1off.it Technical Presentation

MikroTik ISP Datacenter Architecture

06 mar 2025

MikroTik professionals conference

About me
Alessandro Campanella

I have been working in the telecommunications field for over 20 years as a producer, consultant, and trainer, both in Italy and abroad.

I was among the pioneers of wireless technology, conducting experiments on my network and gathering the experiences of the many operators and integrators with whom I have collaborated over the years.

MikroTik Evangelist
Design, analysis, and consulting for ISPs
MikroTik Training and Certification

a.campanella@1off.it
@alessandrocampanella

www.1off.it

MUM 2006

I was there (here)

Returning to the place of the first MUM after 18 years, reconnecting with friends and colleagues, evokes a mix of nostalgia and joy.

It's a powerful reminder of the journey undertaken and the enduring bonds formed.

This reunion is both a reflection on past experiences and a celebration of growth and enduring connections.

I was there Gandalf. I was there 3000 years ago…

Riga 2007 - 1° Train the Trainer

Arnis Riekstins　　Patrik　　me　　Tom　　Maia　　Janis Megis

Module 1
Introduction

My Journey in ISP Network Design

My Story: From Simplicity to a Robust Architecture

When I started working with ISPs, MikroTik wasn’t the first name that came to mind for building a high-performance data center.

The common belief was that only high-end vendors could provide reliable solutions for **scalability**, **redundancy**, and **performance**.

But over the years, I refined my approach, continuously testing and adapting my designs.

I saw MikroTik evolve rapidly, bringing powerful features that — when used correctly — transformed the way I built ISP infrastructures.

My Journey in ISP Network Design

Evolving ISP Networks with MikroTik

- From basic routing to full **Layer 3 hardware offloading**.
- From traditional STP-based failover to fast and **resilient MLAG setups**.
- From simple VLAN segmentation to scalable, **ISP-grade network architectures**.

MikroTik was no longer just an affordable alternative, it became a key enabler for ISP data centers.

My Journey in ISP Network Design

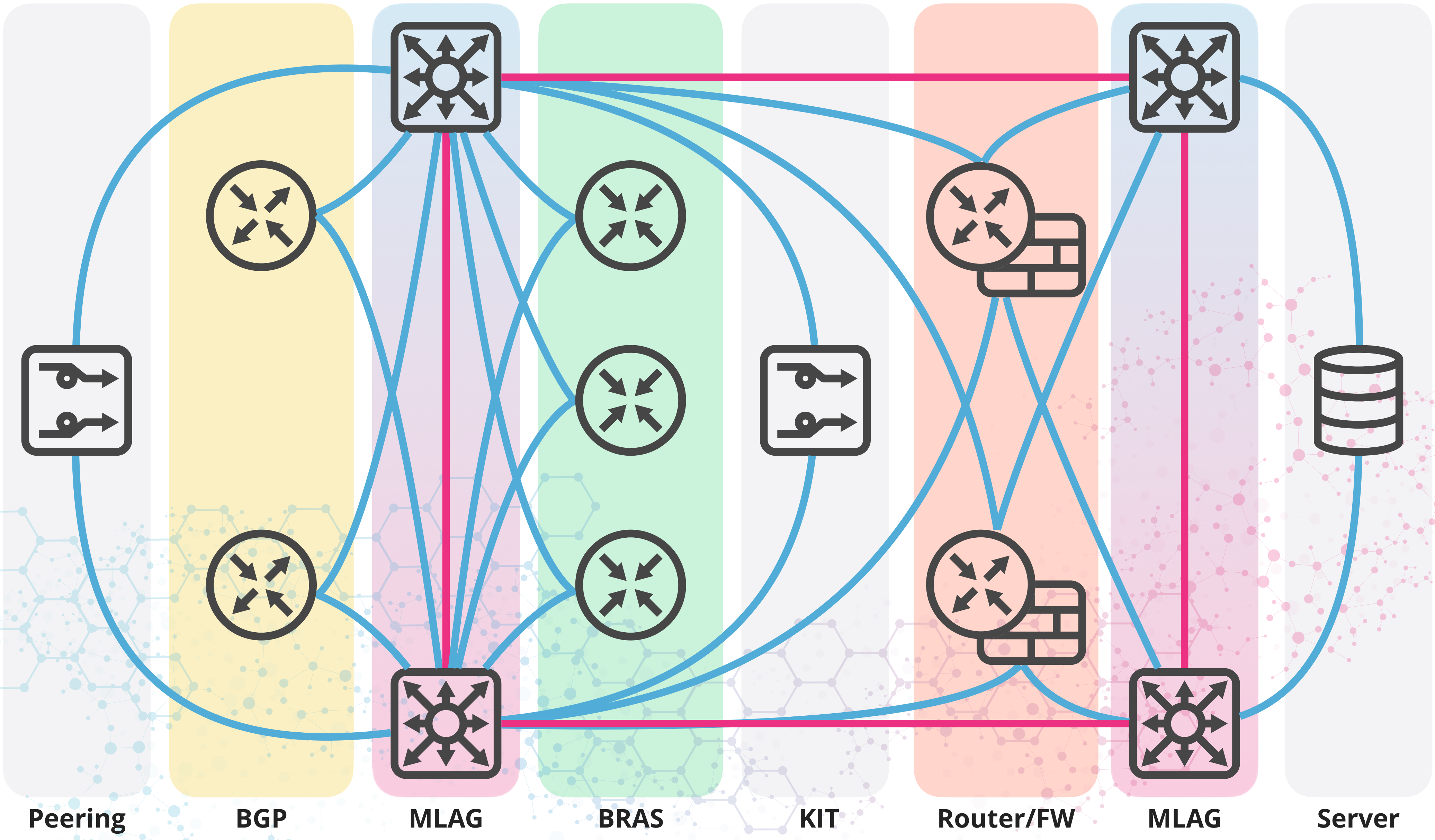
My Approach

The architecture I propose is the result of this experience: a **scalable**, **resilient**, and **high-performance** approach tailored to ISP needs.

A well-designed MikroTik-based network can rival traditional enterprise solutions... **if you know how to use its full potential.**

Module 2
Topology

MikroTik ISP Datacenter Architecture
Designing and Managing an ISP Datacenter with MikroTik



Module 3
MLAG

MLAG: Redundancy Challenges

From Single Points of Failure to a Fully Redundant Network

One of the biggest challenges in ISP networks is ensuring redundancy without sacrificing performance.

Traditional solutions, like Spanning Tree Protocol (STP), block redundant links to prevent loops, but this limits bandwidth utilization and slows down failover in case of a failure.

Imagine a core switch failure in a traditional setup: traffic is disrupted, convergence takes time, and the impact on customers is immediate. ISPs can’t afford this kind of downtime.

MLAG: Redundancy Challenges

From Single Points of Failure to a Fully Redundant Network

This is where Multi-Chassis Link Aggregation (MLAG) changes the game:

- **No More Blocked Links**. Unlike STP, MLAG allows all connections to be active, fully utilizing available bandwidth.
- **Instant Failover**. If a switch or link fails, the other switch takes over immediately, preventing service disruptions.
- **Better Load Balancing**. Traffic is dynamically distributed across multiple links, improving efficiency.
- **Seamless Scalability**. New devices can be integrated without major reconfiguration or downtime.

MLAG: How It Works in RouterOS

Creating a Unified, Redundant Switch Architecture

MLAG (Multi-Chassis Link Aggregation Group) allows two separate MikroTik switches to function as a single logical switch from the perspective of connected devices.

- Devices connected via LACP bonds believe they are linked to a single switch, ensuring uninterrupted operation even if one switch fails.
- Unlike STP, all links remain active, fully utilizing available bandwidth.
- The two switches continuously synchronize their MAC address tables to maintain seamless traffic forwarding.

MLAG: How It Works in RouterOS

How MLAG is Constructed in MikroTik

Two CRS3xx, CRS5xx, CCR2116, or CCR2216 devices form an MLAG pair using RouterOS v7.

The two switches communicate over a dedicated peer port using Inter Chassis Control Protocol (ICCP), sharing MAC tables and system states.

MLAG requires STP, RSTP, or MSTP enabled, and both switches must have the same STP priority and configuration.

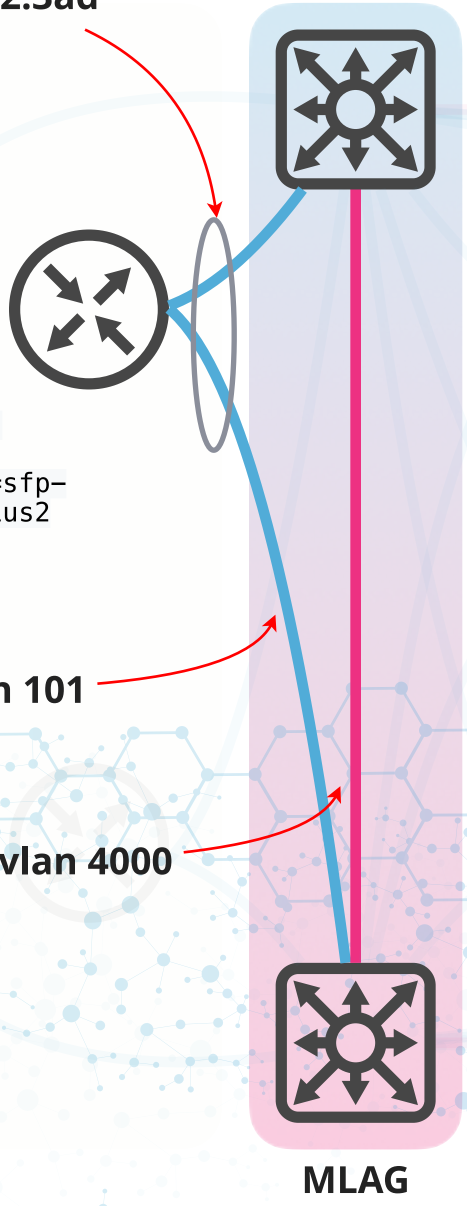
MLAG: How It Works in RouterOS

How MLAG is Constructed in MikroTik

Peer communication should happen over a dedicated untagged VLAN, isolated from the rest of the network using VLAN filtering.

MLAG supports LACP bonding on peer ports, allowing connected devices to distribute traffic across both switches.

