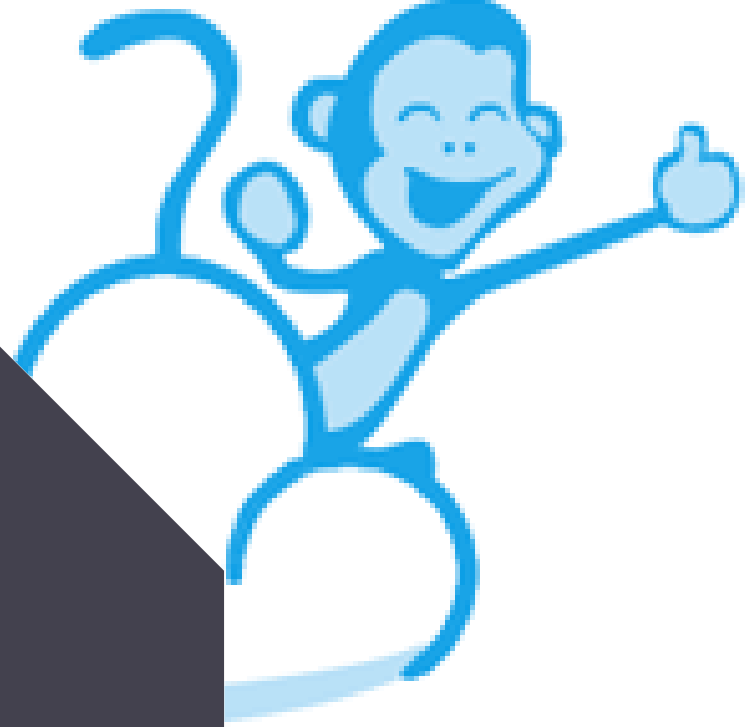
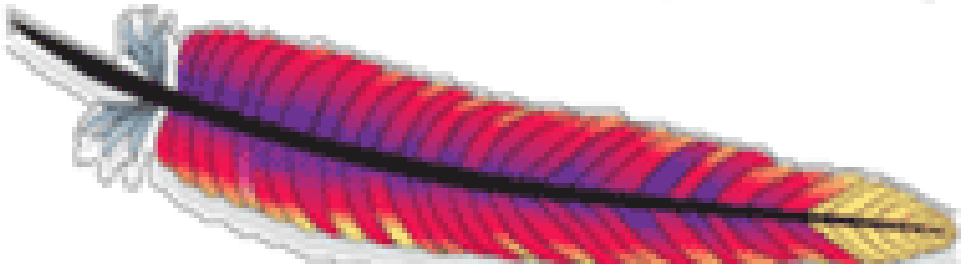
The important of an Out-of-Band Network with Apache CloudStack



udstack

ource cloud computing



Your. Online

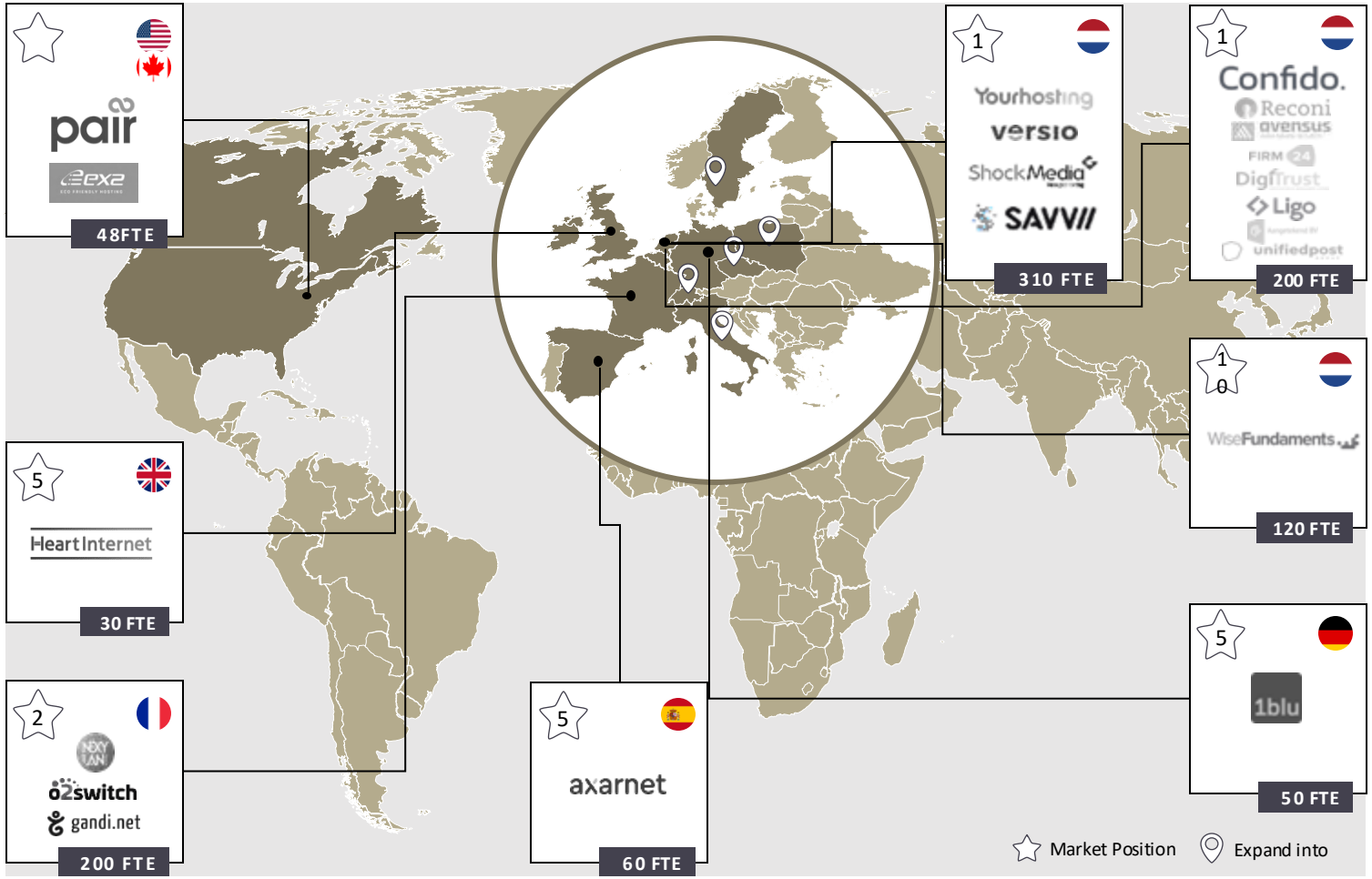
Public

Who am I?

- Wido den Hollander (1986)
- Born and live in the Netherlands
- Married, two sons (2020 and 2022)
- CTO @ Your.Online
  - Started my own hosting company in 2003
  - Techie in my heart
- Open Source & Tech
  - Apache CloudStack developer and PMC member
  - Ceph evangelist
  - IPv6 *fanatic*

Who is Your.Online?

Your.Online is a team of pioneers from all over the world united by the passion of helping businesses succeed online. Our teams of local experts provide highly standardized managed services to high-intent customers to reach their full online potential. We cherish our successful track record in acquiring, developing, and empowering strong local brands to lead their markets



Apache CloudStack @ Your.Online

- We run two large Apache CloudStack deployments
  - Yourhosting in the Netherlands
  - Axarnet in Spain
  - More deployments coming in 2025!

We **love** CloudStack!

What is an Out-of-Band network?

