

# OPEN NETWORKING REVOLUTION

---

Aftab Siddiqui  
Internet Society  
Technical Engagement Manager - Asia Pacific  
[siddiqui@isoc.org](mailto:siddiqui@isoc.org)

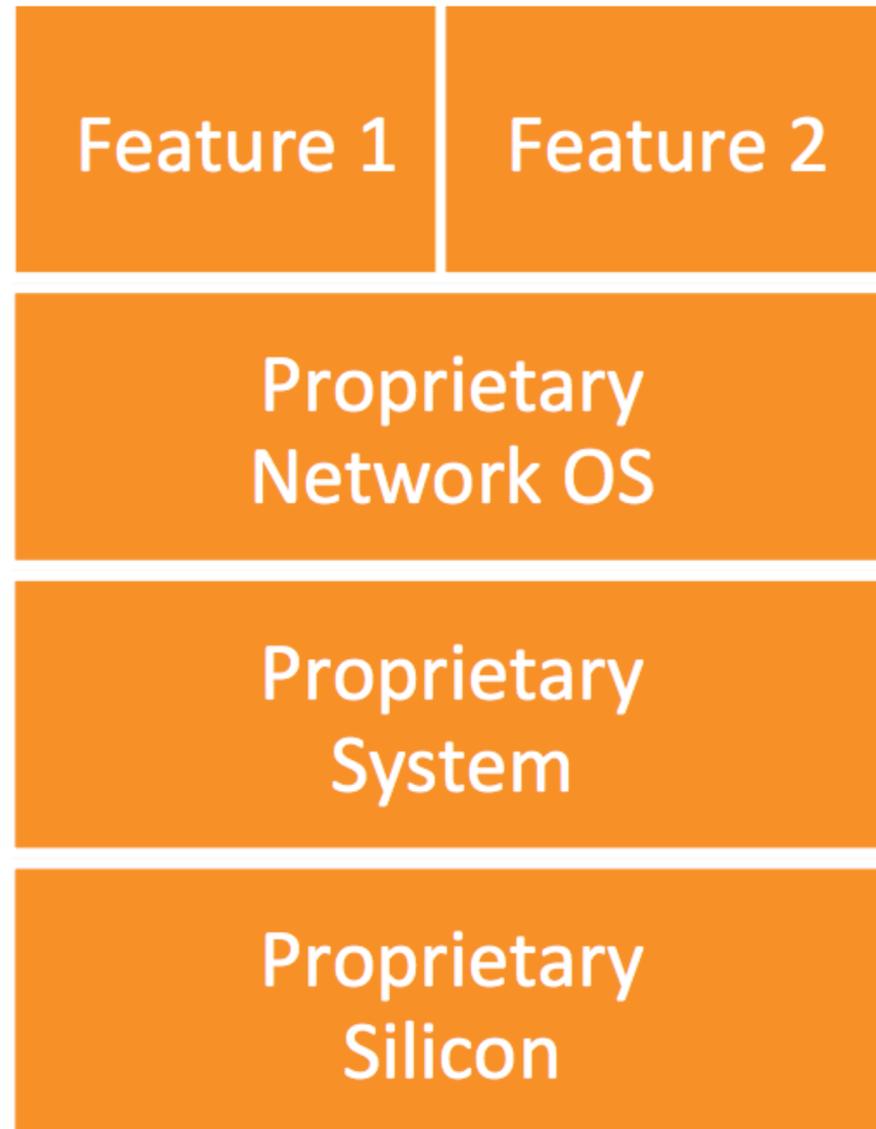
# Open Compute Project

OCP is a collaborative and community focused initiative on redesigning hardware technology to efficiently support the growing demands on compute infrastructure.

## **Projects:**

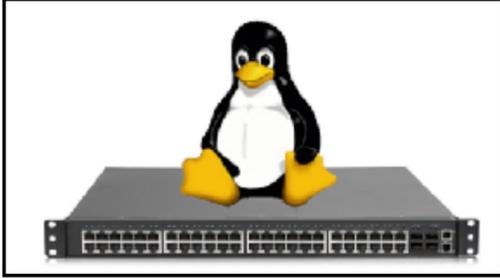
- Storage
- Networking
- Server Design
- Open Rack
- Certification
- Hardware Management
- Data Center

# Traditional Networking



- Proprietary Features
- Few APIs Available
- Locked-in and complex support models
- Mostly proprietary ASICs