The MLAG is not compatible with L3 hardware offloading. When using MLAG, the L3 hardware offloading must be disabled... but only on switches!



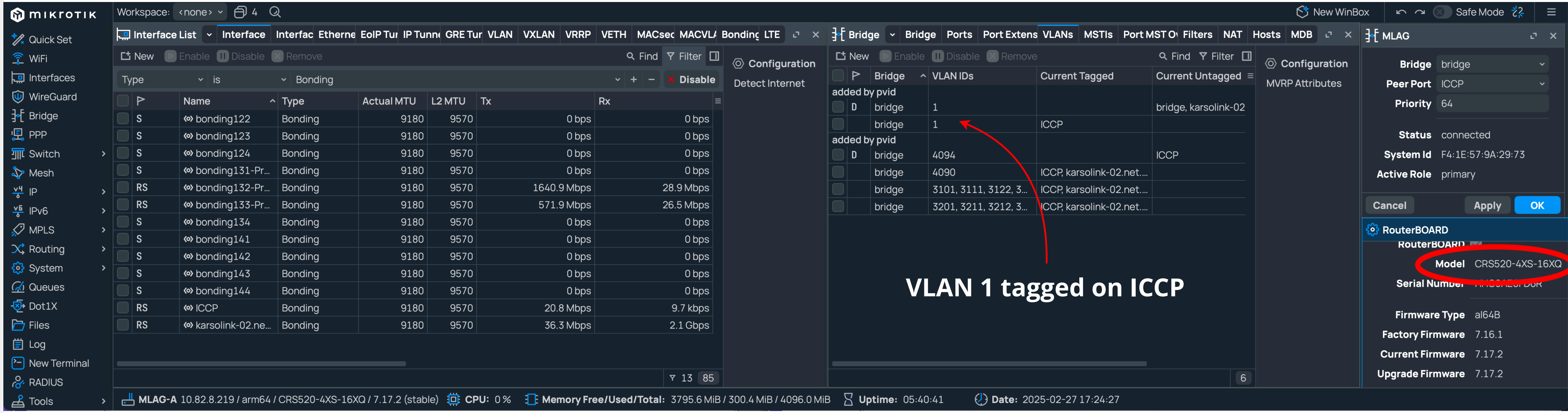

```
/interface bonding
add mode=802.3ad
name=bond1 slaves=sfp-
sfpplus1,sfp-sfpplus2
```

```
/interface bonding
add mlag-id=10 mode=802.3ad name=client-bond slaves=sfp-sfpplus2

/interface bridge
add name=bridge1 vlan-filtering=yes

/interface bridge port
add bridge=bridge1 interface=sfp-sfpplus1-ICCP pvid=4000
add bridge=bridge1 interface=client-bond

/interface bridge vlan
add bridge=bridge1 tagged=sfp-sfpplus1-ICCP vlan-ids=1
add bridge=bridge1 tagged=sfp-sfpplus1-ICCP,client-bond vlan-ids=101

/interface bridge mlag
set bridge=bridge1 peer-port=sfp-sfpplus1-ICCP
```

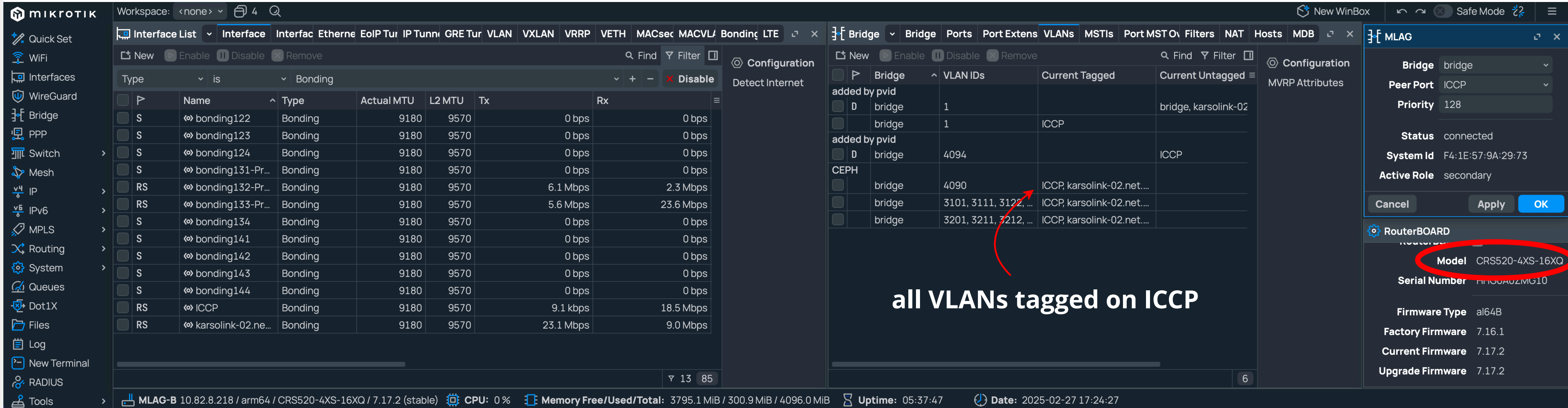
```
/interface bonding
add mlag-id=10 mode=802.3ad name=client-bond slaves=sfp-sfpplus2

/interface bridge
add name=bridge1 vlan-filtering=yes

/interface bridge port
add bridge=bridge1 interface=sfp-sfpplus1-ICCP pvid=4000
add bridge=bridge1 interface=client-bond

/interface bridge vlan
add bridge=bridge1 tagged=sfp-sfpplus1-ICCP vlan-ids=1
add bridge=bridge1 tagged=sfp-sfpplus1-ICCP,client-bond vlan-ids=101

/interface bridge mlag
set bridge=bridge1 peer-port=sfp-sfpplus1-ICCP
```

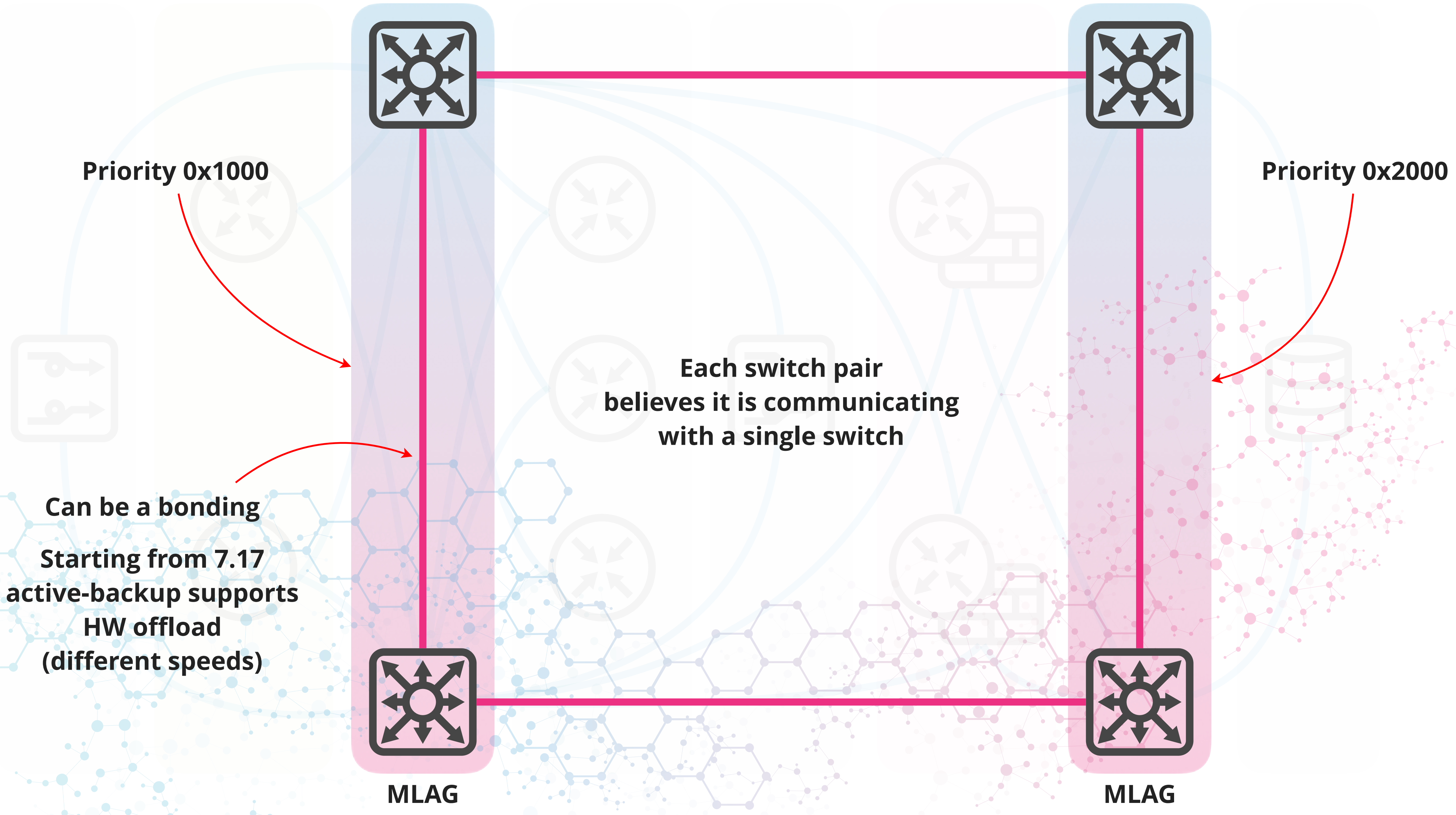
MikroTik ISP Datacenter Architecture
Designing and Managing an ISP Datacenter with MikroTik