An **Out-of-Band (OOB) network** is a *separate* network used to manage and monitor IT infrastructure and devices independently from the primary or “in-band” data network. *OOB* networks provide administrators with secure, dedicated access to devices *even if the primary network is down* or under stress, which is essential for troubleshooting, configuration, and recovery in critical situations.

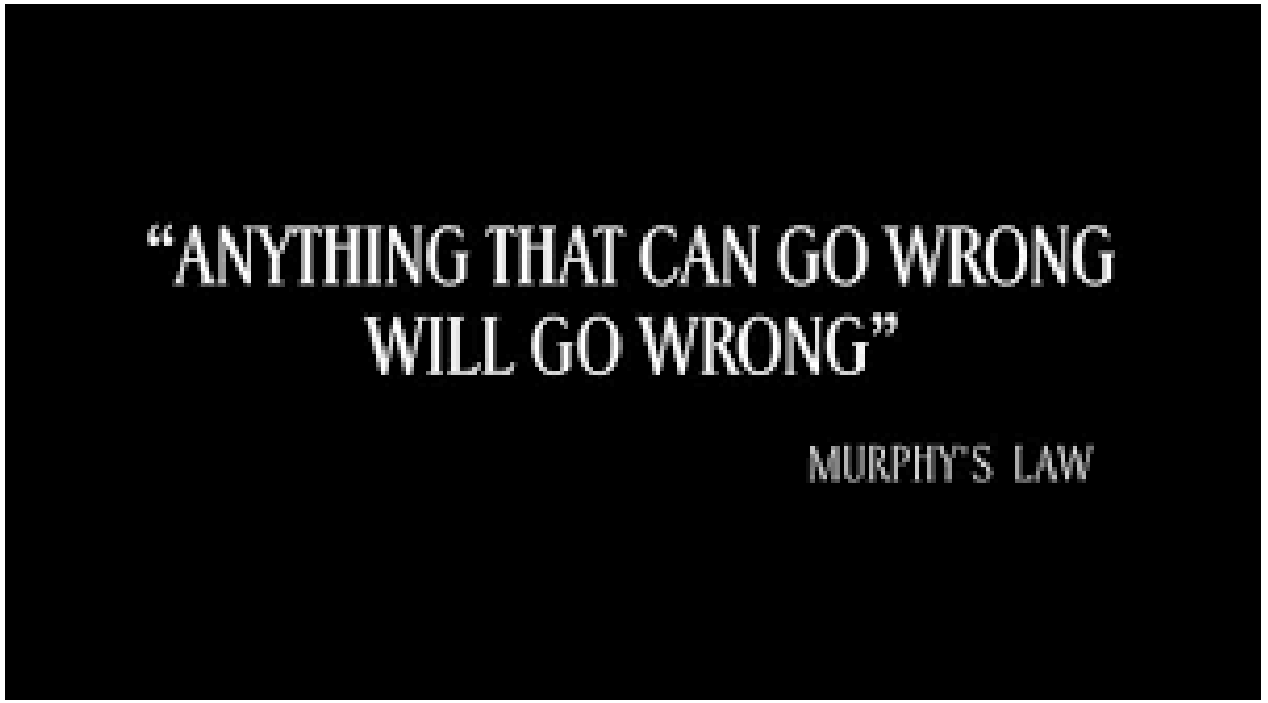
Source: ChatGPT

Why should you want an OOB network?

Because there will be a moment where your primairy network is **down** or under a *lot of stress*

You will not be the first, nor the last to take a trip to the datacenter due to a network outage

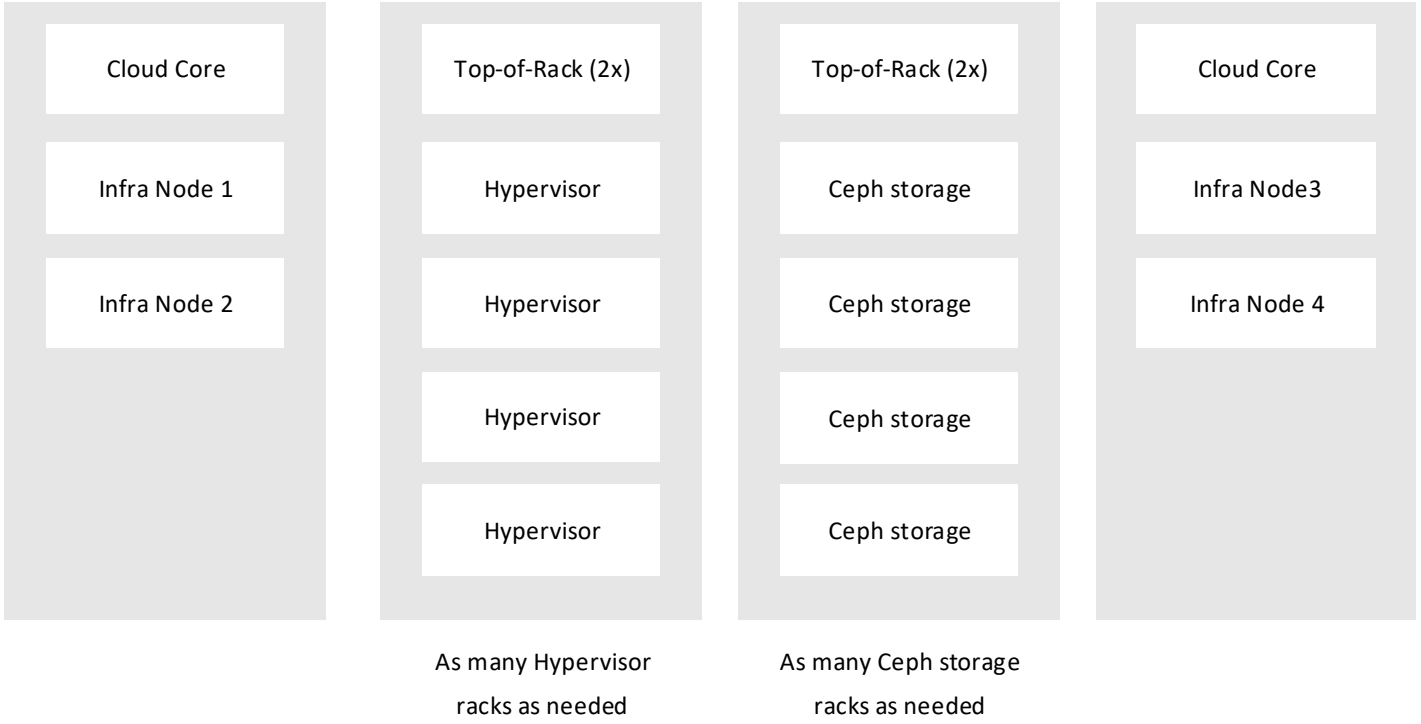
Why should you want an OOB network?



