EVPN:
Or how I learned to stop worrying and love the BGP

Tom Dwyer, JNCIE-ENT #424
Clay Haynes, JNCIE-SEC # 69 JNCIE-ENT # 492

What is EVPN?

- VPN technology that provides L2 or integrated L2+L3 VPN
- Uses MP-BGP – AFI =25 and SAFI=70
- Solves multiple issues with DCI/L2 stretch in an open standard
- Three Data Plane Choices:
  - MPLS
  - Overlay (think VXLAN)
  - PBB

Why I Love the BGP

- L2/L3 Information is advertised via the control plane by BGP
- BGP is scalable
- Supports all Active/Active multi-homing
- Quick Failover for segment/equipment failures
- Doomsday devices blast radius affects one DC instead of all DCs

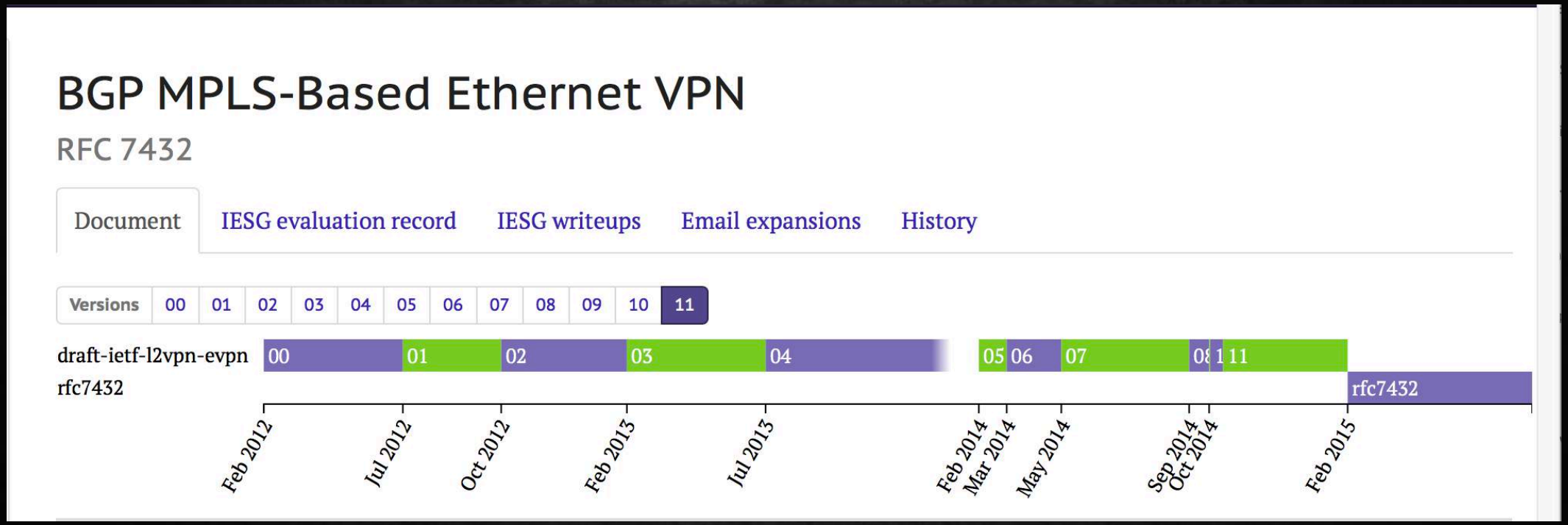
EVPN Terms

- Ethernet Segment
- Ethernet Tag
- Ethernet Segment Identifier (ESI)
- EVPN Instance (EVI)