# Open Networking

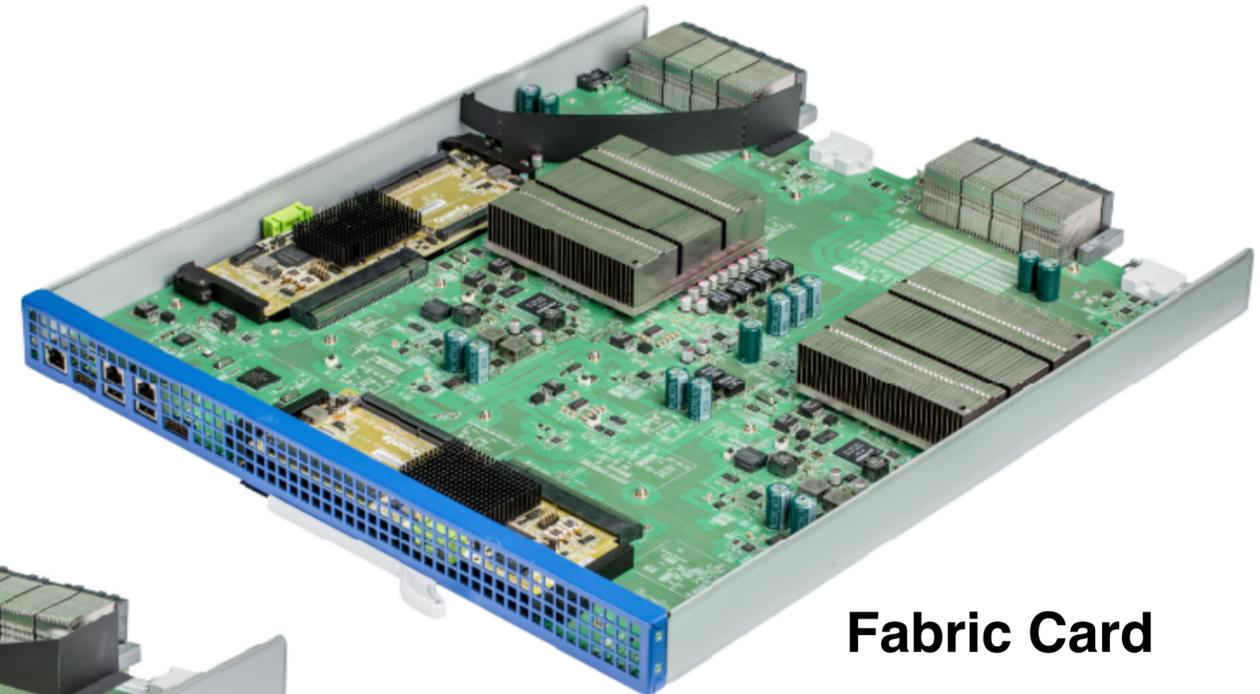
<p><b>Operating System</b></p>		<ul style="list-style-type: none"><li>• Cumulus Linux</li><li>• PicOS</li><li>• IPInfusion</li><li>• SwitchLight OS</li><li>• MS Sonic</li><li>• FBOSS</li><li>• Etc....</li></ul>
<p><b>Open Standard Hardware</b></p>		<ul style="list-style-type: none"><li>• Dell ON</li><li>• HP Altoline</li><li>• EdgeCore</li><li>• Mellanox</li><li>• Etc.....</li></ul>
<p><b>Merchant Silicon</b></p>		<ul style="list-style-type: none"><li>• Broadcom</li><li>• Spectrum</li></ul>

# What are my options?

Vendor	Model	ASIC	Ports	CPU	Memory
Dell	S4048	Trident II	48x10G SFP+ 6x40G QSFP+	Intel Atom C2338	2GB
Dell	S6000	Trident II	32x40G QSFP+	Intel Atom S1220	4GB
Dell	Z9100	Tomahawk	32 x Multirate		
Accton/ EdgeCore/HP	5712	Trident II	48x10G SFP+ 6x40G QSFP+	Intel Rangely C2538	8GB
Accton/ EdgeCore/HP	6712	Trident II	40G x 32	Intel Rangely C2538	8GB
Mellanox	SN2700	Spectrum	32 x Multirate	Spectrum	
Mellanox	SN2100	Spectrum	16 x Multirate	Spectrum	

# Facebook 6-Pack...too complicated?

**Facebook 6-Pack Modular Chassis**



**Fabric Card**



**Line Card  
[16 x 40G]**

# Network Operating System

- Hardware specification of switches are more or less similar but tough task is to pick the right Network Operating System
- Best way is to find major requirement and check the features (no brainer)
- If there is any virtual environment available for those NOS then deploy and test.
- CumulusLinux and IPinfusion both provide virtualized versions of their NOS and without any significant restrictions
- CumulusVX is the most user friendly VM available to-date. It supports all major environments
- All NOS are evolving quite rapidly, keep checking.