Datacenter layout

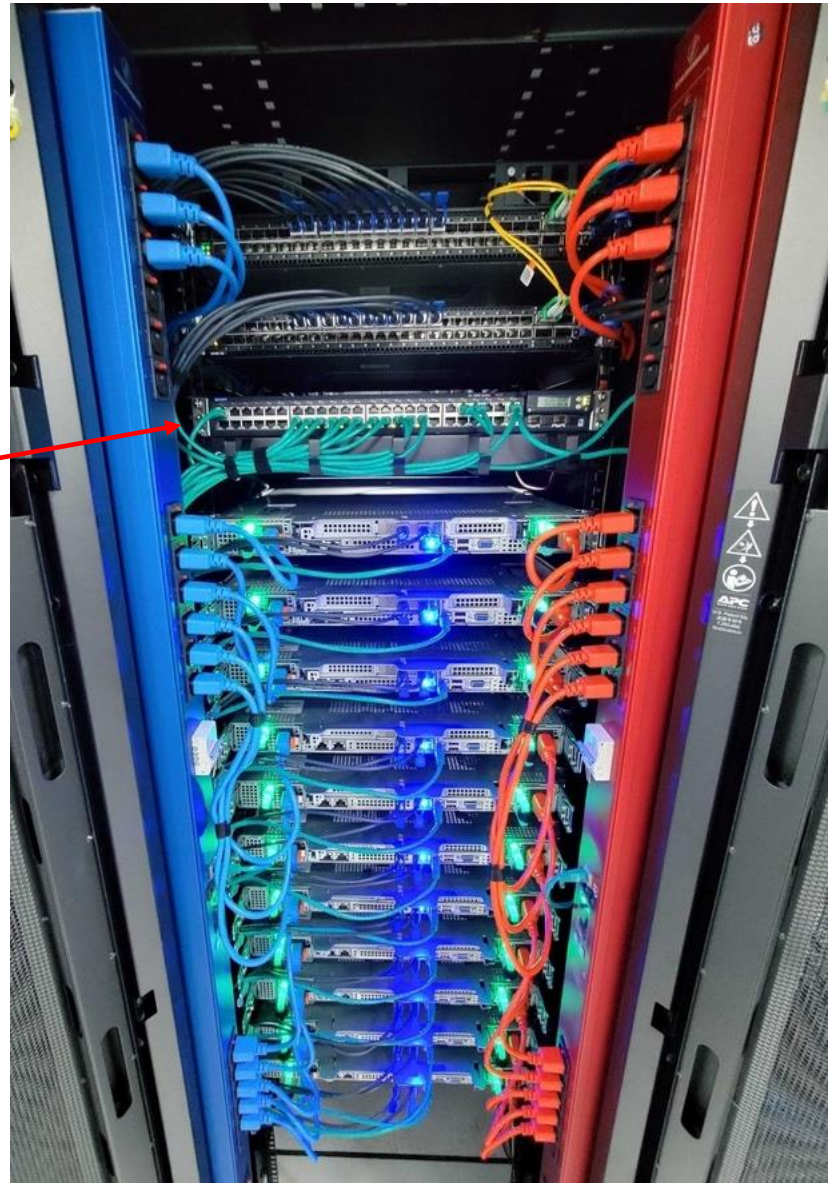
Datacenter layout

- Each rack has a dedicated role
  - Network aggegration
    - OR
  - Hypervisors
    - OR
  - Ceph storage
    - OR
  - TrueNAS storage



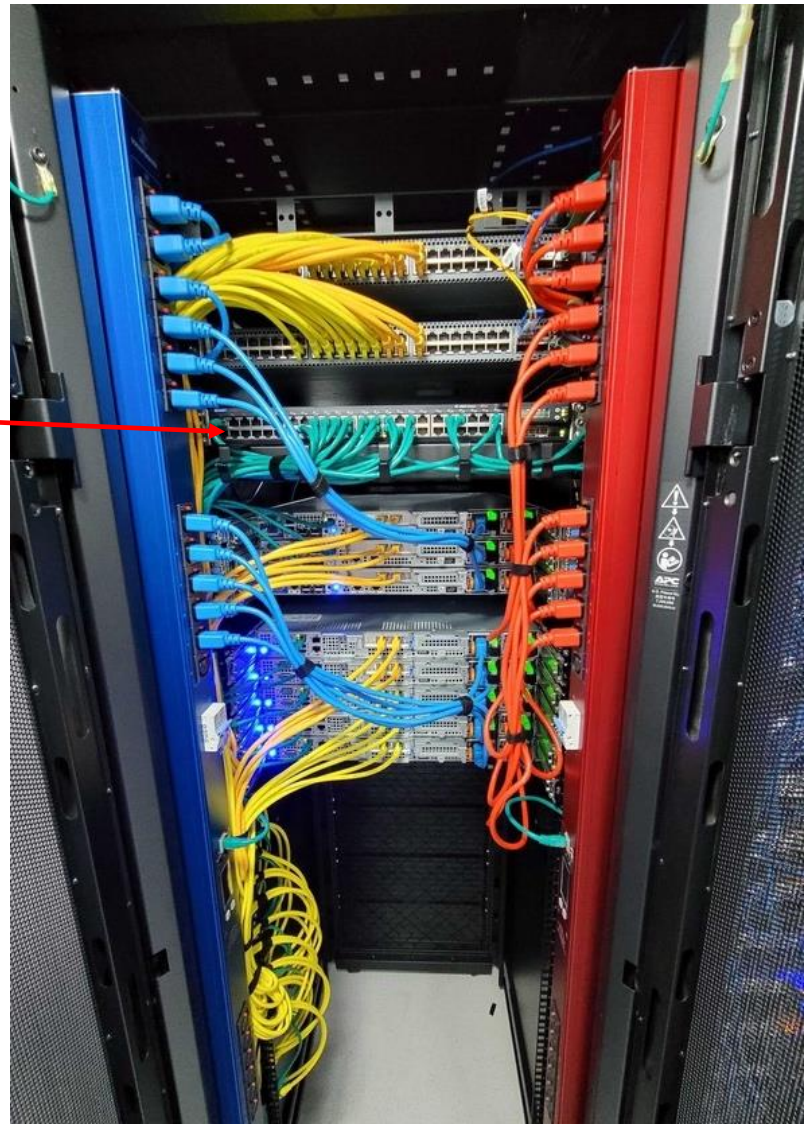
Hypervisor racks

- Contains only **one type** of hypervisors
  - Allows for pre-cabling
  - Keeps them tidy
- Two Top-of-Rack routers
  - 2x100Gb uplink
- Out-of-Band switch
- Add racks when;
  - Additional capacity is required
  - New type of hypervisor is needed
- When replacing hardware after X years, install a new rack and migrate VM workload
  - Discard old rack when no longer in use



Ceph rack

- Contains a single Ceph cluster
- Two top-of-rack
- Out-of-band device
- 3x Ceph Monitor
- As much Ceph OSD machines as fit
  - Physically
    - Most racks are 44~46U in height
  - Power limits
    - Often 32A/5kW



Naming convention

- Hostname of the machine contains multiple elements
  - Role
  - Room
  - Rack
  - Unit
  - Datacenter
- hv-138-a05-18.ams06.cldin.net
  - Role: HV = Hypervisor
  - Room: 138
  - Rack: A05
  - Unit: 18
  - Datacenter: AMS06 (Amsterdam-06)

U	Hostname
46	patchpanel
45	
44	tor-138-a05-46.ams06.cldin.net
43	
42	tor-138-a05-44.ams06.cldin.net
40	
39	**oob-138-a05-39.ams06.cldin.net**
38	
37	hv-138-a05-37.ams06.cldin.net
36	
35	hv-138-a05-35.ams06.cldin.net
34	
33	hv-138-a05-33.ams06.cldin.net
32	
31	hv-138-a05-31.ams06.cldin.net
30	

Patchpanel to route incoming fiber cables through.