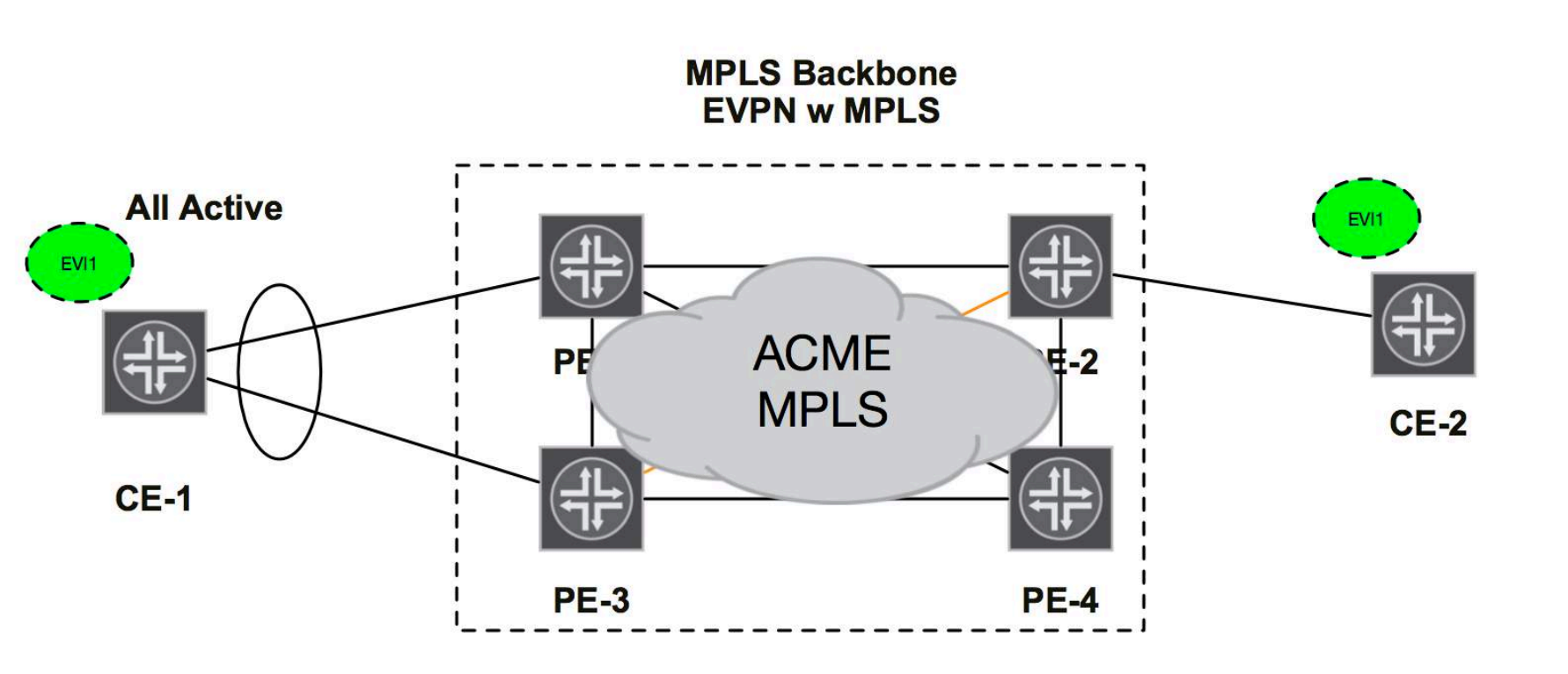
EVPN and MPLS

DCI for the Service Provider

MPLS-Based Ethernet VPN
RFC 7432

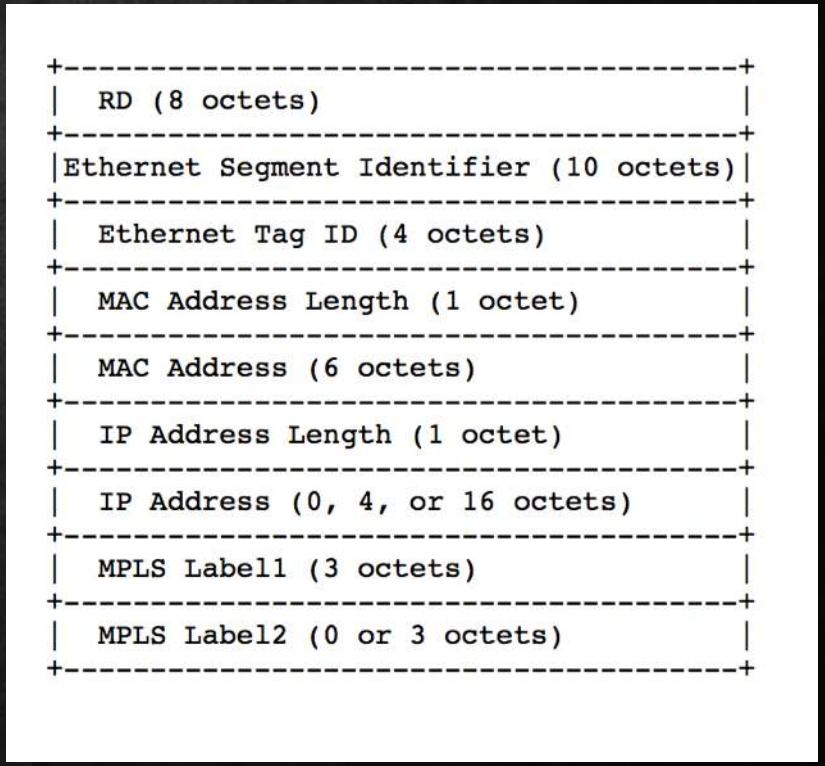


EVPN Sample Topology



MAC Advertisement