# First Impression

```
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?

All the Network Operating Systems are based on Linux and therefore offer the same kind of CLI... CumulusLinux, OcNOS, ONL, Dell OS10 etc

# First Impression

## IOS

```
vlan 100,200

interface ethernet 1/1
  switchport mode access
  switchport access vlan 100

interface ethernet 1/2
  switchport mode access
  switchport access vlan 200
```

## JunOS

```
interfaces {
  ge-0/1/1 {
    unit 0 {
      family bridge {
        interface-mode access;
        vlan-id 100;
      }
    }
  }
  ge-0/1/2 {
    unit 0 {
      family bridge {
        interface-mode access;
        vlan-id 200;
      }
    }
  }
}
```

## Linux

/etc/network/interfaces:

```
auto bridge
iface bridge
  bridge-vlan-aware yes
  bridge-ports swp1 swp2
  bridge-vids 100 200

auto swp1
iface swp1
  bridge-access 100

auto swp2
iface swp2
  bridge-access 200
```

Every NOS offer different port naming convention

Some NOS offer utilities to simplify command line config such as auto complete

# First Deployment Experience

**Target:** Establish Connectivity between multiple PoPs.

**Media:** Dark Fibre

**Service Offering:** Backhaul (Layer 2) and Transit services.

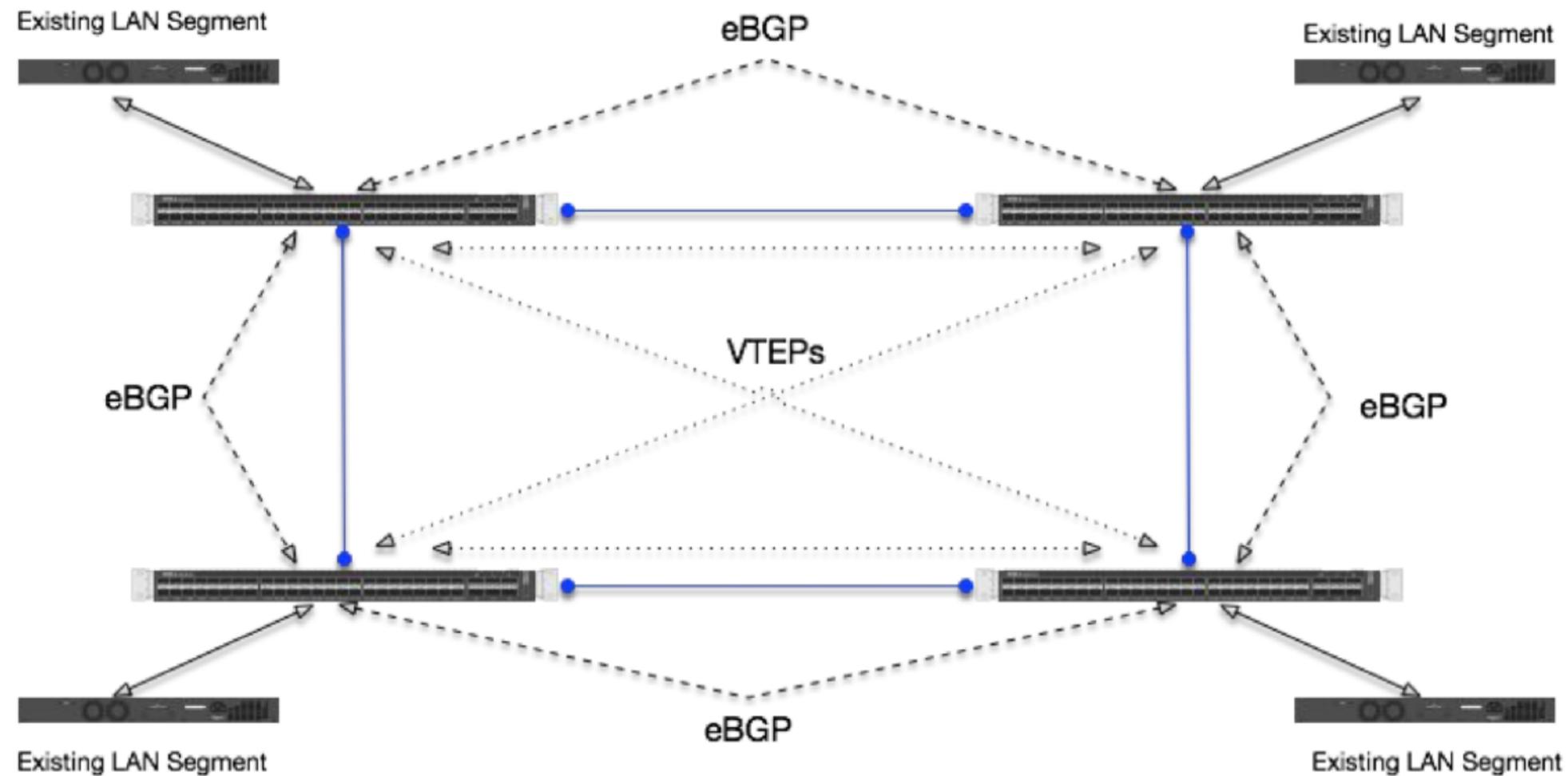
**Budget:** Challenging (we were forced to look into white box switches)

**Project Timelines:** Weeks rather months

**POC:** All services (layer 2 and transit) from 4 PoPs

# First Deployment Experience

- Network Design [Dell S4048-ON switches, CumulusLinux], VXLAN overlay



# Not that Simple