This allows for proper cable management

1U spacing between switches for proper cable routing

Each rack has a dedicated out-of-band switch

Racks are power constrained, not space. 1U spacing for proper airflow and easy cabling

Network connections

- Top-of-Rack and Out-of-Band devices need **at least 48 ports**
- The Ceph Monitor on unit 37 is connected to port 37 on all three switches
  - ToR 46
  - ToR 44
  - **OOB**
- This makes cabling and connections predictable and scalable
  - Less documentation required by following this convention

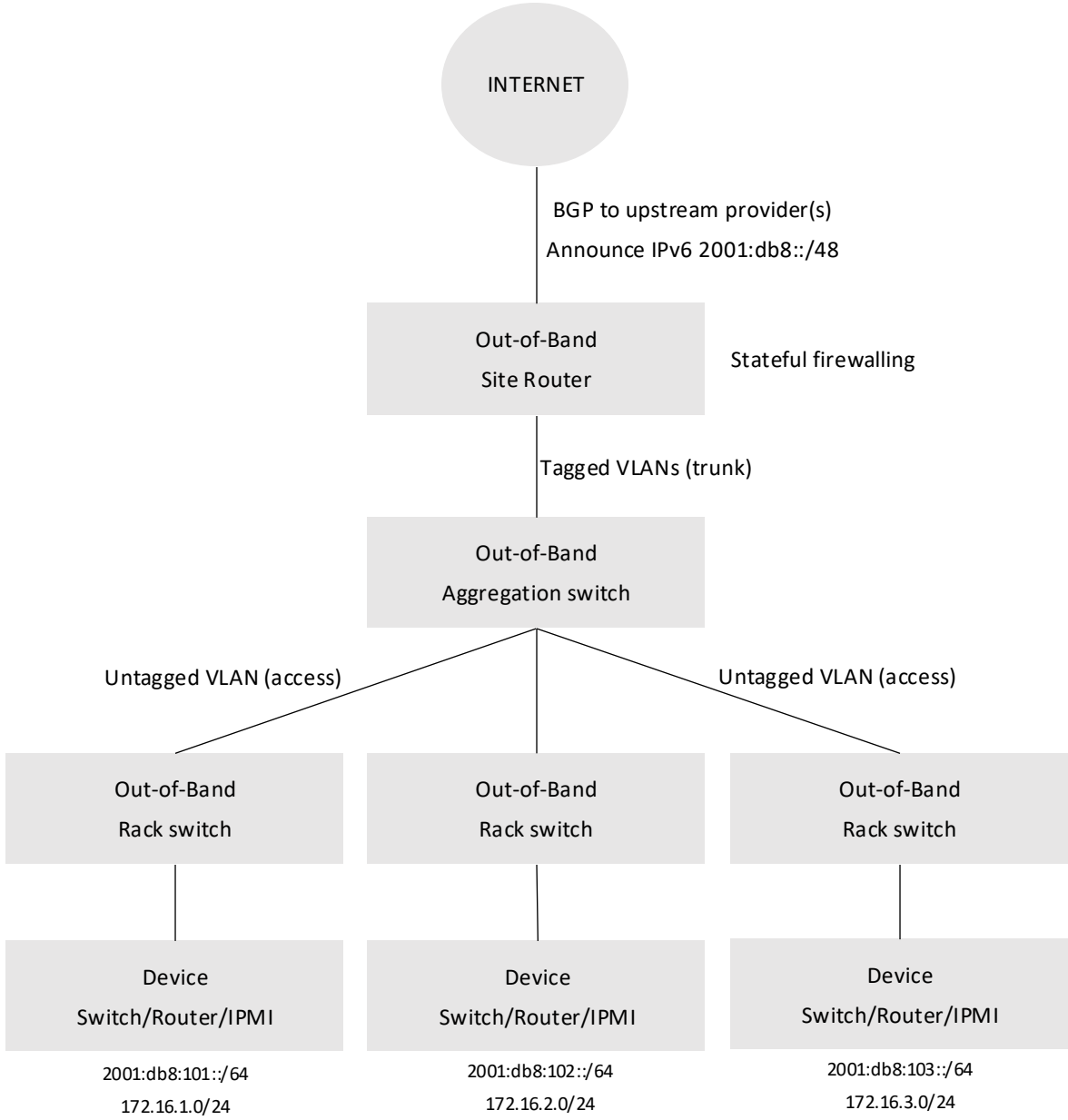
U	Hostname
46	patchpanel
45	
44	tor-138-a05-46.ams06.cldin.net
43	
42	tor-138-a05-44.ams06.cldin.net
40	
39	**oob-138-a05-39.ams06.cldin.net**
38	
37	mon-138-a15-37.ams06.cldin.net
36	mon-138-a15-36.ams06.cldin.net
35	mon-138-a15-35.ams06.cldin.net
34	
33	osd-138-a15-33.ams06.cldin.net
32	osd-138-a15-32.ams06.cldin.net
31	osd-138-a15-31.ams06.cldin.net
30	osd-138-a15-30.ams06.cldin.net

This machine would be connected to port 31 on all three switches. ToR and OOB

Out-of-Band

Out-of-Band

- A proper Out-of-Band network is present to manage devices via a **completely separate** network
  - Switches
  - Routers
  - IPMI/iDRAC
- Switches/Routers have a management port which can be configured in a mgmt VRF
- IPv6-first
  - **/48** subnet for datacenter location
  - **/64** per rack
  - Each rack is it's own VLAN
  - /24 IPv4 with DHCP in each VLAN (fallback)
  - *Stateful firewalling* on **Site Router**
    - MikroTik device



VLAN and subnet per rack

- One VLAN per rack
  - IPv6 /64 subnet (public address space)
    - Part of the larger /48 announced to the internet
  - IPv4 /24 subnet (RFC1918)
- Based on the IP-address you can immediately determine in which **rack** a device is located
- No need for spanning Layer 2 over multiple racks
  - No (R)STP needed

```
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] > /interface print where name=OOB-04-03
Flags: D - dynamic, X - disabled, R - running, S - slave
 #     NAME                                TYPE       ACTUAL-MTU L2MTU  MAX-L2MTU MAC-ADDRESS
 0  R  OOB-04-03                           vlan             1500  1576            64:D1:54:E2:60:6B
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] > /ipv6 address print where interface=OOB-04-03
Flags: X - disabled, I - invalid, D - dynamic, G - global, L - link-local
 #    ADDRESS                                     FROM-POOL INTERFACE
 0  G 2001:5ad0:0:2::1/64                                  OOB-04-03
 1 DL fe80::66d1:54ff:fee2:606b/64                        OOB-04-03
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] > /ip address print where interface=OOB-04-03
Flags: X - disabled, I - invalid, D - dynamic
 #   ADDRESS            NETWORK         INTERFACE
 0   ;;; LAN - RACK-04-03
     172.17.130.1/24    172.17.130.0    OOB-04-03
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] >
```

```
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] > /routing bgp advertisements print peer=AS31577
PEER      PREFIX              NEXTHOP          AS-PATH
AS31577   2001:5ad0::/48      2a00:8a84:97f...
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] >
```

VLAN and subnet per rack

This slide shows a real-life example of one of our active sites in Spain.

All racks are equiped with Out-of-Band as described in this document.