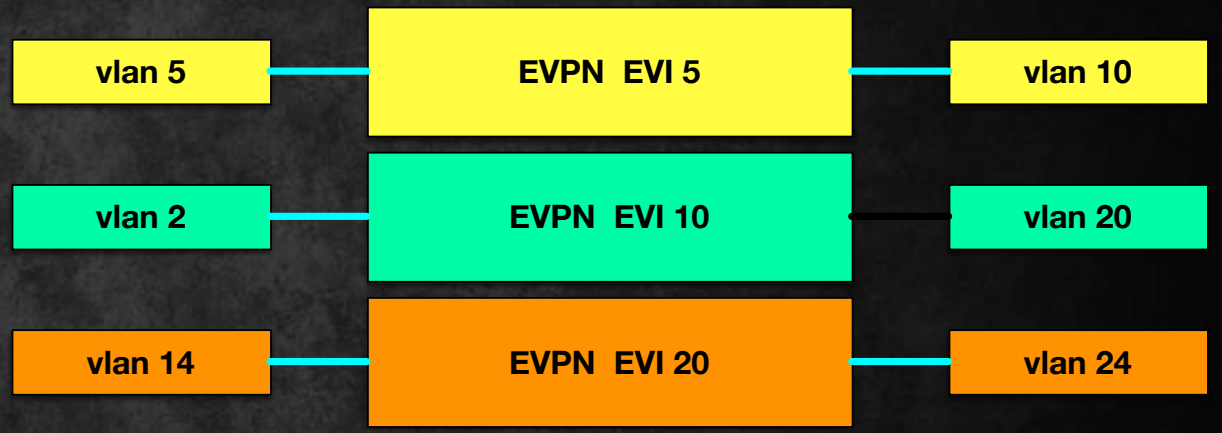
- PE learns local MACs
- Learned MACs are advertised to Remote PE's as Type 2 MAC Address Route
- PE adds extended community to MAC address route for its IRB interface in the respective VLAN