Switch 1

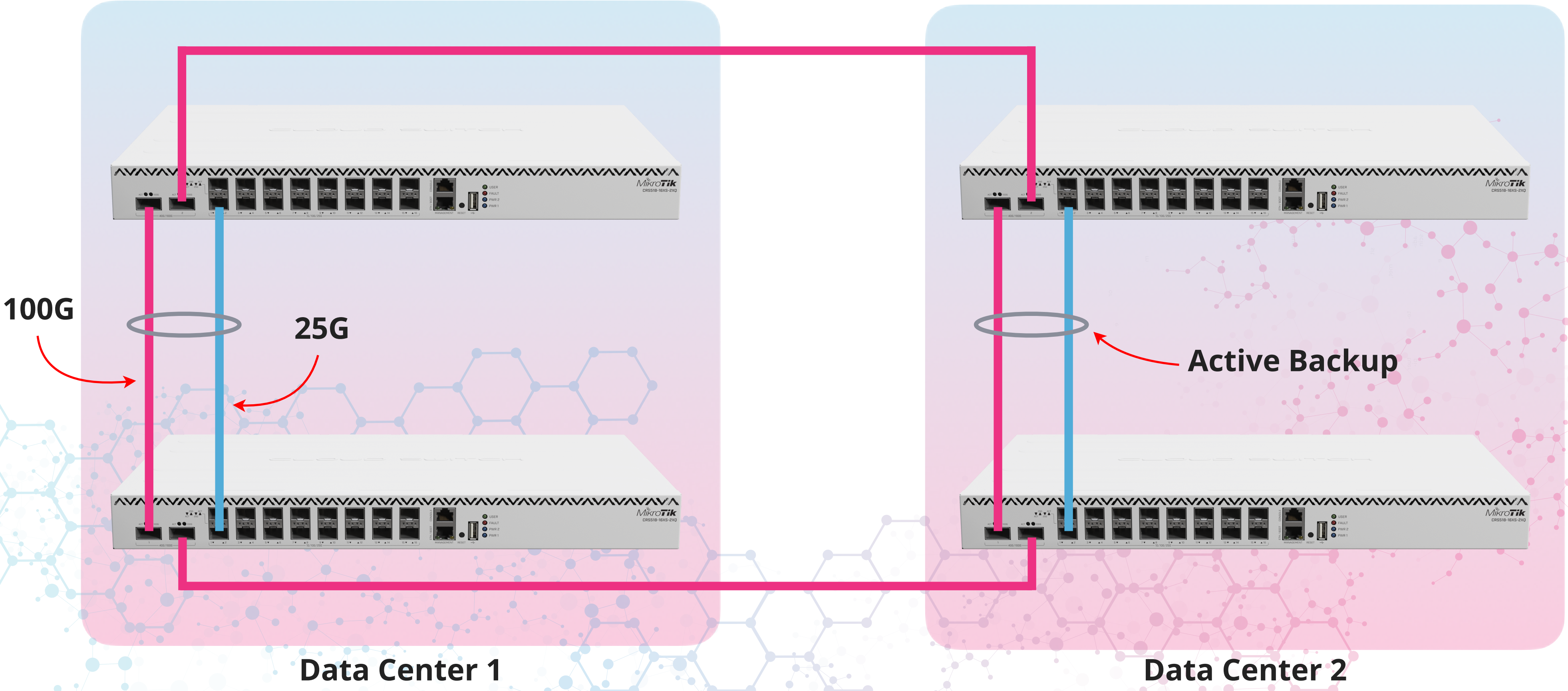


Switch 2




Priority 0x1000
Priority 0x2000
Each switch pair
believes it is communicating
with a single switch
Can be a bonding
Starting from 7.17
active-backup supports
HW offload
(different speeds)
MLAG
MLAG

MLAG: 100G with 25G ICCP failover



Module 4
VLAN

VLAN Design

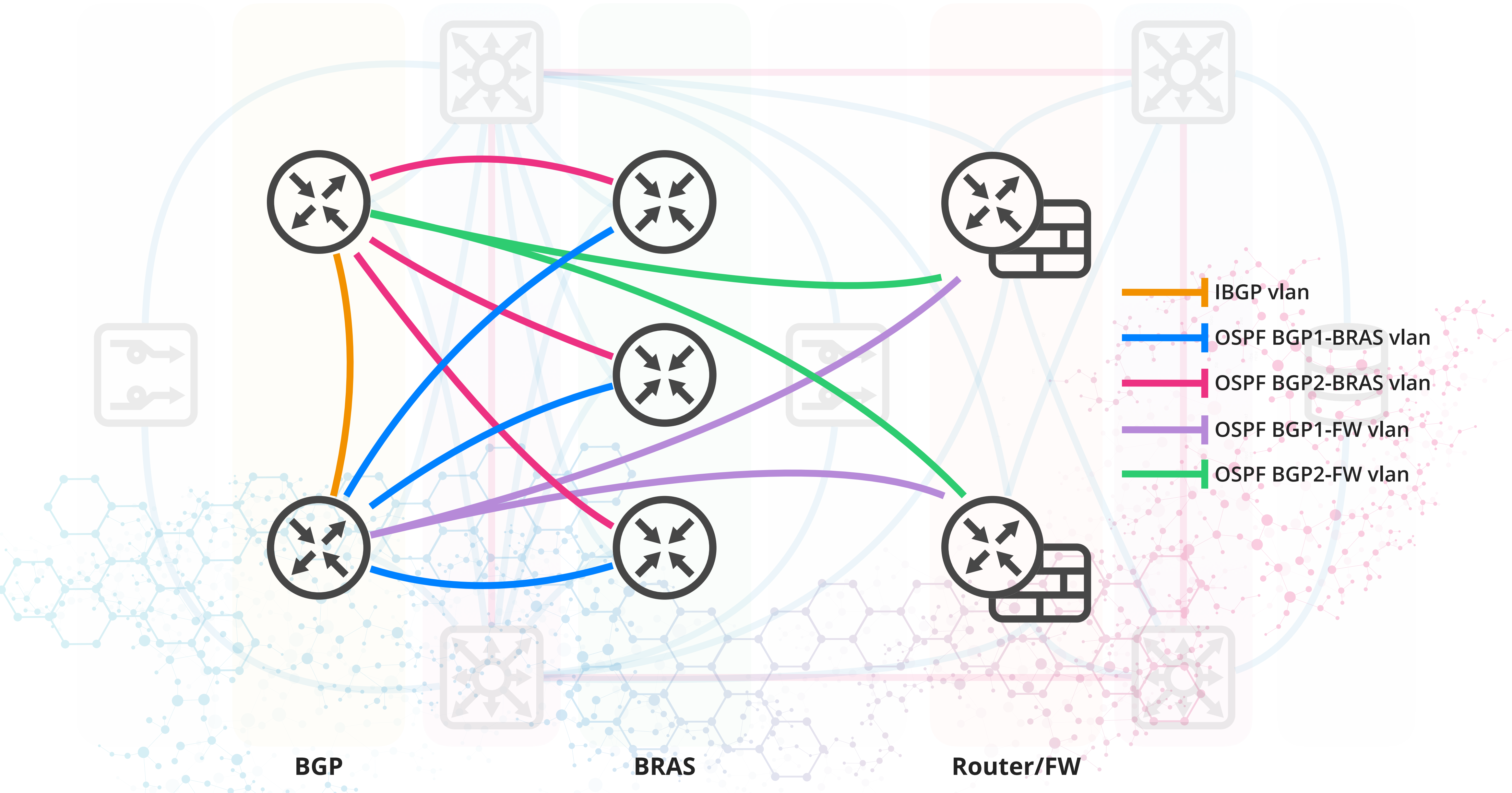
Why VLAN Design Matters in an MLAG Setup

In a MLAG environment, proper VLAN configuration is crucial for ensuring seamless communication between devices.

Unlike traditional setups, where VLAN traffic is confined to a single switch, MLAG allows VLANs to be distributed across both switches in a redundant and scalable way.

- Both MLAG switches must have the same VLAN settings to avoid inconsistencies.
- Ensures traffic is forwarded correctly based on VLAN membership.
- Separating device to device traffic improves security and performance.

MikroTik ISP Datacenter Architecture
Designing and Managing an ISP Datacenter with MikroTik



VLAN Design

Optimizing Network Architecture with VLAN Segmentation

In a network architecture involving **two BGP routers** (BGP1 and BGP2), **three BRAS devices**, and **two firewalls**, implementing five distinct VLANs ensures **efficient routing, redundancy, and security**.

iBGP is used to maintain **dynamic routing updates**, with BGP1 and BGP2 acting as **route reflectors** to optimize scalability and avoid a full-mesh iBGP configuration.

IBGP vlan
OSPF BGP1-BRAS vlan
OSPF BGP2-BRAS vlan
OSPF BGP1-FW vlan
OSPF BGP2-FW vlan

VLAN Design

iBGP Synchronization Between BGP1 and BGP2

This VLAN establishes a direct Layer 2 connection between BGP1 and BGP2, facilitating internal BGP (iBGP) sessions for routing table synchronization.

Assign dedicated interfaces on both routers to this VLAN, configure IP addresses on these interfaces, and establish both IPv4 and IPv6 iBGP peering to ensure consistent routing information.