Item	Value
Location	VLC03, Valencia, Spain
IPv6 Aggregate	2001:5ad0::/48
IPv4 Local	172.17.128.0/19
Rooms	2
Racks	22

Rack	VLAN	IPv6	IPv4
01	3001	2001:5ad0::/64	172.17.128.0/24
02	3002	2001:5ad0:0:1::/64	172.17.129.0/24
03	3003	2001:5ad0:0:2::/64	172.17.130.0/24
04	3004	2001:5ad0:0:3::/64	172.17.131.0/24
05	3005	2001:5ad0:0:4::/64	172.17.132.0/24
06	3006	2001:5ad0:0:5::/64	172.17.133.0/24
07	3007	2001:5ad0:0:6::/64	172.17.134.0/24

IPv6 address determination

- Site Router will sent IPv6 Router Advertisements in each VLAN
  - Contains /64 prefix to be used in that network
- Using StateLess Address Auto Configuration (SLAAC, RFC4862) devices ***can*** obtain an address based on the prefix + their MAC address
  - IPMI/iDRAC will do this by default
- MAC address of a device can be obtained from the MAC-address table of Out-of-Band switch in rack
  - Remember that all devices are connected to the port number which *corresponds with their rack Unit*.

```
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] > /ipv6 neighbor print where mac-address=10:7D:1A:FD:33:2A
Flags: R - router
 0   address=2001:5ad0:0:9:127d:1aff:fefd:332a interface=OOB-04-10 mac-address=10:7D:1A:FD:33:2A
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] >
```

Prefix = 2001:5ad0:0:9::/64

MAC = 10:7D:1A:FD:33:2A

Address = 2001:5ad0:0:9:127d:1aff:fefd:332a (SLAAC, **RFC4862**)

IPv6 address rack's OOB switch

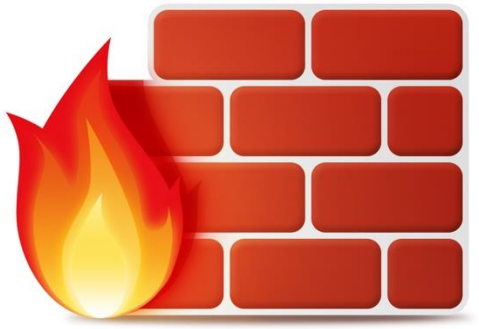
- Single VLAN configured on the switch
- OOB switch has a static IPv6 addres
- Last digit matches the Unit number in the rack

Juniper EX4200 OOB switch

```
wdh@oob-138-c13-41> show configuration interfaces vlan
unit 0 {
   family inet6 {
      address 2001:5ad0:141:38::41/64;
   }
}


{master:0}
wdh@oob-138-c13-41> show configuration vlans
default {
   l3-interface vlan.0;
}

{master:0}
wdh@oob-138-c13-41> show vlans
Name         Tag     Interfaces
default
                ge-0/0/0.0*, ge-0/0/1.0, ge-0/0/2.0, …, …


{master:0}
wdh@oob-138-c13-41>
```

Firewalling

- The Site Router performs Statefull firewalling
- TCP connections are only allowed from trusted sources
  - Private VPN servers
  - Monitoring servers
  - Management servers
- ICMP is essential for IPv6 and allowed, rest is filtered
- Management is done via SSH and HTTPS and thus encrypted



```
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] > /ipv6 firewall filter export
/ipv6 firewall filter
add action=accept chain=input comment="MikroTik own connections" connection-state=established,related,untracked
add action=accept chain=input protocol=icmpv6
add action=accept chain=input port=179 protocol=tcp src-address-list=BGP
add action=accept chain=input port=3784 protocol=udp src-address-list=BGP
add action=accept chain=input dst-port=22,443,1194 protocol=tcp src-address-list=management
add action=accept chain=input dst-port=22 protocol=tcp src-address-list=monitoring
add action=accept chain=input dst-port=161 protocol=udp src-address-list=monitoring
add action=accept chain=input connection-state=established,related src-address-list=local-ipv6
add action=reject chain=input log=yes log-prefix="ipv6 input drop: " reject-with=icmp-admin-prohibited
add action=accept chain=forward protocol=icmpv6
add action=accept chain=forward dst-address-list=local-ipv6 src-address-list=management
add action=accept chain=forward dst-address-list=local-ipv6 port=22 protocol=tcp src-address-list=monitoring
add action=accept chain=forward dst-address-list=local-ipv6 port=161 protocol=udp src-address-list=monitoring
add action=accept chain=forward dst-address-list=local-ipv6 protocol=tcp src-address-list=monitoring src-port=10051
add action=accept chain=forward src-address-list=local-ipv6
add action=accept chain=forward connection-state=established,related dst-address-list=local-ipv6
add action=accept chain=forward dst-address-list=!local-ipv6 dst-port=53 protocol=udp src-address-list=local-ipv6
add action=accept chain=forward dst-address-list=!local-ipv6 dst-port=53 protocol=tcp src-address-list=local-ipv6
add action=accept chain=forward dst-address-list=local-ipv6 protocol=udp src-port=53
add action=accept chain=forward dst-address-list=local-ipv6 protocol=tcp src-port=53
add action=accept chain=forward dst-address-list=local-ipv6 dst-port=123 packet-size=0-512 protocol=udp src-port=123
add action=accept chain=forward dst-address-list=local-ipv6 dst-port=80,443 protocol=tcp src-address-list=CloudStack
add action=reject chain=forward log=yes log-prefix="ipv6 forward drop: " reject-with=icmp-admin-prohibited
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] >
```

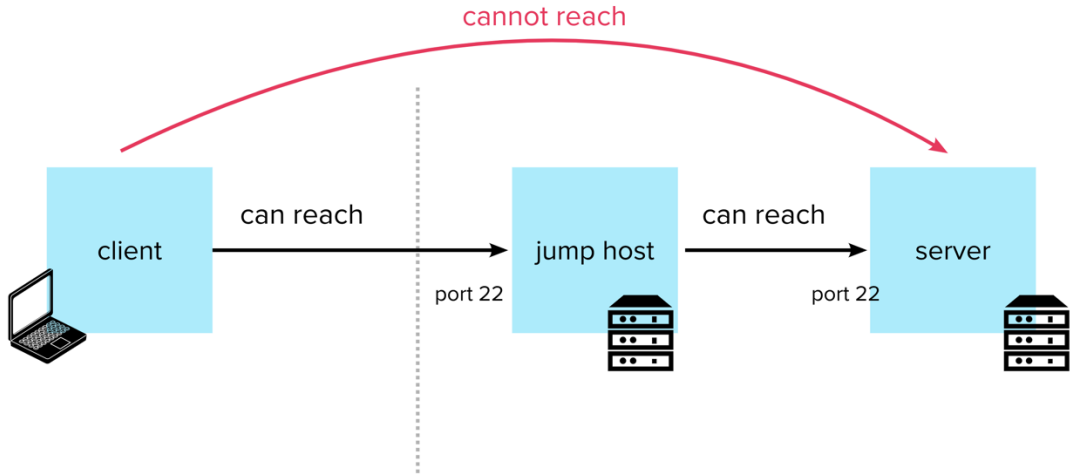
```
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] > /ipv6 firewall address-list export
/ipv6 firewall address-list
add address=2001:5ad0::/48 list=local-ipv6
add address=2001:db8:100::/64 comment=monitoring.cldin.net list=monitoring
add address=2a0c:8e90:400::/40 comment="CLDIN Trusted" list=management
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] >
```

