

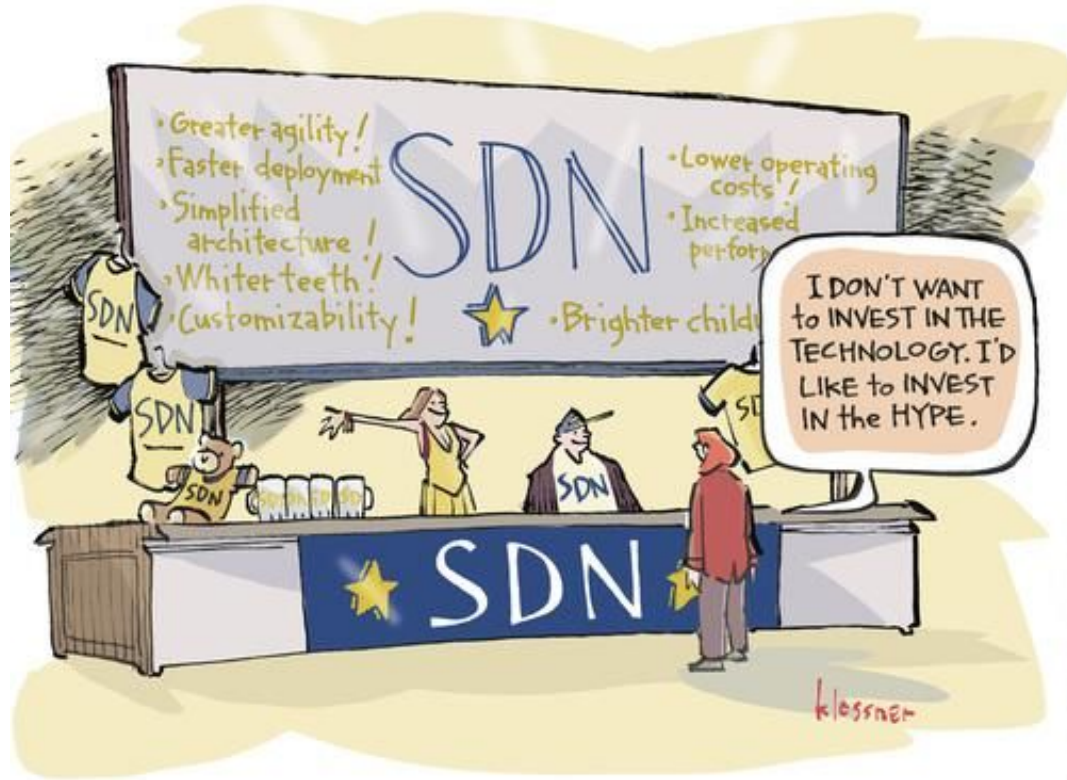
Whitebox Switches Deployment Experience

[Engineer's Perspective]

Aftab Siddiqui

eintellego Networks

I'm only going to talk about reality..No Hype....



Introduction

- What is Open Network or Whitebox in Networking?
 - Whitebox = Bare-metal = Britebox = Open Network
- But still, what is it?
 - Hardware with merchant silicon without software. Same as servers.
- Another cheap and low performance switch range in the market?
 - Yes, they are comparatively cheaper but definitely not low performance
 - All switches are non-blocking
- Who are the vendors?
 - There are many
 - HP, Edgecore/Accton, Dell, Quanta
- What about software to run these switches?
 - Cumulus Linux, PicaOS etc

Why?