```
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

# Basic Automation

## Variables Required:

hname = HostnameA

swp = Number of Interfaces

lip = Loopback IP

bint = Bond interface name (e.g. bond0)

bslaves = Member of bond interfaces (separate with , or space)

bip = Bond IP

localasn= Local ASN (Eg: 420000XXXX where XXXX is POP-ID)

remotehname = HostnameB

\* nip = Neighbour IP for BGP (e.g. 30.10.0.2)

\* remoteasn = Remote ASN (eg: 420000XXXX where XXXX is POP-ID)

\* - Populated Automatically on selection of Remote Host

# Standard Configuration

- **Create User Credentials**
- **Enable Routing (e.g. Quagga)**
- **NTP**
- **DNS**
- **MOTD**
- **SWP Interfaces**

# Basic Automation

## POP-A:

```
auto swp4  
iface swp4  
mtu 9216
```

```
auto vni-600030  
iface vni-600030  
vxlan-id 600030  
vxlan-local-tunnelip 30.30.0.1  
vxlan-remoteip 30.10.6.1
```

```
auto br-vl104  
iface br-vl104  
bridge-ports swp4.104 vn1-600030  
bridge-stp on
```

## POP-B:

```
auto swp4  
iface swp4  
mtu 9216
```

```
auto vni-600030  
iface vni-600030  
vxlan-id 600030  
vxlan-local-tunnelip 30.10.6.1  
vxlan-remoteip 30.30.0.1
```

```
auto br-vl104  
iface br-vl104  
bridge-ports swp4.104 vn1-600030  
bridge-stp on
```

# First Deployment Experience

- Whitebox Switches are good even in enterprise and ISPs as well. You do need 10G/40G 25G/50G switches in your network.
- CumulusLinux worked well for “almost” everything we needed But review your requirement before selecting NOS.
- VXLAN can solve many problems to help de-clutter layer 2 network. MTU can be a killer though (50 extra Bytes to accommodate) and it breaks LACP and LLDP.
- Operationally simple and economical deployment IF you have proper automation.



**Thanks**

**Any Questions**

# First Deployment Experience

- VXLAN Packet Captures

```
▶ Frame 13: 209 bytes on wire (1672 bits), 209 bytes captured (1672 bits) on interface 0
▶ Ethernet II, Src: Dell_01:4b:30 (14:18:77:01:4b:30), Dst: Dell_01:5d:30 (14:18:77:01:5d:30)
▶ Internet Protocol Version 4, Src: 1.1.1.1, Dst: 2.2.2.2
▶ User Datagram Protocol, Src Port: 37103 (37103), Dst Port: 4789 (4789)
▶ Virtual eXtensible Local Area Network
▶ Ethernet II, Src: CiscoInc_85:3f:79 (00:13:80:85:3f:79), Dst: CiscoInc_87:88:21 (00:1f:ca:87:88:21)
▶ Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.2
▶ Internet Control Message Protocol
```

```
▶ Frame 2: 568 bytes on wire (4544 bits), 568 bytes captured (4544 bits) on interface 0
▶ Ethernet II, Src: Dell_f7:67:09 (34:17:eb:f7:67:09), Dst: Dell_f8:0d:09 (34:17:eb:f8:0d:09)
▶ Internet Protocol Version 4, Src: 30.0.0.1, Dst: 30.0.0.2
▶ User Datagram Protocol, Src Port: 10123 (10123), Dst Port: 4789 (4789)
▶ Virtual eXtensible Local Area Network
▶ Ethernet II, Src: CiscoInc_37:1d:a0 (00:13:19:37:1d:a0), Dst: CiscoInc_87:88:20 (00:1f:ca:87:88:20)
▶ 802.1Q Virtual LAN, PRI: 0, CFI: 0, ID: 100
▶ Internet Protocol Version 4, Src: 60.0.0.2, Dst: 60.0.0.1
▶ Internet Control Message Protocol
```