/40 management CIDR

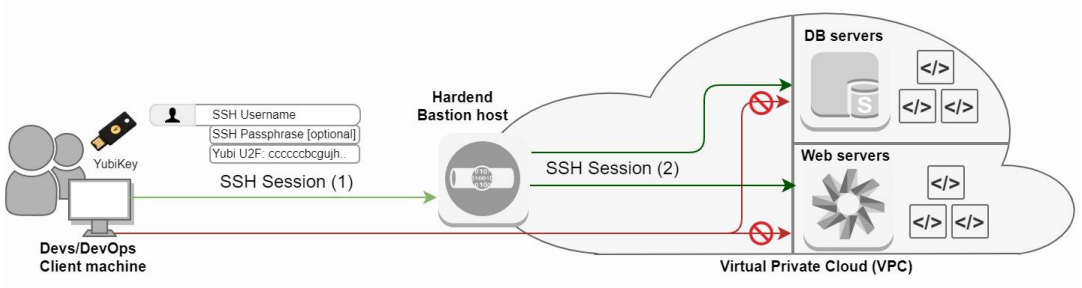
- We have allocated a /40 subnet for our management
  - 2a0c:8e90:400::/40
- From this /40 we announce separate /48 subnets
  - A /48 subnet is the smallest you can announce for IPv6 through BGP on the internet
  - We only have to whitelist a single /40 in our firewalls
  - 256 /48 subnets fit into a /40 (2^8=256)
- /48 subnets
  1. Our VPN
  2. SSH Jump (bastion) Host (OpenSSH proxyjump, option –J)
  3. Backup VPN
  4. Backup SSH Jump Host
- Our jump hosts only allow SSH keys signed by a Yubikey
  - https://cryptsus.com/blog/how-to-configure-openssh-with-yubikey-security-keys-u2f-otp-authentication-ed25519-sk-ecdsa-sk-on-ubuntu-18.04.html

```
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] > /ipv6 firewall address-list export
/ipv6 firewall address-list
add address=2001:5ad0::/48 list=local-ipv6
add address=2001:db8:100::/64 comment=monitoring.cldin.net list=monitoring
add address=2a0c:8e90:400::/40 comment="CLDIN Trusted" list=management
[wdh@oob-rtr-04-01-38.vlc03.cldin.net] >
```

ssh -J jump.domain.tld wido@router01.mycloud.tld

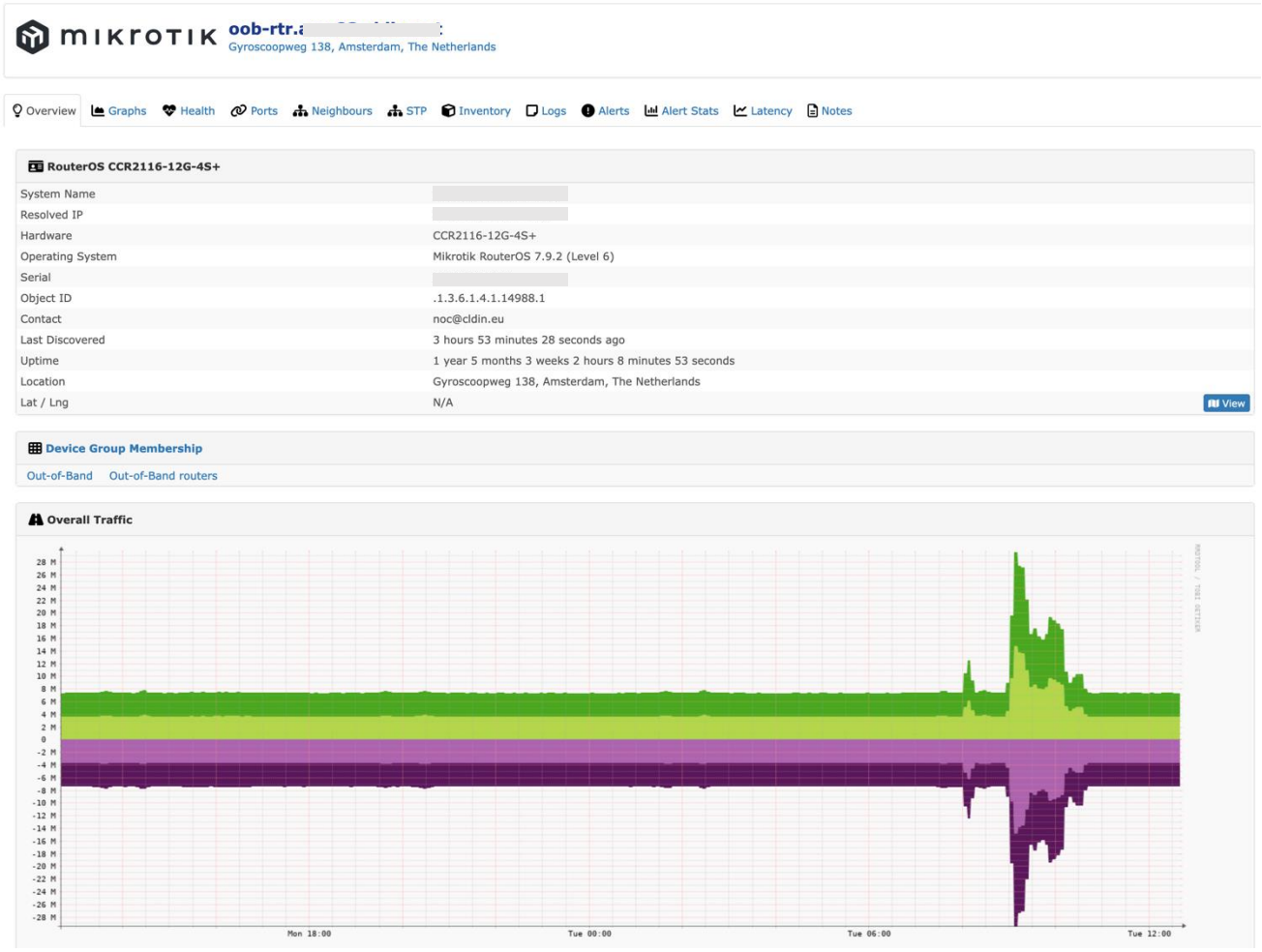


Source: oooops.dev



Daily use

- Out-of-Band is used for **daily operations**
  - SNMP monitoring of Routers/Switches
    - We use LibreNMS from a *remote* location
  - SSH for management
- Devices can download firmware updates via OOB
- By using the OOB for daily operations people know that **it works**
- The OOB is a fully functional L3 network
  - Completely separate from primary L3 network
  - Low on bandwidth, but very important
  - In case of DDoS attack all devices can still be reached