- Many options to choose from
- Multiple Operating Systems to Bring Out the Best of the Same Box
- Multiple Boxes to Bring out the best from Same Operating System
- Select the Software for your Needs Instead of Replacing Systems
- Select the hardware for your Needs without replacing the Operating System
- End to End Common Hardware/Software
- Unlock New Capabilities of Network Hardware
- Drive Down the Economics with Standardized Hardware
- Capture Fast Moving Merchant Silicon Innovation
- Simplify Support, Sparing, Logistics and Re-use

Under the hood

Operating System:



Chassis:



Operating System:



Any Network
Operating System

Open Standard
Hardware

Merchant Silicon

Hardware Specs - 10G Switches

Model	Port Density	CPU	RAM	Buffer	ASIC
Dell S4048	48x10G SFP+ 6x40G QSFP+	Intel x86 (C2338)	2GB	12MB	Broadcom Trident II
Accton AS5712	48x10G SFP+ 6x40G QSFP+	Intel x86 (C2538)	8GB	12MB	Broadcom Trident II
Quanta T3048-LY8	48x10G SFP+ 6x40G QSFP+	Intel x86 (C2538)	4GB	12MB	Broadcom Trident II

Note: Specs taken from respective datasheets

Hardware Specs - 40G Switches

Model	Port Density	CPU	RAM	Buffer	ASIC
Dell S6000	32x40G QSFP+	Intel x86 (S1220)	4GB	12MB	Broadcom Trident II
Accton AS6712	32x40G QSFP+	Intel x86 (C2538)	8GB	12MB	Broadcom Trident II
Quanta T5032-LY6	32x40G QSFP+	Intel x86 (C2758)	4GB	12MB	Broadcom Trident II

Note: Specs taken from respective datasheets

Vendor Selection

- Which vendor is providing better switch?
 - All vendors. Specs are almost similar
- Which hardware vendor should I pick?
 - Your choice. This is the whole idea behind whitebox. 'Choice'
- Which Network OS should I select?
 - Your choice. Depends upon the functionality you need
 - Not all of them provide all features
 - PicOS and Cumulus both use debian kernel
 - PicOS provide QinQ, Cumulus doesn't
 - Cumulus provide controller less VXLAN, PicOS doesn't
 - PicOS CLI is IOS like, Cumulus is pure linux

Use Cases

- I'm not Facebook, Google, AWS or any other web scale giant, why should I care?
 - Because you still need Top of the Rack (ToR) switches and/or Access/Aggregation switches in your network
 - It's all about economics, choice, scalability and nothing else
- Will it help me to do something related to SDN?
 - Yes, this is the first step towards something practical in the SDN world but let's not talk about it
 - Start with Software Assisted Networking (SAN)
- Great, where should I deploy?
 - Deploy as ToR in DC or as Access switch or anywhere you like. Start from somewhere

First Experience with Open Network Switches

Target: Establish Connectivity between multiple PoPs.

Media: Dark Fibre

Service Offering: Backhaul (Layer 2)

Budget: Challenging (Only because of this we were forced to look into white box switches)

Project Timelines: Weeks rather months

POC: 4 PoPs

Selection of Hardware

Requirement: 24+ x 10G SFP+ ports and 2+ x 40G QSFP+ Ports

Available Options: Many [48x10G SFP+, 6x40G QSFP+] switches available from different vendors as mentioned on previous slides

Selection Criteria [self imposed]:

- Who can deliver the switches ASAP?
- Who maintains local/regional inventory?
- Who has clueful local SEs?

Hardware Selected: Dell S-4048 ON (purely on above criteria)

Selection of Network OS

Requirement: Switch OS [all features required in a switch OS]

Available Options: Cumulus Linux and Pica8 (PicOS)

Selection Criteria [self imposed]:

- Who can provide demo OS?
- Who has some non-DC deployments closer to our scenario?
- Who has clueful local SEs?

NOS Selected: Cumulus Linux

- Cumulus VX available to try and evaluate for free
- An awesome Techfield Day event available on YouTube to understand the structure
- Local SEs available to share some knowledge

First Impression - Hardware

Looks like a normal switch.