VLAN Design

Connectivity from BGPs to BRAS Devices and Firewalls

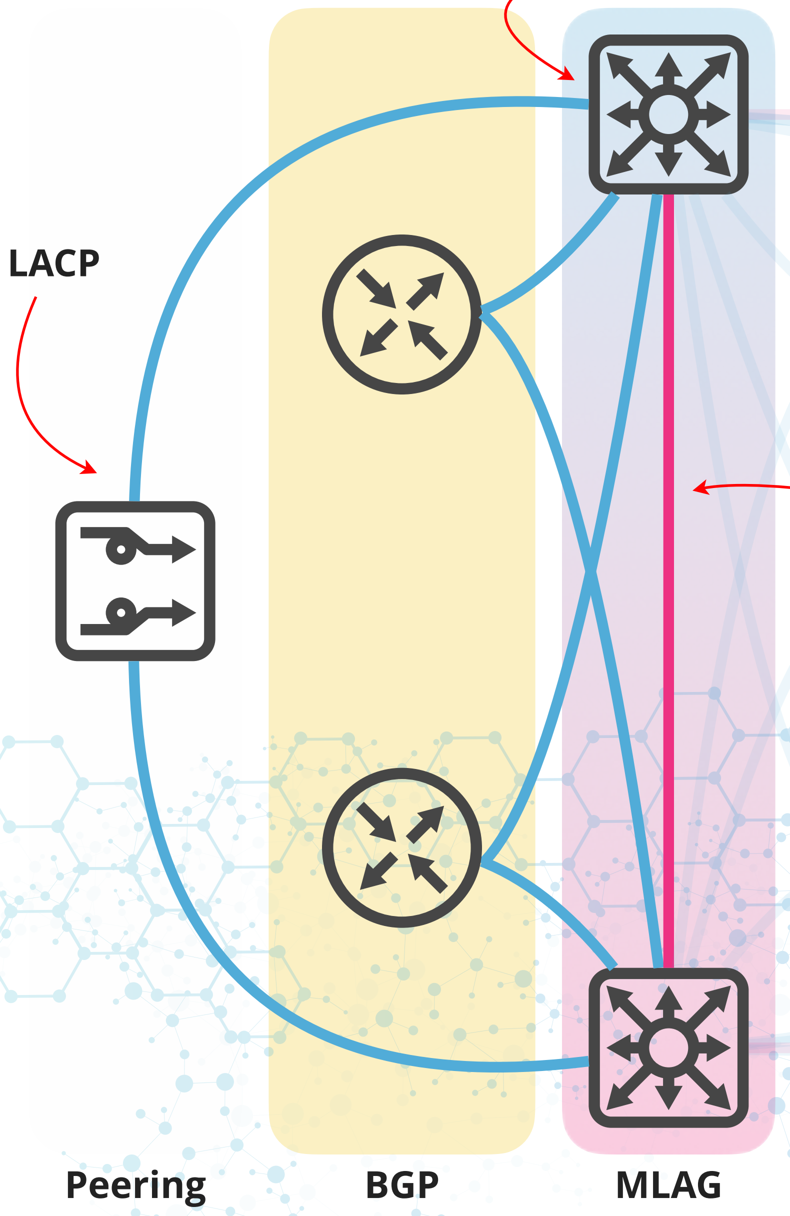
This VLANs connect both BGP routers to all three BRAS devices and to the two Firewalls, allowing the BRAS and the Firewalls to establish iBGP sessions with BGP1 and BGP2 on their loopback addresses.

Since BGP1 and BGP2 are configured as route reflectors, the BRAS devices and Firewalls do not need a full-mesh iBGP connection.

OSPF is enabled on this VLANs but only to distribute loopback IP addresses (IPv4 and IPv6).

This configuration ensures that customer and Datacenter routes are propagated efficiently and BRASs and Firewalls obtain only default routes.

EDGE port, VLAN ingress filtering (admit only untagged), PVID of the Peering VLAN




```
/interface bonding
add mlag-id=20 mode=802.3ad name=peering-bond slaves=sfp-sfpplus3
add mlag-id=30 mode=802.3ad name=bgp1-bond slaves=sfp-sfpplus4
add mlag-id=40 mode=802.3ad name=bgp2-bond slaves=sfp-sfpplus5

/interface bridge port
add bridge=bridge edge=yes frame-types=admit-only-untagged-and-priority-tagged \
   interface=peering-bond pvid=101
add bridge=bridge interface=bgp1-bond
add bridge=bridge interface=bgp2-bond

/interface bridge vlan
add bridge=bridge1 tagged=sfp-sfpplus1-ICCP,peering-bond,bgp1-bond,bgp2-bond \
   vlan-ids=101
```

Peering VLAN
untagged on Switches EDGE ports
tagged on ICCP and bonded interfaces


```
/interface bonding
add mlag-id=20 mode=802.3ad name=peering-bond slaves=sfp-sfpplus3
add mlag-id=30 mode=802.3ad name=bgp1-bond slaves=sfp-sfpplus4
add mlag-id=40 mode=802.3ad name=bgp2-bond slaves=sfp-sfpplus5

/interface bridge port
add bridge=bridge edge=yes frame-types=admit-only-untagged-and-priority-tagged \
   interface=peering-bond pvid=101
add bridge=bridge interface=bgp1-bond
add bridge=bridge interface=bgp2-bond

/interface bridge vlan
add bridge=bridge1 tagged=sfp-sfpplus1-ICCP,peering-bond,bgp1-bond,bgp2-bond \
   vlan-ids=101
```

VLAN Design

Configuring Peering Interfaces

In a peering setup, the configuration of physical interfaces plays a critical role in ensuring stability, security, and flexibility. Peering connections can be established in two main ways:

- **Using LACP (Link Aggregation Control Protocol)** – When multiple links are available, LACP can be used to bundle them into a single logical interface, increasing bandwidth and providing failover capabilities.
- **Single Port Assignments** – If different peering partners need to be assigned to specific switches, each connection can be terminated on separate interfaces, distributing the load across the MLAG pair.

VLAN Design

Configuring Peering Ports in Edge Mode

To maintain stability and security, peering ports should be set as Edge Ports (also called PortFast in other networking vendors).

- **No STP Convergence Delays** – Edge ports immediately transition to forwarding state without going through STP listening/learning phases, reducing downtime when links are flapped.
- **No Risk of Loops** – Since peering ports connect directly to routers, they do not need to participate in STP calculations. If deemed appropriate, BPDU Guard can be enabled.

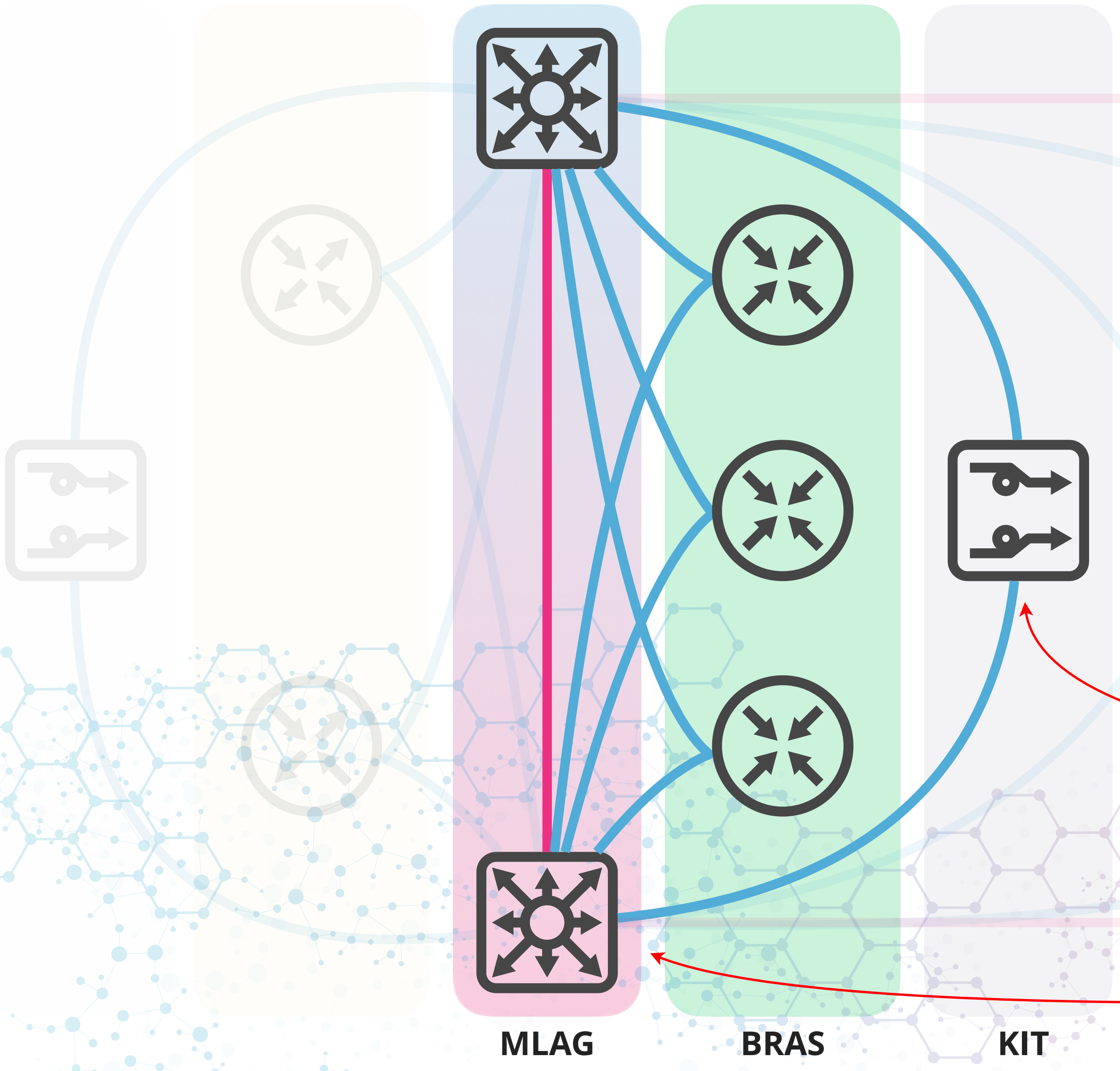
VLAN Design

Filtering VLANs and Allowing Only Untagged Traffic

For security and operational consistency, ingress VLAN filtering should be applied, ensuring that only untagged traffic is allowed on peering interfaces.

This prevents unauthorized VLANs from being propagated across the network and ensures that each peering session is handled in its designated VLAN.

- Peering VLAN traffic should be untagged to avoid inconsistencies between our network and the peering partner.
- VLAN filtering prevents accidental trunking of VLANs that could interfere with traffic segmentation.




```
/interface bonding
add mlag-id=50 mode=802.3ad name=kit-bond \
   slaves=sfp-sfpplus6
add mlag-id=60 mode=802.3ad name=bras1-bond \
   slaves=sfp-sfpplus7
add mlag-id=70 mode=802.3ad name=bras2-bond \
   slaves=sfp-sfpplus8
add mlag-id=80 mode=802.3ad name=bras2-bond \
   slaves=sfp-sfpplus9

/interface bridge port
add bridge=bridge edge=yes tag-stacking=yes\
   frame-types=admit-only-vlan-tagged \
   interface=kit-bond pvid=201
add bridge=bridge interface=bras1-bond
add bridge=bridge interface=bras2-bond
add bridge=bridge interface=bras3-bond

/interface bridge vlan
add bridge=bridge tagged=sfp-sfpplus1-ICCP,
   kit-bond,bras1-bond,bras2-bond,bras3-bond \
   vlan-ids=201
```

LACP

EDGE port, tag-stacking
VLAN ingress filtering (admit only tagged)
PVID of the KIT VLAN

VLAN Design

Configuring Interfaces for Wholesale Delivery Kits

In wholesale service delivery, the provider may hand over VLANs in two ways:

- **QinQ (802.1ad)** – Where the Service VLAN (S-VLAN) encapsulates the Customer VLAN (C-VLAN), commonly used for wholesale access.
- **Double 802.1Q VLAN (802.1q)** – Where only a single VLAN tag type is assigned, either for the service or the customer.