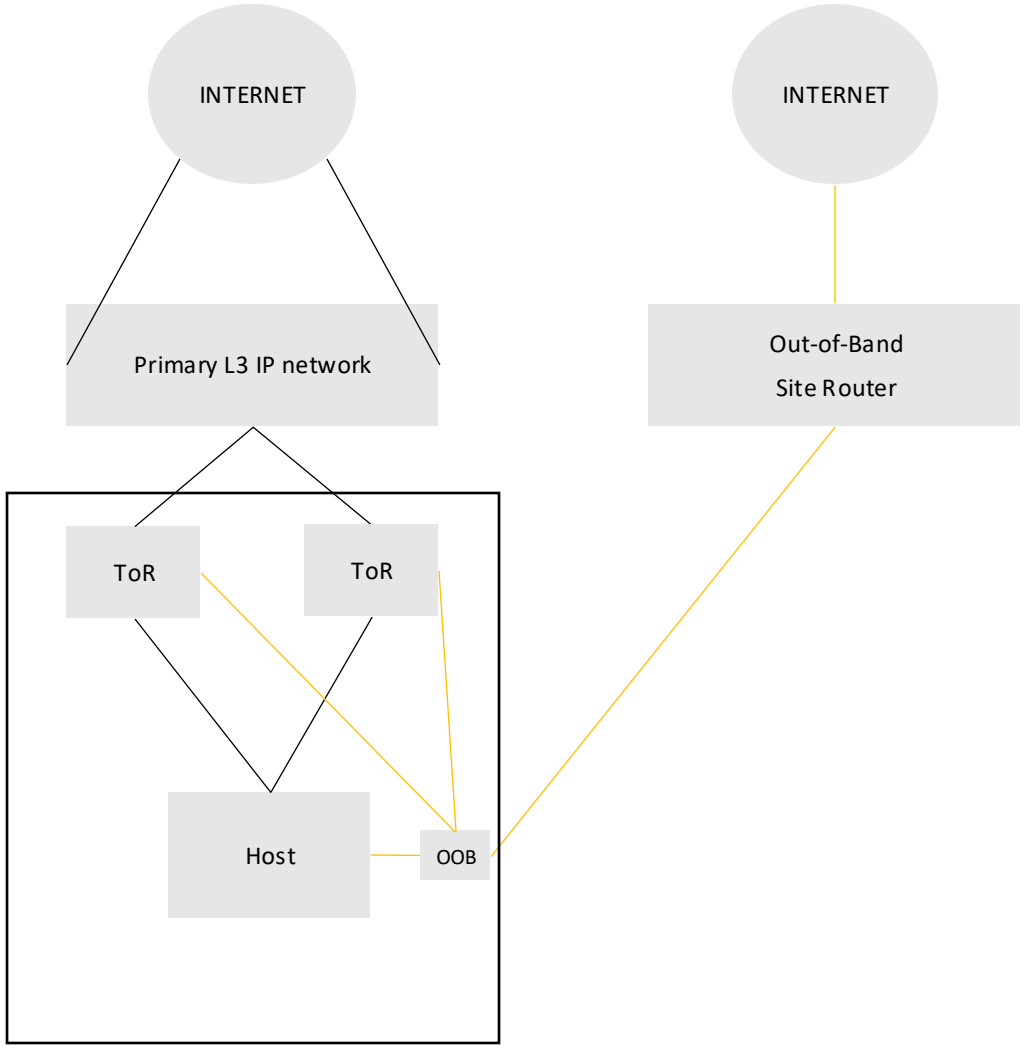


I don’t have IPv6 at home/office/mobile/whatever

- Use/build a VPN!
  - Just as we do
- We have build a VPN solution ourselves
  - Based on OpenVPN
  - Pritunl software
  - We route a /64 to the OpenVPN machine
  - Each client gets a /128 IPv6 address assigned
- This VPN does NOT terminate on the MikroTik routers on site
  - It’s just somewhere on the internet
  - All our management is encrypted via SSH
  - The MikroTik firewalls allow you to connect to the OOB devices when you originate from the VPN

It’s an additional network!

- Effectively you have two fully functional L3 networks!
- The OOB network is less redundant
  - But not unreliable!
- No NAT, a true routing L3 network



OOB with CloudStack

IPMI, iLO, iDRAC and Redfish

- **Intelligent Platform Management Interface (IPMI)** is a standardized interface used for managing computer systems and monitoring their operation, primarily in data centers and enterprise environments. Developed by Intel, IPMI provides out-of-band management capabilities, meaning it allows administrators to remotely monitor, manage, and troubleshoot systems independently of the operating system or even if the system is powered off.
- IPMI typically uses **UDP port 623** for communication

- **Redfish** is a modern, RESTful API standard designed for managing and monitoring data center hardware, such as servers, storage systems, and networking equipment. Developed by the Distributed Management Task Force (DMTF), Redfish addresses some of the limitations of older management protocols like IPMI by providing a more secure, flexible, and extensible interface that's also easier to use.
- Redfish communicates over a network using **HTTP(S)** as its transport protocol.

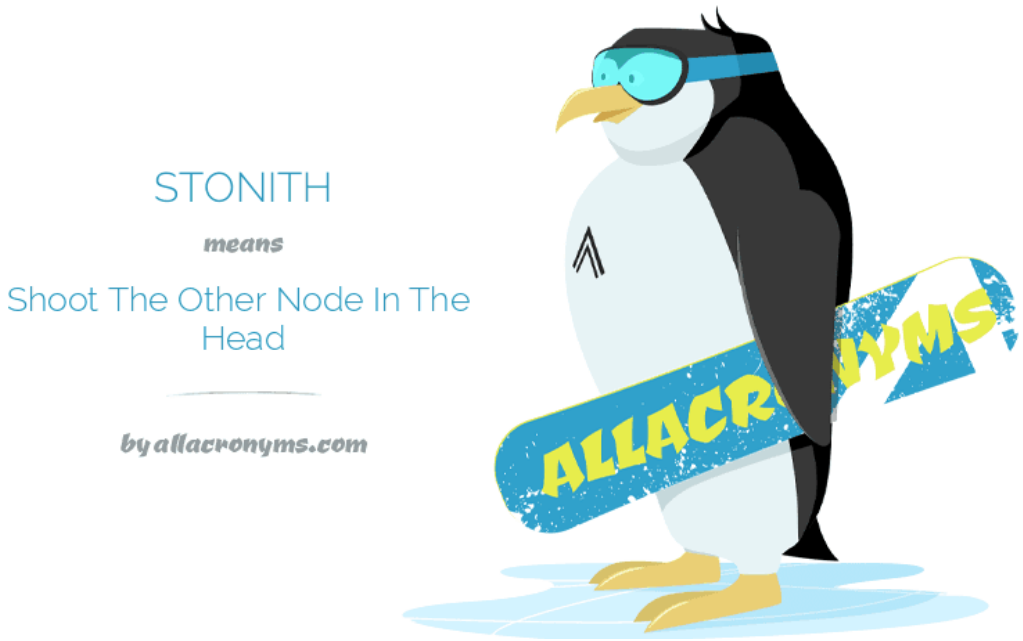
Vendor	BMC	IPMI?	Redfish?
SuperMicro	IPMI	Yes	Yes
Dell	iDRAC	Yes	Yes
HPE	iLO	Yes	Yes

Vendors have different names for their Baseboard Management Controller (BMC).

Host fencing through Out-of-Band

- How do you reliably determine of a node is dead?
  - No more Agent connection?
  - Not responding to ping?
  - What if only WAN is down, and it still has running VMs?
  - Storage locking?

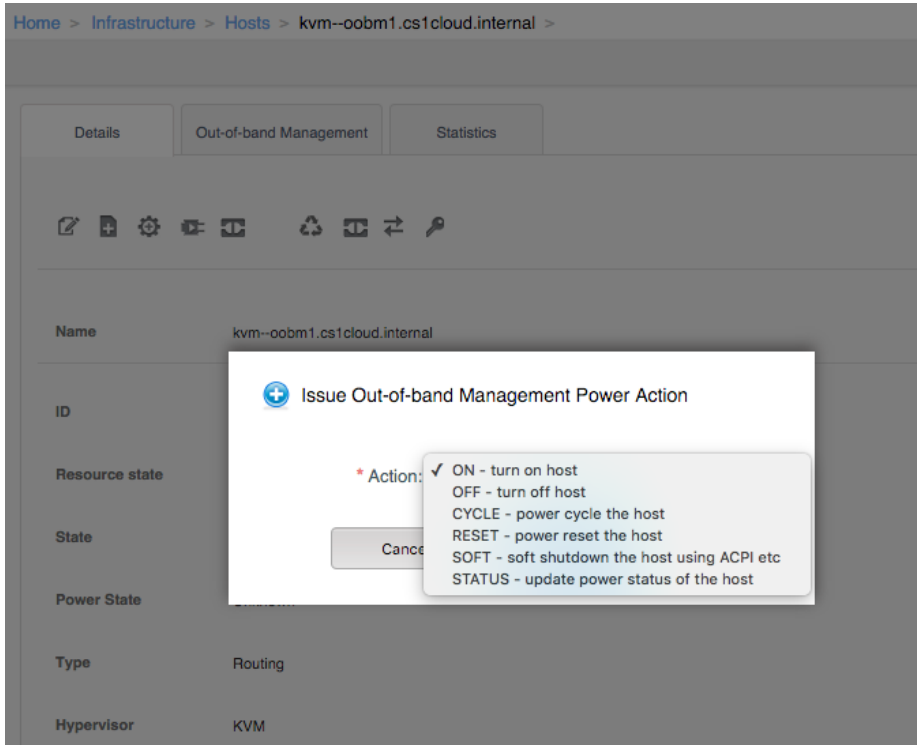
- STONITH = **S**hoo**T** the **O**ther **N**ode **I**n **T**he **H**ead
  - Use the BMC to Power Cycle the Host which is assumed *down*