First Impression - Network OS [Cumulus]

Installed Cumulus VX on Virtualbox for testing

```
Linux cumulus 3.2.68-6 #3.2.68-6 SMP Mon Oct 26 15:45:09 PDT 2015 x86_64
Welcome to Cumulus VX (TM)

Cumulus VX (TM) is an open-source LINUX (R) distribution. License files are included with every package
installed in the system and can be viewed in the /usr/share/*/doc/copyright files.

The registered trademark Linux (R) is used pursuant to a sub-license from LMI, the exclusive licensee of
Linus Torvalds, owner of the mark on a world-wide basis.

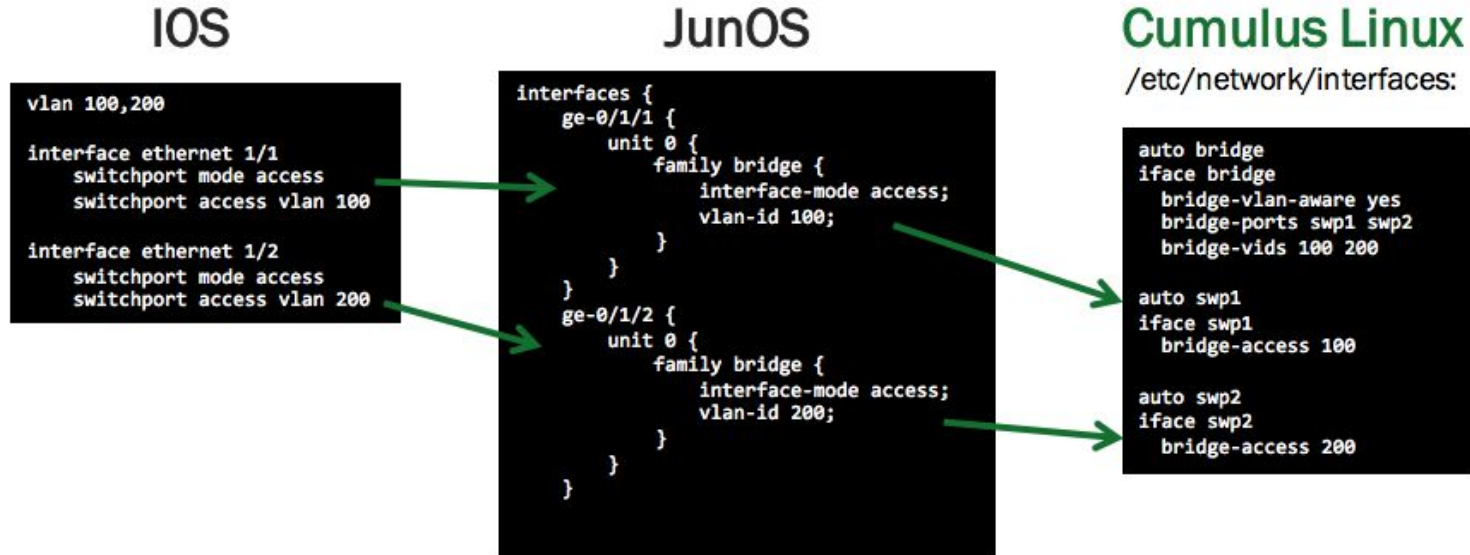
Last login: Thu Jan 21 02:53:11 2016 from 22.0.12.59
cumulus@cumulus$ █
```



Where is the CLI?

First Impression - Network OS [Cumulus]

Thanks to detailed documentation on CumulusLinux website, I got some idea how to kick start the config.