Our objective is to transport these VLANs transparently to the BRAS without altering their structure, allowing correct end-to-end VLAN mapping for the service provider and the customer.

To achieve this, we implement **Tag-Stacking** and **accept only tagged packets**.

Module 5
L3 Hardware Offloading

L3 Hardware Offloading

Enhancing Routing Performance

Layer 3 Hardware Offloading (L3HW) enables MikroTik devices to **delegate routing tasks from the main CPU to a dedicated switch chip**.

This transition significantly boosts network performance, allowing for high-speed data transmission that would otherwise strain the CPU.

- **Increased Throughput**
- **Reduced Latency**
- **Optimized CPU Utilization**
- **Improved Scalability**

L3 Hardware Offloading

Prerequisites

L3HW is available on devices equipped with specific switch chips (CRS3xx Series Switches, CRS5xx Series Switches, CCR2116 and CCR2216 Routers).

Some functionalities, such as **MLAG**, **VRF** and **VRRP**, are incompatible with L3HW and must be disabled to utilize hardware offloading.

These three features are essential, can L3HW still be used?

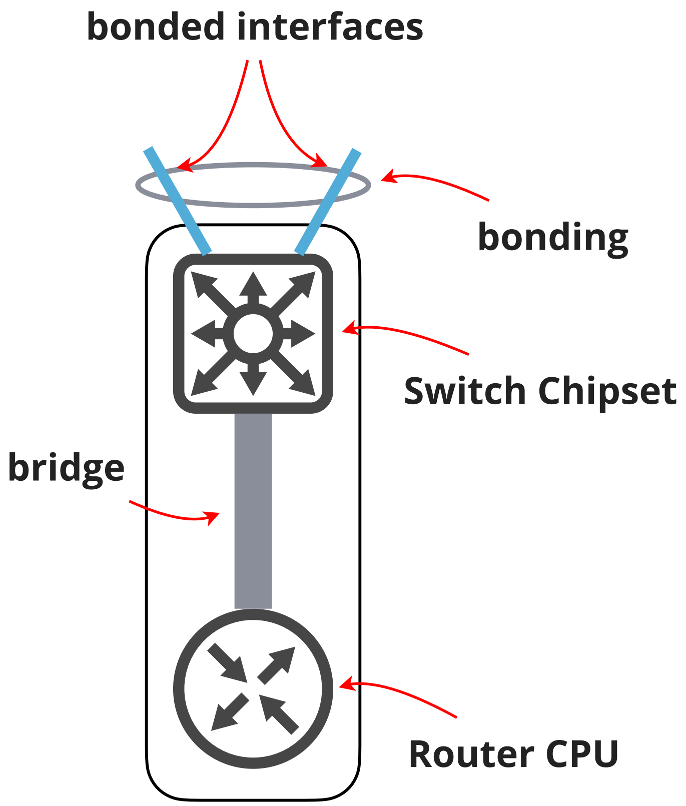
L3 Hardware Offloading

Tip

Visualize your router as comprising two components:

- **Switch Chipset**: Manages Layer 2 (L2) traffic, functioning similarly to a dedicated switch.
- **Router CPU**: Handles Layer 3 (L3) traffic, performing routing operations.

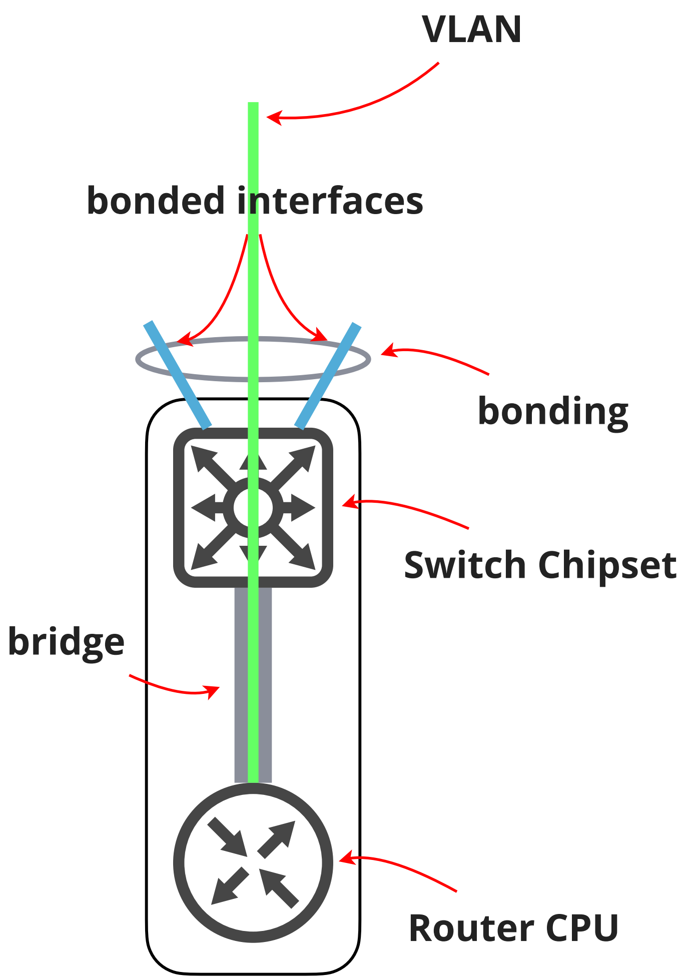
These elements are integrated within a single chassis, collaborating to provide both switching and routing functionalities.



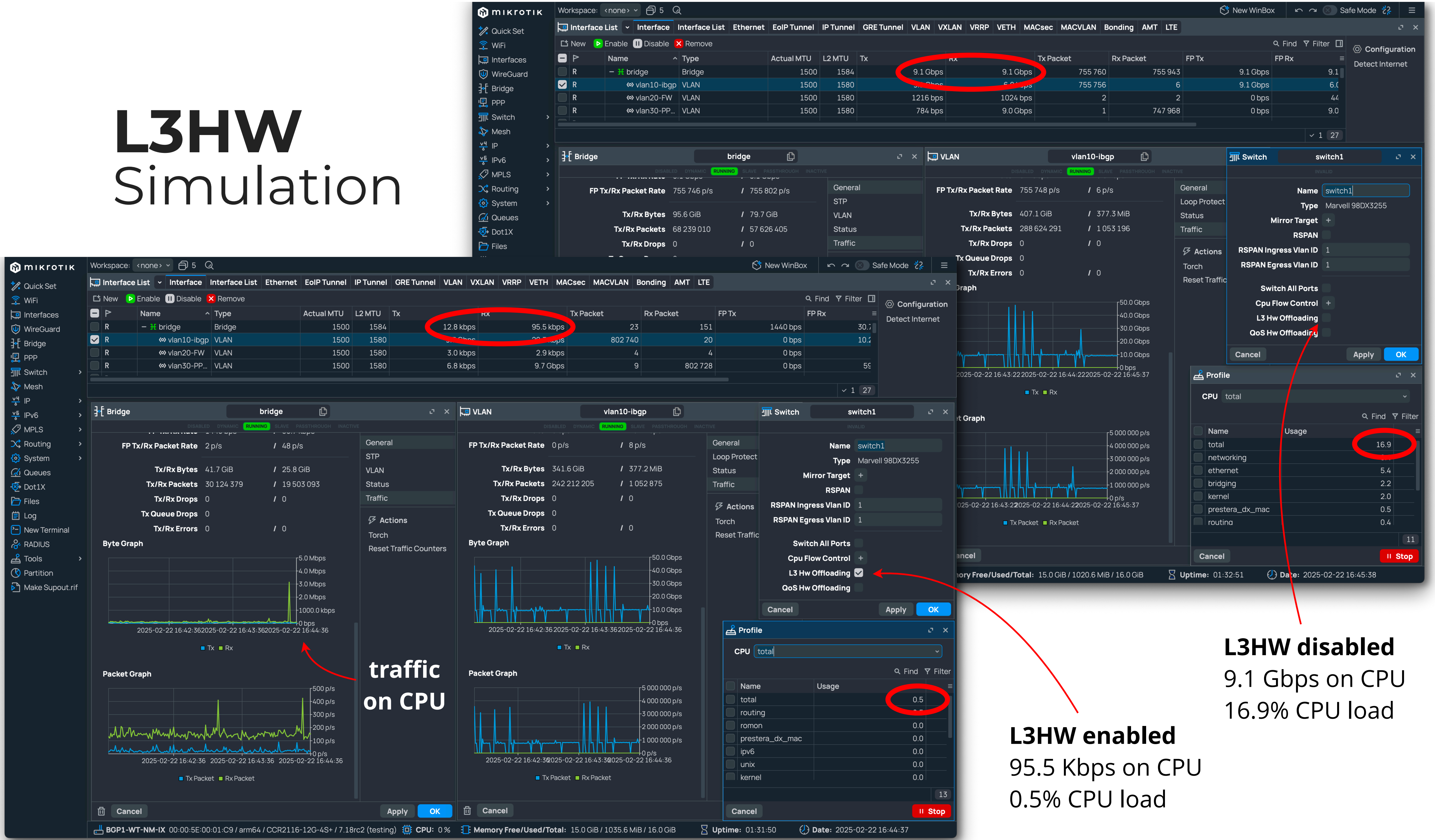
L3 Hardware Offloading

L3HW inter-VLAN routing

- Configure a bridge to serve as the central point for traffic management.
- Add only the bonding interface to the bridge.
- Activate VLAN filtering on the bridge.
- Assign VLAN tags to both the bonding interface and the bridge
- Define VLAN interfaces on the bridge.
- Enable L3HW.