MAC Advertisement – Services

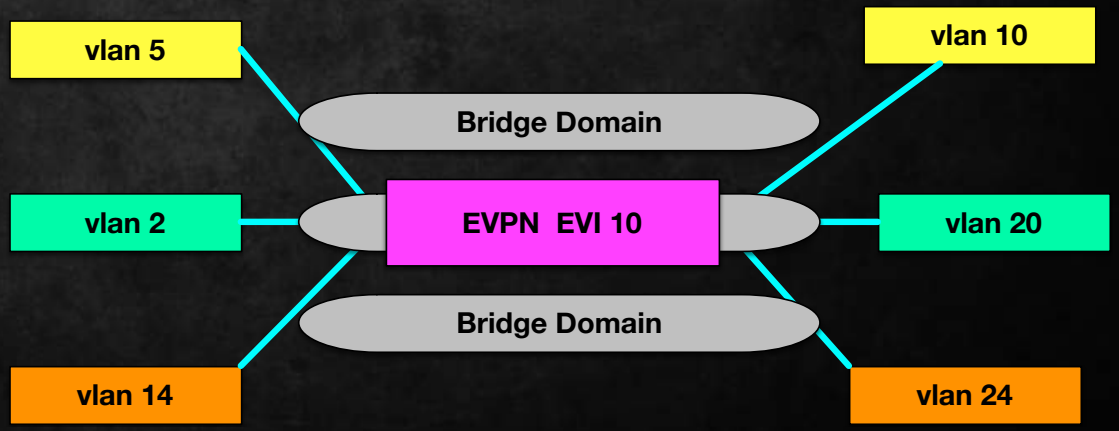
Vlan Base Service Interface

- Single bridge domain per EVI
- 1:1 mapping between Vlan ID and EVI
- Ethernet tag in route update set to 0
- Vlan translation can occur at Egress PE
- Label created per EVI



Vlan Aware Bundle

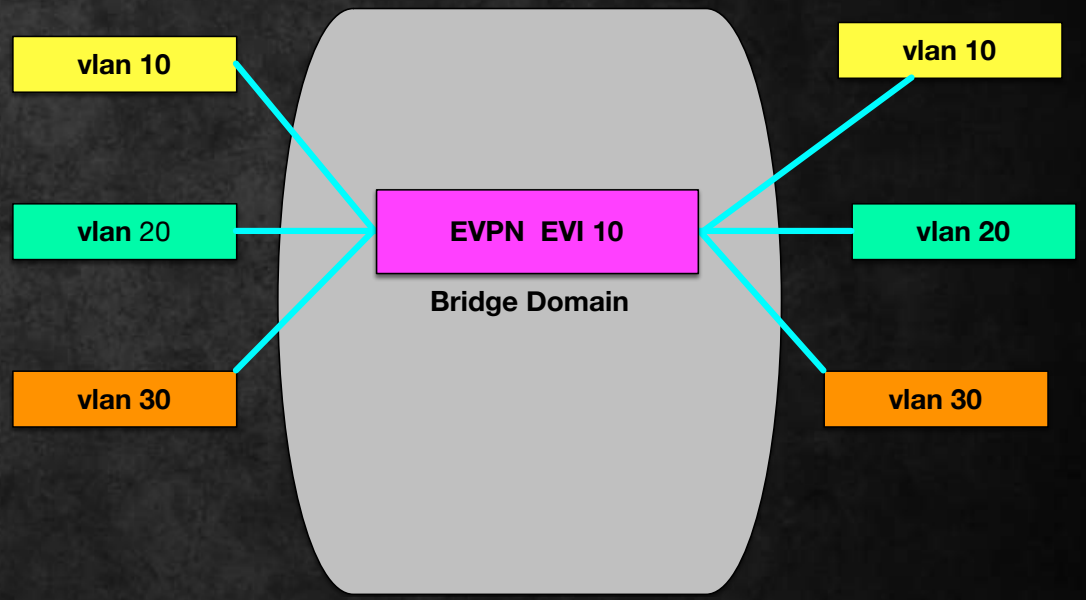
- Multiple VLANs
- N:1 mapping between Vlan ID and EVI
- Ethernet tag in route is set to the tag value
- Mutiple bridge domains, one per vlan
- Label created per vlan