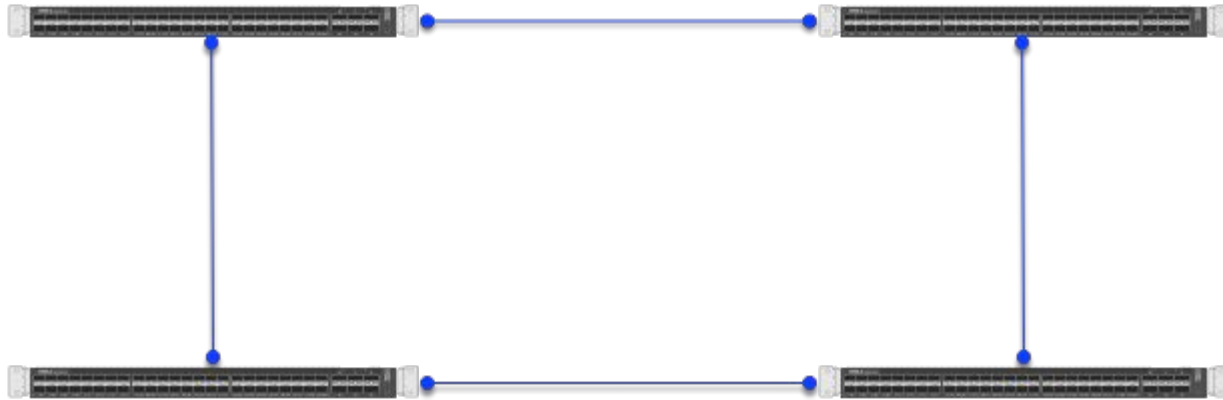
First Impression - Network OS [Cumulus]

Routing: Quagga [Good to see some familiar CLI]

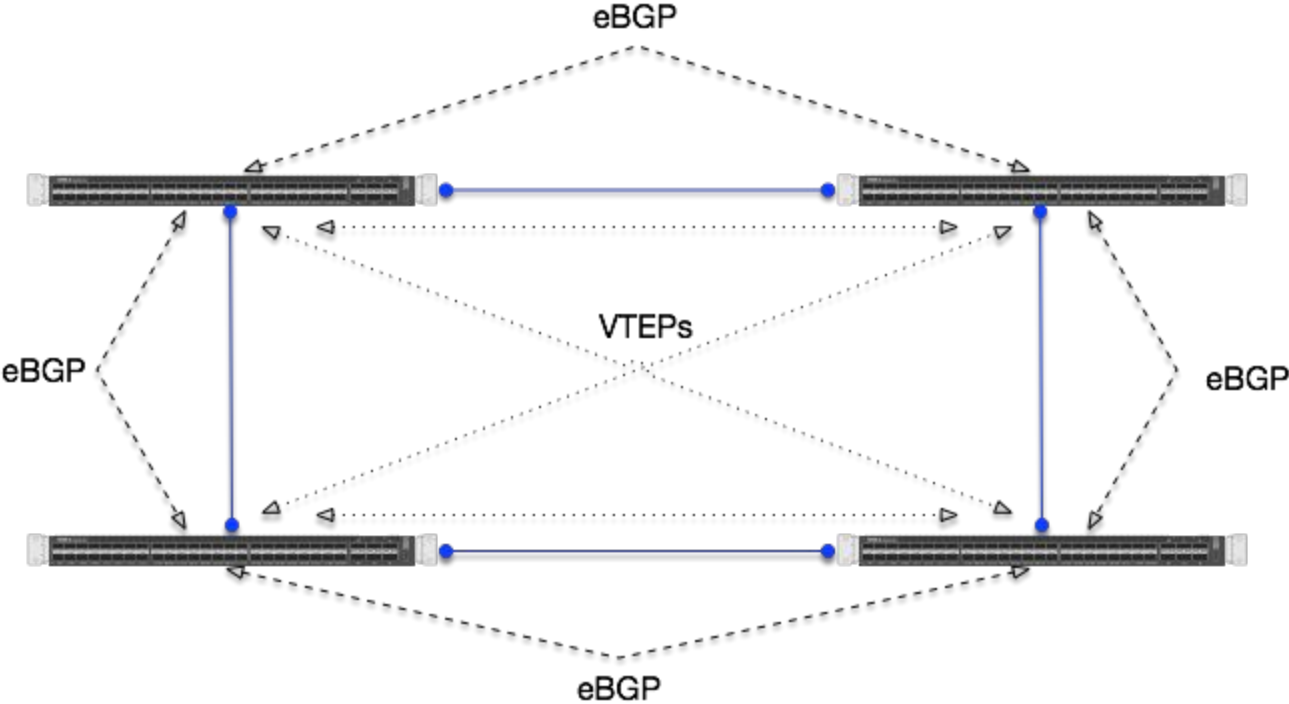
```
[root@cumulus:~# vtysh  
Hello, this is Quagga (version 0.99.23.1).  
Copyright 1996-2005 Kunihiro Ishiguro, et al.  
  
[cumulus# show int desc  
Interface      Status  Protocol  Description  
eth0           up      up  
lo             up      up  
swp1           up      up  
swp2           down    down  
swp3           down    down  
swp4           down    down  
swp5           down    down  
swp6           down    down  
swp7           down    down  
cumulus# █
```