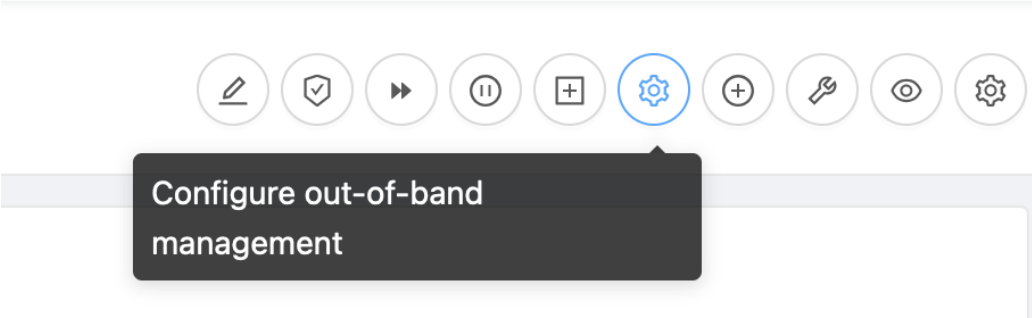


KVM HA via OOB Fencing

- Per host Out-of-Band details need to be enabled and configured
- Choose IPMI or Redfish protocol
- CloudStack Management Server will now start to poll the “Power Status” of each node
- Management Server now has the option to RESET and CYCLE a Host

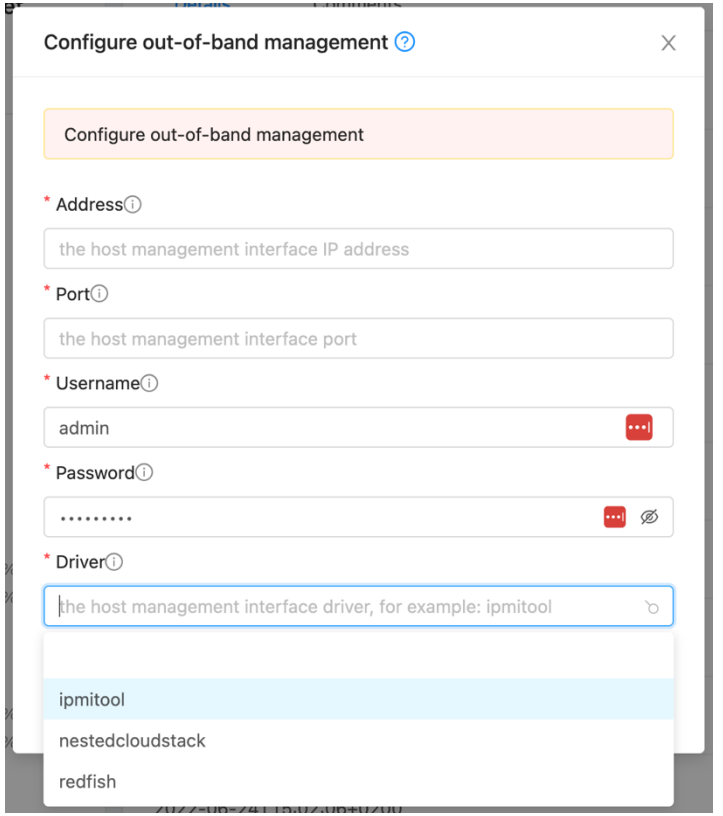
- HA is more reliable by using the Host’s BMC to Fence the Host

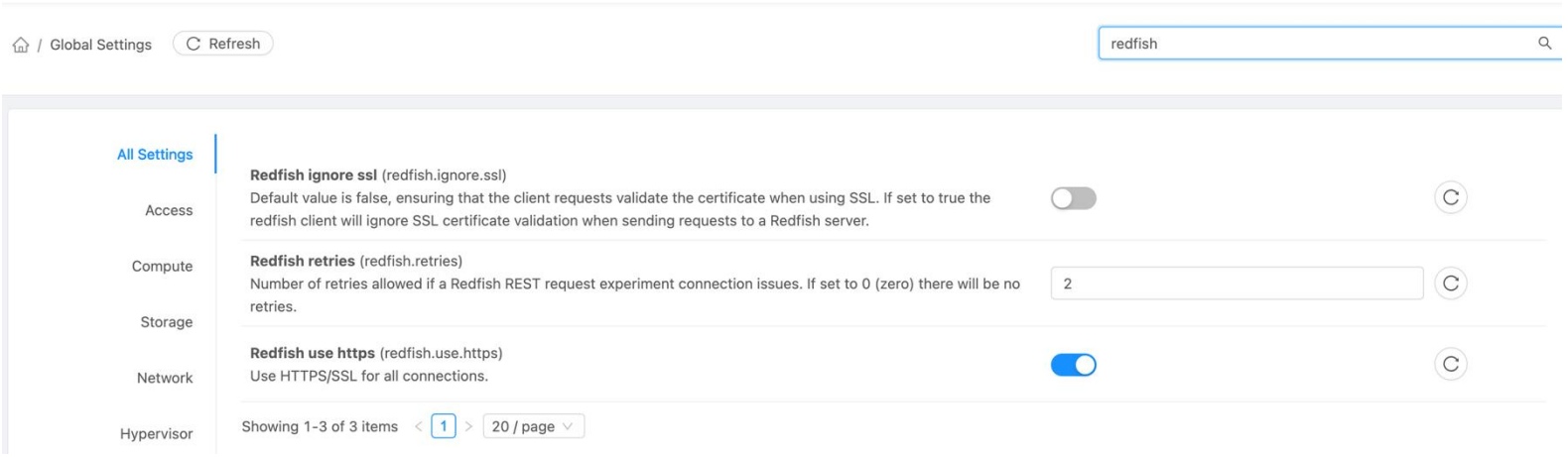




KVM HA via OOB Fencing

- Make sure the Management Server is allowed to connect to the OOB network where the BMC of the host is
- Choose your driver
  - IPMI (UDP port 623)
  - Redfish (HTTP or HTTPS)





Summary

- A good OOB network is more than JUST a backup
- Use your OOB network for all your daily tasks
  - SNMP, SSH, etc
- Treat your OOB network is a first-class network
- Use IPv6 to directly access the devices
- Firewalls prevent unauthorized traffic
- OOB support in CloudStack can reliably Fence Hosts

@widodh
wido@denhollander.io
blog.widodh.nl

More information

- https://blog.apnic.net/2024/11/12/out-of-band-network-design-for-service-provider-networks/
- https://www.daryllswer.com/out-of-band-network-design-for-service-provider-networks/


Your.
Online