MAC Advertisement – Services

Vlan Bundle Service Interface

- Single bridge domain per EVI
- Many –to-one mapping VLAN ID and EVI
- Ethernet tag in route update set to 0
- MACs unique across VLANs
- Vlan translation NOT ALLOWED

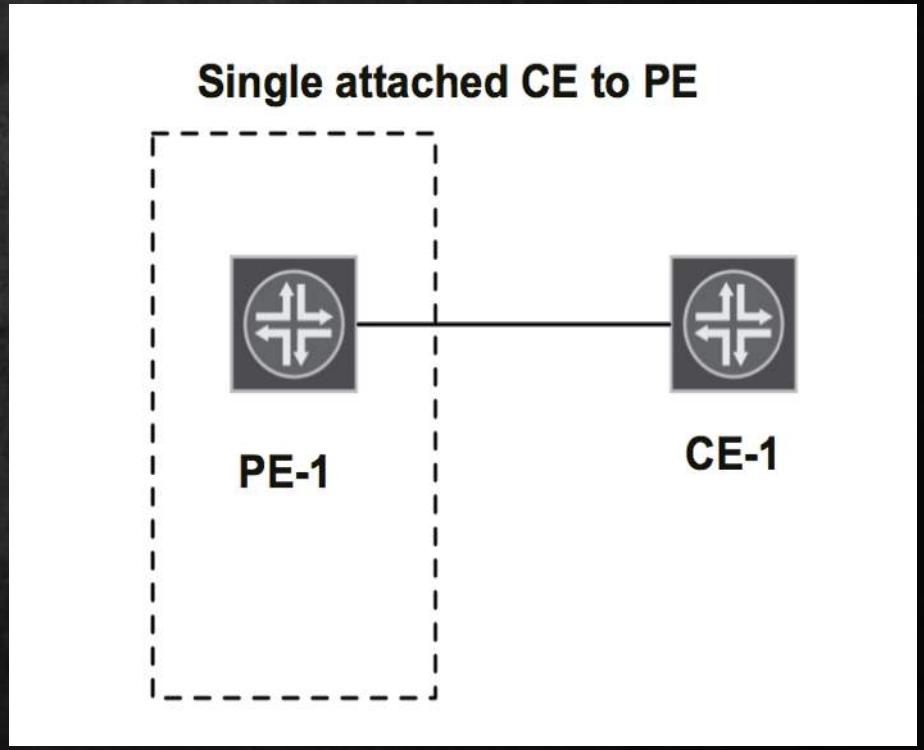


EVPN Multi-homing

- Single