Design

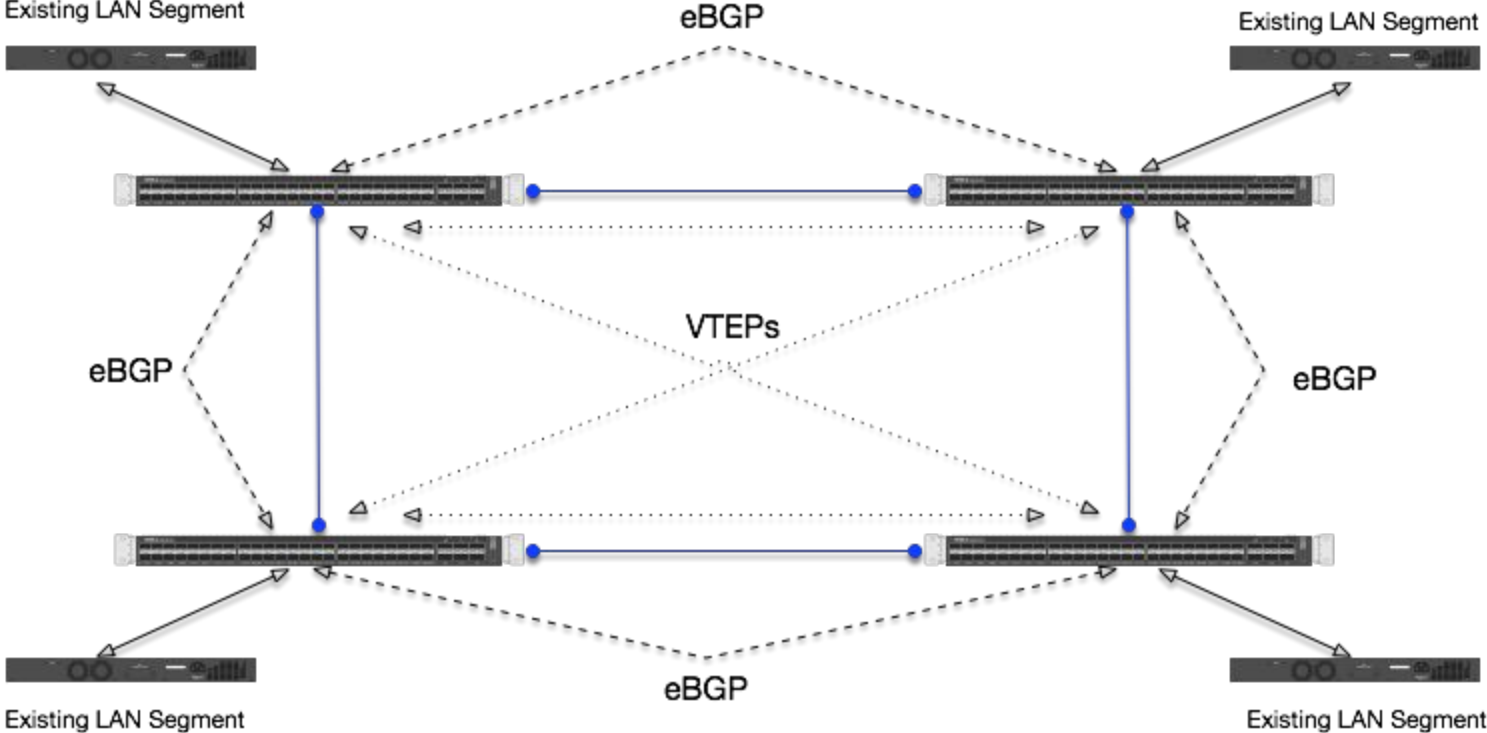
POC - 4 PoPs



Design



Design



Design Summary

Ever growing Layer 2 network

Can't change the existing network design

Achieved the goal by implementing VXLAN

But,

Broke LACP between existing network and 802.1p marking

Troubleshooting nightmare for operations team

```
cumulus@CML-02$ netshow interface
```

```
-----  
To view the legend, rerun "netshow" cmd with the "--legend" option  
-----
```

Name	Speed	MTU	Mode	Summary
UP br-200	N/A	1500	Bridge/L3	IP: 192.168.0.3/24 Tagged Members: swp2.200 802.1q Tag: 200 STP: Disabled
UP br-699	N/A	1500	Bridge/L2	Untagged Members: vni-600003 Tagged Members: swp1.699 802.1q Tag: 699 STP: Disabled
UP eth0	1G	1500	Mgmt	IP: 22.0.12.75/24
UP lo	N/A	16436	Mgmt	IP: 127.0.0.1/8, 2.2.2.2/32, ::1/128
UP swp2	1G(SFP+)	1500	Trunk/L2	Bridge Membership: Tagged: br-200(200)
UP swp48	1G(SFP+)	1500	Interface/L3	IP: 10.0.0.2/30
UP vni-2000	N/A	1500	IntTypeUnknown	
UP vni-600003	N/A	1500	Access/L2	Untagged: br-699

```
cumulus@CML-02$ brctl showmacs br-200
```

port	name	mac	addr	vlan	is	local?	ageing	timer
swp2.200	00:23:9c:19:1c:40	0	no	7.55				
swp2.200	14:18:77:01:5d:02	0	yes	0.00				

```
cumulus@CML-02$ bridge fdb show
```

```