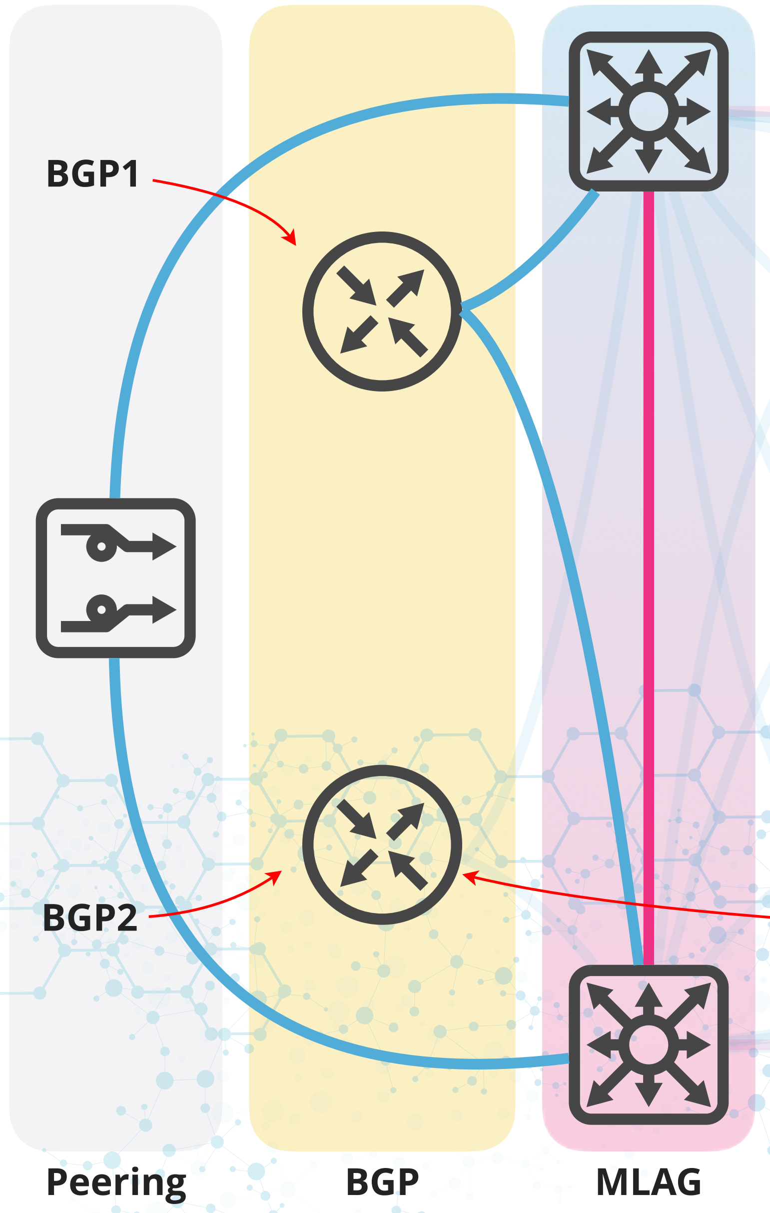


L3HW
Simulation



L3HW enabled
95.5 Kbps on CPU
0.5% CPU load

L3HW disabled
9.1 Gbps on CPU
16.9% CPU load




```
/interface bridge
add name=bridge protocol-mode=mstp vlan-filtering=yes

/interface bonding
add lacp-rate=1sec mode=802.3ad name=MLAG slaves=sfp-sfpplus1,sfp-sfpplus2

/interface bridge port
add bridge=bridge interface=MLAG

/interface vlan
add interface=bridge name=vlan10-ibgp vlan-id=10
add interface=bridge name=vlan20-Firewall vlan-id=20
add interface=bridge name=vlan30-BRAS vlan-id=30
add interface=bridge name=vlan101-Peering1 vlan-id=101
add interface=bridge name=vlan102-Peering2 vlan-id=102

/interface bridge vlan
add bridge=bridge comment=ibgp tagged=bridge,MLAG vlan-ids=10
add bridge=bridge comment=Firewall tagged=bridge,MLAG vlan-ids=20
add bridge=bridge comment=BRAS tagged=MLAG,bridge vlan-ids=30
add bridge=bridge comment=Peering1 tagged=MLAG,bridge vlan-ids=101
add bridge=bridge comment=Peering2 tagged=bridge,MLAG vlan-ids=102

/interface ethernet switch l3hw-settings set autorestart=yes ipv6-hw=yes

/interface ethernet switch set 0 l3-hw-offloading=yes
```

Same configuration on BGP2

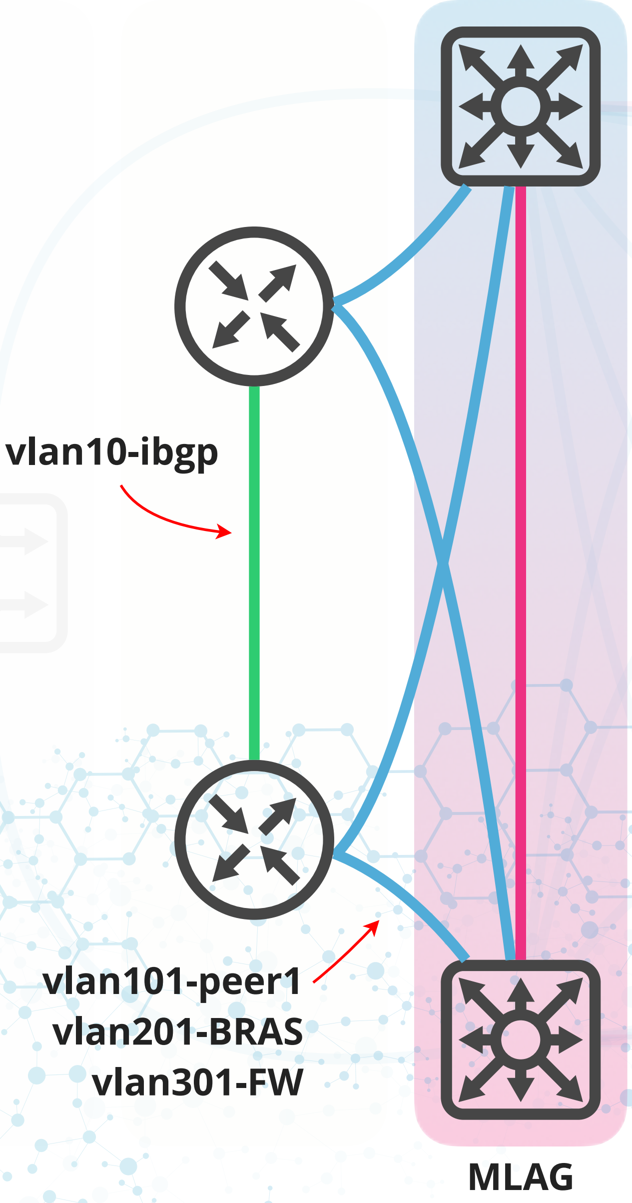
L3 Hardware Offloading

iBGP Through Direct Connection Interface

L3HW requires the connection interface to be in the bridge to enable hardware-accelerated routing for iBGP sessions.

However, adding the interface to the bridge causes STP to block the port, preventing VLAN traffic from passing through.

Multiple Spanning Tree Protocol (MSTP) is required to allow the iBGP VLAN to traverse the interconnection interface while maintaining loop prevention and efficient traffic routing.




```
/interface bonding
add mode=802.3ad name=MLAG slaves=sfp-sfpplus1,sfp-sfpplus2
add mode=802.3ad name=BGP1-BGP2 slaves=sfp-sfpplus3,sfp-sfpplus4

/interface bridge
add name=bridge1 protocol-mode=mstp vlan-filtering=yes \
   region-name=MLAG region-revision=1

/interface bridge port
add bridge=bridge1 interface=MLAG
add bridge=bridge1 interface=BGP1-BGP2

/interface vlan
add interface=bridge1 name=vlan101 vlan-id=101
add interface=bridge1 name=vlan201 vlan-id=201
add interface=bridge1 name=vlan301 vlan-id=301
add interface=bridge1 name=vlan10-ibgp vlan-id=10

/interface bridge vlan
add bridge=bridge1 tagged=MLAG,bridge1 vlan-ids=101
add bridge=bridge1 tagged=MLAG,bridge1 vlan-ids=201
add bridge=bridge1 tagged=MLAG,bridge1 vlan-ids=301
add bridge=bridge1 tagged=BGP1-BGP2,bridge1 vlan-ids=10

/interface bridge msti
add bridge=bridge1 identifier=1 priority=0x1000 vlan-mapping=10
add bridge=bridge1 identifier=1 priority=0x8000 vlan-mapping=101,201,301
```

MSTP is essential otherwise VLAN 10 must pass through the switches

L3 Hardware Offloading

L3HW does not support VRRP, but...

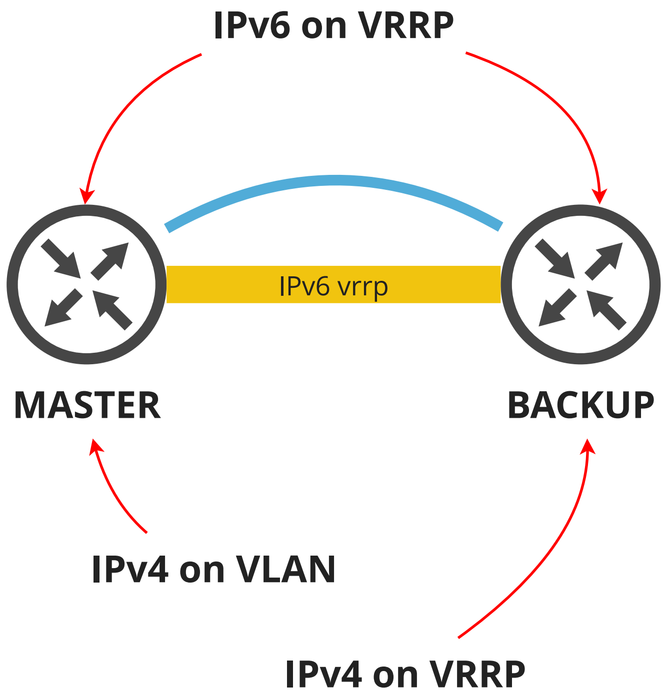
Establish VRRP using IPv6 link-local addresses.

Assign the IPv6 peering address to the VRRP interfaces on both Master and Backup routers.

On Master, configure the IPv4 address directly on the VLAN interface associated with VRRP.

On Backup, assign the IPv4 address to the VRRP interface.

This setup allows all IPv4 routes on the Master to be offloaded to hardware, optimizing performance.