- Active-Standby
- Active-Active

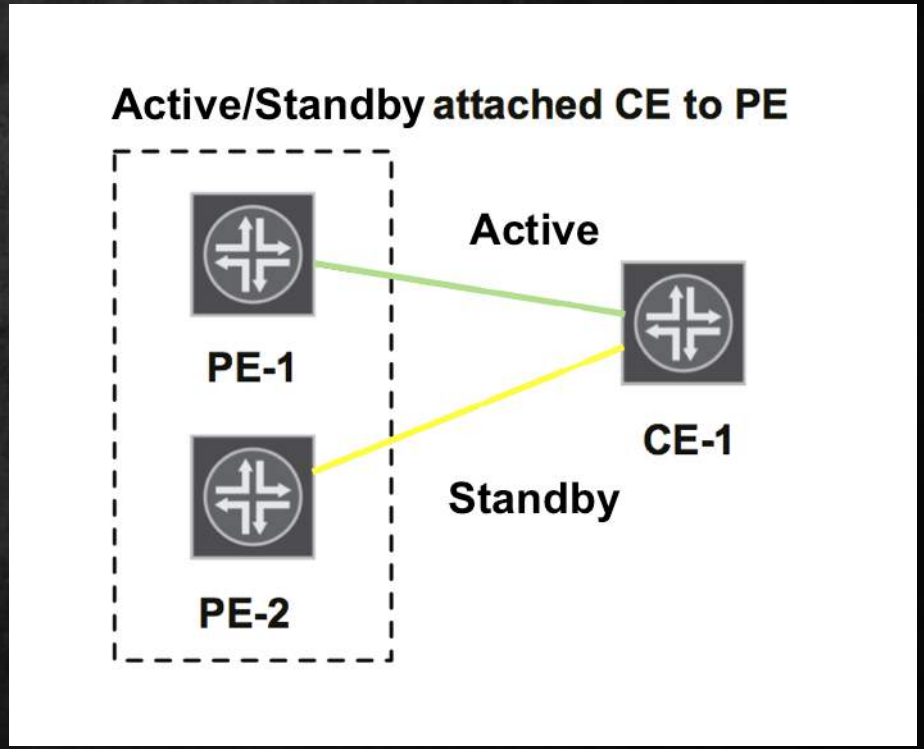
EVPN Multi-homing

- Single-homed Instance
- No ESI



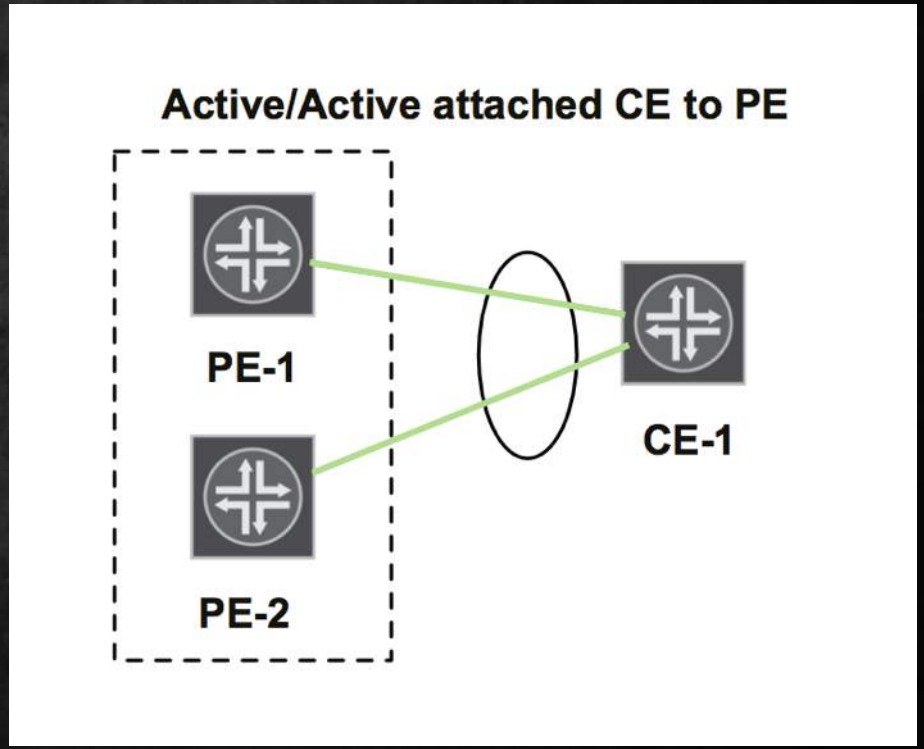
EVPN Multi-homing

- One PE is selected as a Designated Forwarder
- Only Designated Forwarder can pass traffic
- ESI is included as a community