00:23:9c:19:1c:40 dev swp2.200 vlan 0 master br-200  
14:18:77:01:5d:02 dev swp2.200 vlan 0 master br-200 permanent  
00:00:00:00:00:00 dev vni-2000 dst 1.1.1.1 vlan 65535 self permanent  
00:00:00:00:00:00 dev vni-600003 dst 30.30.0.10 vlan 65535 self permanent  
f6:ad:12:bd:55:07 dev vni-600003 vlan 0 master br-699 permanent  
14:18:77:01:5d:01 dev swp1.699 vlan 0 master br-699 permanent
```

```
cumulus@CML-02$ sudo cat /proc/net/vlan/config  
sudo: unable to resolve host CML-02  
[sudo] password for cumulus:  
VLAN Dev name | VLAN ID  
Name-Type: VLAN_NAME_TYPE_RAW_PLUS_VID_NO_PAD  
swp2.200 | 200 | swp2  
swp1.699 | 699 | swp1
```

Troubleshooting and configuration isn't very simple.

Solution: Ansible

Conclusion

- Whitebox Switches are good even in enterprise and ISPs as well
- CumulusLinux works well for almost everything you need in a Layer 3 switch
- VXLAN can solve many problems to help declutter layer 2 network. MTU can be a killer though (50 extra Bytes to accommodate) and it breaks LACP, LLDP, 802.1p (because of Trident II)
- Linux as switching OS is difficult for operations team to manage but tools like Ansible are available to make it simple
- Evaluate your problem vs solution before jumping into another band wagon