPE  TARGET-SCOPE
AbH ::/0        200:1ce:127::1   ip6    200     40      30
  bH 200:1ce:127::1/128  200:1ce:127::1   ip6    200     40      30
AoH 200:1ce:127::1/128  fe80::de2c:6eff:fe8a:982d%vlan3100 ip6    110     20      10
AoH 200:1ce:127::2/128  fe80::de2c:6eff:fe8a:990f%vlan3200 ip6    110     20      10
UoH 200:1ce:127::3/128  lo-ipv6          ip6    110     20      10
UoH 200:1ce:a1a:3100::/64  vlan3100        ip6    110     20      10
AcH 200:1ce:a1a:3100::/64  vlan3100        ip6     0      10
UoH 200:1ce:a1a:3200::/64  vlan3200        ip6    110     20      10
AcH 200:1ce:a1a:3200::/64  vlan3200        ip6     0      10
AbH 200:1ce:1001::/48    200:1ce:127::2   ip6    200     40      30
AbH 200:1ce:1002::/48    200:1ce:127::2   ip6    200     40      30
AbH 200:1ce:1003::/48    200:1ce:127::2   ip6    200     40      30
AcH fe80::%sfp-sfpplus3/64 sfp-sfpplus3     ip6     0      10
AcH fe80::%sfp-sfpplus4/64 sfp-sfpplus4     ip6     0      10
```

S3.1 | l3 fw offload – requirements

- Design considerations and requirements

- No VLANs, bonding or subinterfaces
- Must use two ports – one per path
- Hw offload must be enabled – all port level hw offload config must be disabled.
- Fastpath config required

436	SACFhs	tcp	10.250.77.101:55948	203.0.113.101:5201	established	23h59m55s	100.5Mbps	350.8kbp>	R	✖	sfp28-2-To-KM4.4	Ethernet	1500	1584	0 bps
437	SACFhs	tcp	10.250.77.101:55900	203.0.113.101:5201	established	23h59m55s	67.8Mbps	232.6kbp>	R	✖	sfp28-3	Ethernet	1500	1584	0 bps
439	SACFhs	tcp	10.250.77.101:55980	203.0.113.101:5201	established	23h59m55s	84.3Mbps	286.4kbp>	R	✖	sfp28-4	Ethernet	1500	1584	0 bps
440	SACFhs	tcp	10.250.77.101:55796	203.0.113.101:5201	established	23h59m55s	78.1Mbps	319.7kbp>	R	✖	sfp28-5	Ethernet	1500	1584	0 bps
441	SACFhs	tcp	10.250.77.101:55896	203.0.113.101:5201	established	23h59m55s	47.3Mbps	220.1kbp>	R	✖	sfp28-6	Ethernet	1500	1584	0 bps
442	SACFhs	tcp	10.250.77.101:55906	203.0.113.101:5201	established	23h59m55s	96.5Mbps	301.0kbp>	R	✖	sfp28-7	Ethernet	1500	1584	0 bps
443	SACFhs	tcp	10.250.77.101:55826	203.0.113.101:5201	established	23h59m55s	59.2Mbps	294.4kbp>	R	✖	sfp28-8	Ethernet	1500	1584	0 bps
444	SACFhs	tcp	10.250.77.101:55966	203.0.113.101:5201	established	23h59m55s	82.5Mbps	276.1kbp>	R	✖	sfp28-9	Ethernet	1500	1584	42.4 Mbps
445	SACFhs	tcp	10.250.77.101:55792	203.0.113.101:5201	established	23h59m55s	58.1Mbps	298.4kbp>	R	✖	sfp28-10	Ethernet	1500	1584	9.7 Gbps
446	SACFhs	tcp	10.250.77.101:55932	203.0.113.101:5201	established	23h59m55s	67.8Mbps	259.1kbp>	R	✖	sfp28-11	Ethernet	1500	1584	0 bps
447	SACFhs	tcp	10.250.77.101:55972	203.0.113.101:5201	established	23h59m55s	53.3Mbps	302.0kbp>	R	✖	sfp28-12-To-Edge-Point90M	Ethernet	1500	1584	0 bps
448	SACFhs	tcp	10.250.77.101:55924	203.0.113.101:5201	established	23h59m55s	101.0Mbps	304.0kbp>		✖	vlan777	VLAN			0 bps
449	SACFhs	tcp	10.250.77.101:55920	203.0.113.101:5201	established	23h59m55s	76.0Mbps	280.9kbp>		✖	vlan778	VLAN			0 bps
450	SACFhs	tcp	10.250.77.101:55954	203.0.113.101:5201	established	23h59m55s	91.6Mbps	253.8kbp>							
451	SACFhs	tcp	10.250.77.101:55840	203.0.113.101:5201	established	23h59m55s	199.0Mbps	441.5kbp>							
452	SACFhs	tcp	10.250.77.101:55898	203.0.113.101:5201	established	23h59m55s	85.4Mbps	262.0kbp>							
453	SACFhs	tcp	10.250.77.101:55780	203.0.113.101:5201	established	23h59m55s	89.9Mbps	326.2kbp>							

Thank you!

Thank you for joining us today! This is a large topic with plenty of nuances, if you'd like to brainstorm with us your deployment, network architecture, or software ecosystem, do not hesitate to contact us using the information below.

We are a full-service networking firm that can help identify areas of improvement, design network architecture, as well as plan and execute your migration windows.

IP ArchiTechs

consulting@iparchitech.com

11757 W Ken Caryl Ave, Littleton, CO, 80127

+1 (855) 645-7684