MikroTik ISP Datacenter Architecture

Designing and Managing an ISP Datacenter with MikroTik

Workspace: <none> 2 | New WinBox | Safe Mode

MIKROTIK

- Quick Set
- WiFi
- Interfaces
- WireGuard
- Bridge
- PPP
- Switch
- Mesh
- IP
- IPv6
- MPLS
- Routing
- System
- Queues
- Dot1X
- Files
- Log
- New Terminal
- RADIUS
- Tools
- Partition
- Make Supout.rif

Route List

New | Enable | Disable | Remove | Find | Filter | all

Active is yes

	Dst. Address	Distance	Routing Table	AS Path
DAbH	1.6.146.0/24	20	main	41327,6453,4755,9583
DAbH	1.6.147.0/24	20	main	41327,9583
DAbH	1.6.148.0/22	20	main	41327,9583
DAbH	1.6.152.0/22	20	main	41327,9583
DAbH	1.6.156.0/22	20	main	41327,9583
DAbH	1.6.156.0/24	20	main	41327,9583
DAbH	1.6.157.0/24	20	main	41327,9583
DAbH	1.6.158.0/24	20	main	41327,9583
DAbH	1.6.160.0/22	20	main	41327,9583
DAbH	1.6.161.0/24	20	main	41327,1299,9583
DAbH	1.6.163.0/24	20	main	41327,3320,9583
DAbH	1.6.164.0/22	20	main	41327,9583
DAbH	1.6.165.0/24	20	main	41327,3320,9583
DAbH	1.6.166.0/24	20	main	41327,6453,4755,9583
DAbH	1.6.167.0/24	20	main	41327,9583
DAbH	1.6.168.0/22	20	main	41327,9583
DAbH	1.6.168.0/24	200	main	28716,6453,4755,9583
DAbH	1.6.169.0/24	200	main	28716,6453,4755,9583
DAbH	1.6.172.0/22	20	main	41327,1299,9583
DAbH	1.6.172.0/24	20	main	41327,2914,9583
DAbH	1.6.176.0/22	20	main	41327,9583
DAbH	1.6.177.0/24	20	main	41327,1299,9583
DAbH	1.6.180.0/22	20	main	41327,1299,9583
DAbH	1.6.183.0/24	200	main	28716,6453,4755,9583
DAbH	1.6.184.0/22	20	main	41327,9583
DAbH	1.6.187.0/24	20	main	41327,3320,9583
DAbH	1.6.188.0/22	20	main	41327,9583
DAbH	1.6.188.0/24	20	main	41327,9583
DAbH	1.6.191.0/24	20	main	41327,9583
DAbH	1.6.192.0/22	20	main	41327,9583
DAbH	1.6.194.0/24	20	main	41327,9583
DAbH	1.6.195.0/24	20	main	41327,9583
DAbH	1.6.196.0/22	20	main	41327,9583
DAbH	1.6.197.0/24	20	main	41327,1299,9583
DAbH	1.6.198.0/24	20	main	41327,1299,9583

L3HW

1 | 921060 | 1408781

1.4M routes

L3 Hw Settings

General | Advanced | Monitor

Error Code	0
State	OK
IPv4 Routes Total	971263
IPv4 Routes HW	28614
IPv4 Routes CPU	942736
IPv4 Shortest HW Prefix	24
IPv4 Hosts	175
IPv6 Routes Total	12128
IPv6 Routes HW	322
IPv6 Routes CPU	11805
IPv6 Shortest HW Prefix	48
IPv6 Hosts	114
Route Queue Size	0
Route Queue Rate	2
Route Process Rate	2
Nexthop Cap	8192
Nexthop Usage	310
VXLAN MTU Drop Packet	0
FastTrack IPv4 Conns	0
FastTrack Queue Size	0
FastTrack Queue Rate	0
FastTrack Process Rate	0
FastTrack HW Min Speed	0
FastTrack HW Offloaded	0
FastTrack HW Unloaded	0
LPM Cap	6720
LPM Usage	6535
LPM Bank Cap	336
LPM Bank Usage	336
	336

actual offloaded prefixes

Cancel | Apply | OK

BGP1-WT-NM-IX 00:00:5E:00:01:C9 / arm64 / CCR2116-12G-4S+ / 7.18rc2 (testing) | CPU: 1 % | Memory Free/Used/Total: 15.0 GiB / 1041.6 MiB / 16.0 GiB | Uptime: 01:03:49 | Date: 2025-02-22 16:16:31

L3 Hardware Offloading

shortest-hw-prefix

If the entire routing table does not fit into the hardware memory, partial offloading is applied.

In this scenario, the longest prefixes are hardware-offloaded, while shorter ones are redirected to the CPU. ***(/24 is shorter than /16)***

This parameter shows the shortest route prefix (/x) that is offloaded to the hardware memory, all prefixes shorter than this are processed by the CPU.

An "ipv4-shortest-hw-prefix=0" value indicates that the entire routing table is offloaded to the hardware memory.

L3 Hardware Offloading

Efficient Management of IPv4 and IPv6 Prefixes

In high-performance networking, the ability of a router to manage a substantial number of IPv4 and IPv6 prefixes is crucial.

CCR2216 and CCR2116 models both offer Layer 3 Hardware Offloading (L3HW) capabilities to optimize routing. However, they differ significantly in their prefix handling capacities.

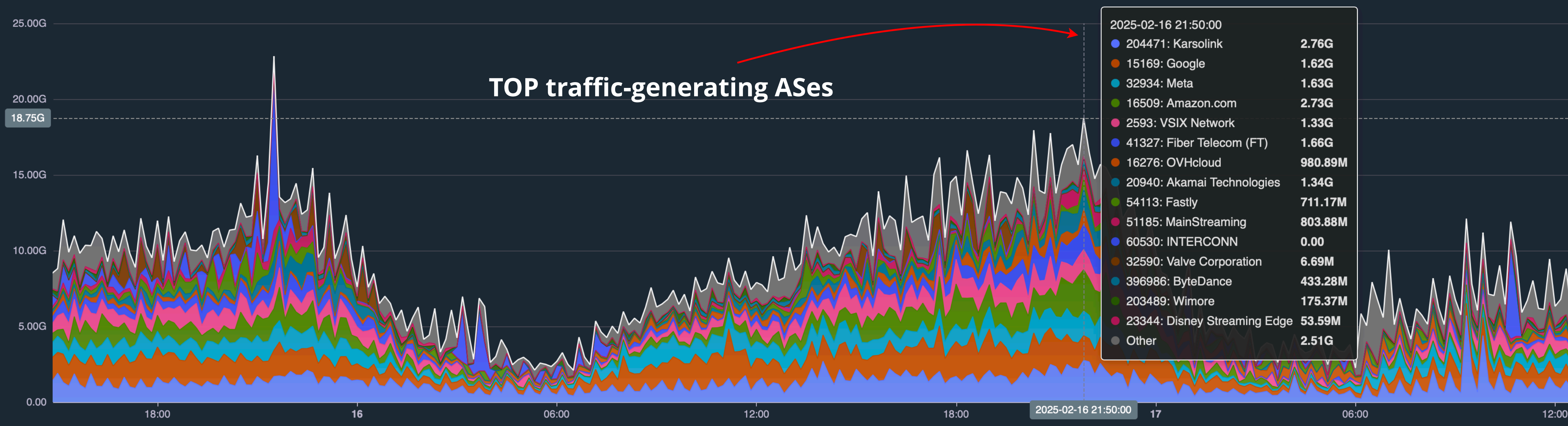
Model	Offloadable IPv4 Prefixes	Offloadable IPv6 Prefixes
CCR2116-12G-4S+	16,000 - 36,000	4,000 - 6,000
CCR2216-1G-12XS-2XQ	60,000 - 120,000	15,000 - 20,000

MikroTik ISP Datacenter Architecture
Designing and Managing an ISP Datacenter with MikroTik

Akvorado
v1.11.2-8-g580...

Home | Visualize | Documentation

February 15, 2025 2:52 PM — February 17, 2025 2:52 PM | Stacked areas | 15 | L3% | SrcAS | InIfBoundary = external AND SrcAS != 0



SRC AS	MIN	MAX	AVERAGE
204471: Karsolink	222.01Mbps	5.71Gbps	1.26Gbps
15169: Google	150.45Mbps	3.26Gbps	1.24Gbps
32934: Meta	53.31Mbps	1.71Gbps	859.45Mbps
16509: Amazon.com	19.64Mbps	3.56Gbps	733.83Mbps
2593: VSIX Network	4.37Mbps	2.11Gbps	707.31Mbps
41327: Fiber Telecom (FT)	1.91Mbps	1.74Gbps	457.80Mbps
16276: OVHcloud	11.15Mbps	1.95Gbps	394.57Mbps
20940: Akamai Technologies	5.44Mbps	1.65Gbps	313.99Mbps
54113: Fastly	1.91Mbps	2.45Gbps	294.41Mbps
51185: MainStreaming	31.46Mbps	1.35Gbps	261.68Mbps
60530: INTERCONN	5.00bps	10.15Gbps	257.91Mbps
32590: Valve Corporation	117.00bps	5.28Gbps	206.87Mbps
396986: ByteDance	47.75Kbps	433.28Mbps	146.22Mbps

L3 Hardware Offloading

Efficient Management of IPv4 and IPv6 Prefixes

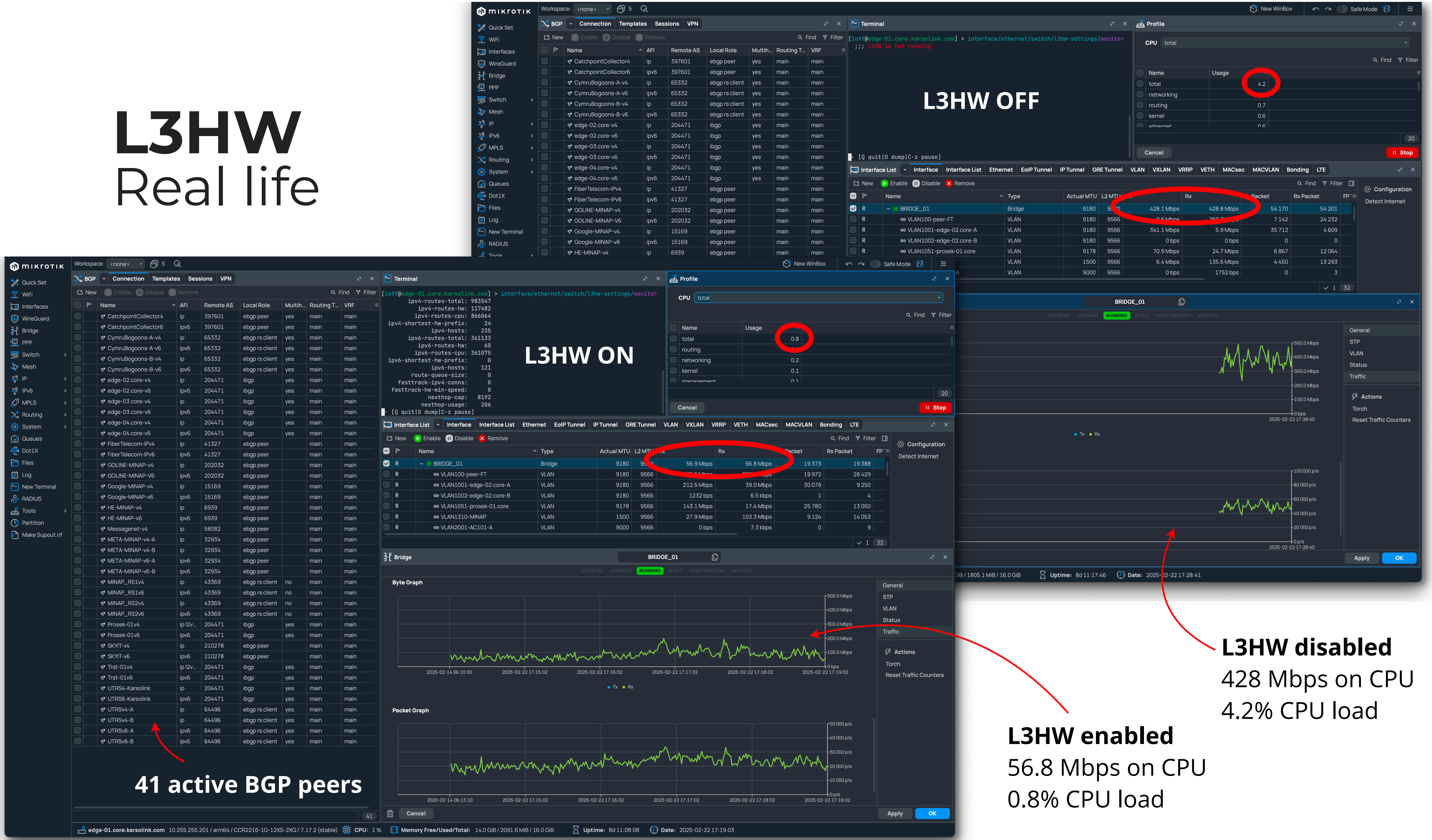
Identify High-Traffic Prefixes

- Monitor network traffic to pinpoint prefixes that generate significant load.
- Focus on Autonomous Systems (AS) or specific IP ranges that are critical to your operations.

Implement Routing Filters

- Utilize routing filters to control which routes are offloaded to hardware.
- Apply filters to selectively enable or disable hardware offloading based on criteria such as AS numbers or prefix lengths.

L3HW
Real life



L3HW disabled
428 Mbps on CPU
4.2% CPU load

L3HW enabled
56.8 Mbps on CPU
0.8% CPU load

L3 Hardware Offloading

L3 Hardware Offloading and VRF Compatibility

L3HW is not compatible with VRF, which means all user traffic must pass through the main VRF.

Despite this limitation, it is possible to maintain security and network segmentation by using a separate VRF for management traffic.

By creating a dedicated VRF for the management interface and moving all access services (like SSH, WinBox, and SNMP) to this VRF, management traffic remains isolated from user data.

L3 Hardware Offloading

Leveraging Layer 3 Hardware Offloading with NAT

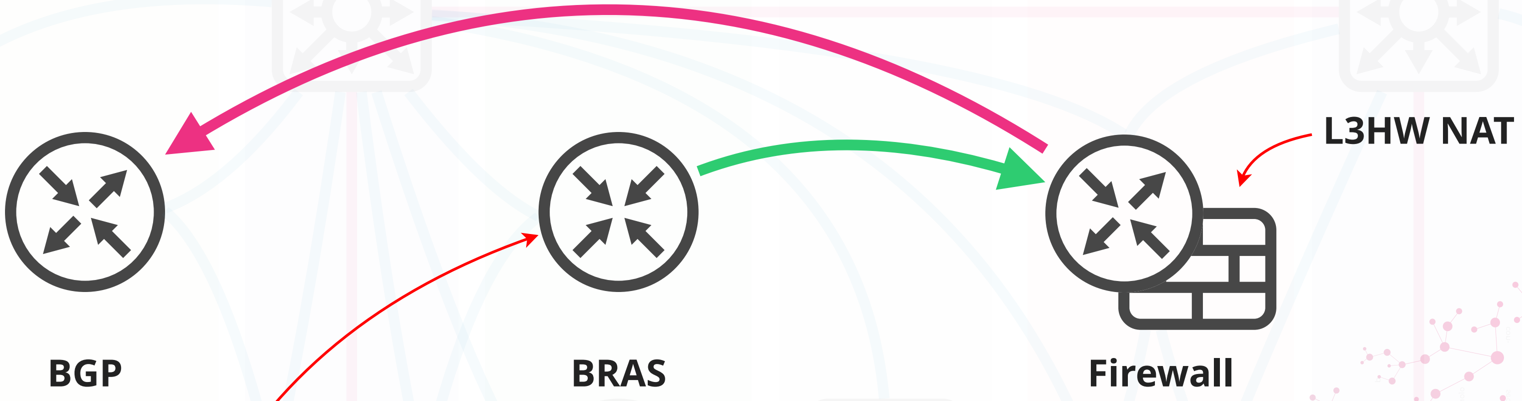
NAT operations are typically processed by the CPU.

To offload NAT processing to hardware, connections must be eligible for FastTrack. Without FastTrack, NAT traffic remains CPU-bound.

When the HW limit of NAT entries is reached, other connections will fall back to the CPU. MikroTik's smart connection offload algorithm ensures that the connections with the most traffic are offloaded to the hardware.

CCR2116 supports up to 2.25K NAT entries, while CCR2216 supports up to 4K.

Starting from version 7.18, VXLAN can be offloaded. However, VTEPs are not supported on bonded or VLAN interfaces making it challenging to use MLAG.



L3HW cannot be used on BRAS because PPPoE interfaces cannot be offloaded

```
/interface bridge
add name=bridge1 vlan-filtering=yes

/interface bridge port
add bridge=bridge1 interface=sfp-plus1 pvid=1000
add bridge=bridge1 interface=sfp-plus2 pvid=2000

/interface bridge vlan
add bridge=bridge1 tagged=bridge1 vlan-ids=1000
add bridge=bridge1 tagged=bridge1 vlan-ids=2000

/interface vlan
add interface=bridge1 name=vlan1000 vlan-id=1000
add interface=bridge1 name=vlan2000 vlan-id=2000

/ip firewall nat add action=masquerade chain=srcnat out-interface=vlan2000

/ip firewall filter add action=fasttrack-connection chain=forward

/interface/ethernet/switch set switch1 l3-hw-offloading=yes
```

Module 6
HW QoS

HW QoS

Optimizing Traffic Flow and Performance

In ISP datacenters, Quality of Service (QoS) is essential for managing network traffic efficiently.

It ensures that critical applications like voice, video, and real-time services receive priority and low latency, while other traffic types are managed without impacting overall performance.

MikroTik implements QoS using Marvell Prestera DX switch chips, supporting advanced features from RouterOS v7.15+.

HW QoS

Key Features

- **QoS Marking**: Classifies traffic at ingress, altering headers to prioritize packets.
- **QoS Enforcement**: Manages congestion with traffic shaping and rate limiting.
- **Active Queue Management (AQM)**: Utilizes WRED and ECN to prevent buffer overflow and maintain low latency.
- **Priority Flow Control (PFC)**: Allocates bandwidth for high-priority traffic, minimizing packet loss.

Module 7
Conclusions

Conclusions

MLAG and Resilience

MLAG provides high availability and redundancy by enabling dual-homing without STP blocking.

Precise VLAN and MSTP configurations are essential to avoid loops and ensure efficient traffic flow.

Conclusions

L3 Hardware Offloading (L3HW)

L3HW significantly enhances routing performance, but compatibility limitations exist with VRF, VRRP, and PPPoE.

Careful design is needed to leverage its benefits without compromising functionality.

Selective route offloading optimizes hardware utilization, particularly on CCR2116.

Conclusions

NAT and L3HW Integration

NAT offloading is supported but limited to 2.25K entries on CCR2116 and 4K on CCR2216.

A strategic approach is needed to balance FastTrack and non-FastTrack traffic, ensuring efficient resource utilization.

Conclusions

QoS and Traffic Prioritization

MikroTik allows the creation of advanced QoS systems suitable for complex ISP environments.

However, due to the variety of QoS mechanisms, it is crucial to evaluate and test different QoS models to find the most effective solution for specific network needs.

Deep knowledge and continuous tuning are required to balance latency, throughput, and resource usage effectively.

Conclusions

VXLAN Support and Future Evolution

VXLAN support has just been released in RouterOS 7.18, marking a significant step forward for network virtualization.

However, VTEPs are not supported on bonded or VLAN interfaces, complicating MLAG implementations.

This release paves the way for EVPN, indicating MikroTik's commitment to evolving towards next-generation data center networking.

Conclusions

Final Note: A Proven Yet Personal Approach

This network architecture is the result of my personal approach to designing ISP datacenters using MikroTik's advanced capabilities. It is built on years of hands-on experience, continuous optimization, and adapting to MikroTik's evolving features.

Although this method reflects my unique perspective.

This approach is not just theoretical, it's battle-tested and continues to evolve with each deployment, ensuring it meets the demanding requirements of ISP environments.

1off.it Technical Presentation

Thank you
Q&A

Alessandro Campanella

a.campanella@1off.it

@alessandrocampanella

MikroTik professionals conference

1off