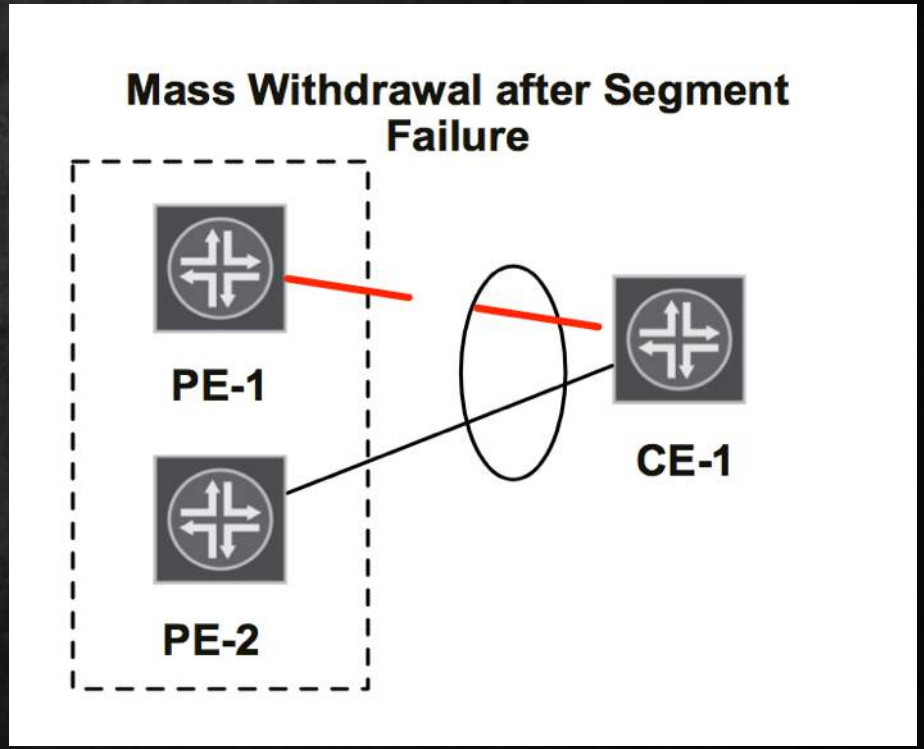
EVPN Multi-homing

- Multiple PE's can forward traffic for specific ESI
- Only Designated Forwarder can pass BUM traffic
- Split Horizon



EVPN MAC Mass withdrawal

- Local PE withdraws routes on ESI link failure
- Remote PE's are signaled to update/remove Next-Hops for routes associated with ESI
- If PE fails a new designated forwarder is elected and remote PE's update Next-Hops for routes associated with ESI



EVPN MAC Mobility

- Each EVPN route has extended community - MAC Mobility Sequence Number
- PE's may not detect MAC Address moves during certain events (vMotion)
- Upon a vMotion Event:
  - New PE learns MAC locally after vMotion
  - New PE advertises MAC to EVPN peers and increments MAC Mobility Sequence Number
  - Remote PE's see higher Sequence Number and prunes old MAC Route
  - Original PE sees the new route and withdraws the old route advertisement

Doomsday Devices!

- Each PE will learn the MAC or ARP entry before traffic is passed
- Local MACs are utilized for local traffic
- PE Proxy-ARPs for hosts present in the EVPN Route table
- All other traffic is dropped if not local nor present in the EVPN Route Table
- If the same MAC Route is learned 5 times in 180 seconds, the route is suppressed for period of time, limiting MAC Flapping

EVPN and VXLAN

DCI for the rest of us

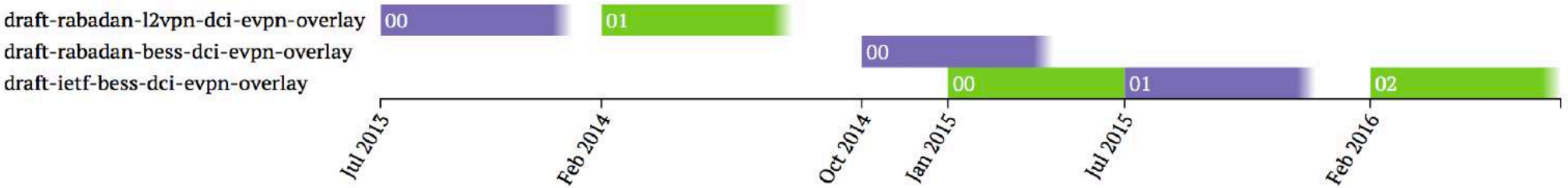
EVPN Overlay (NVO)

Interconnect Solution for EVPN Overlay networks

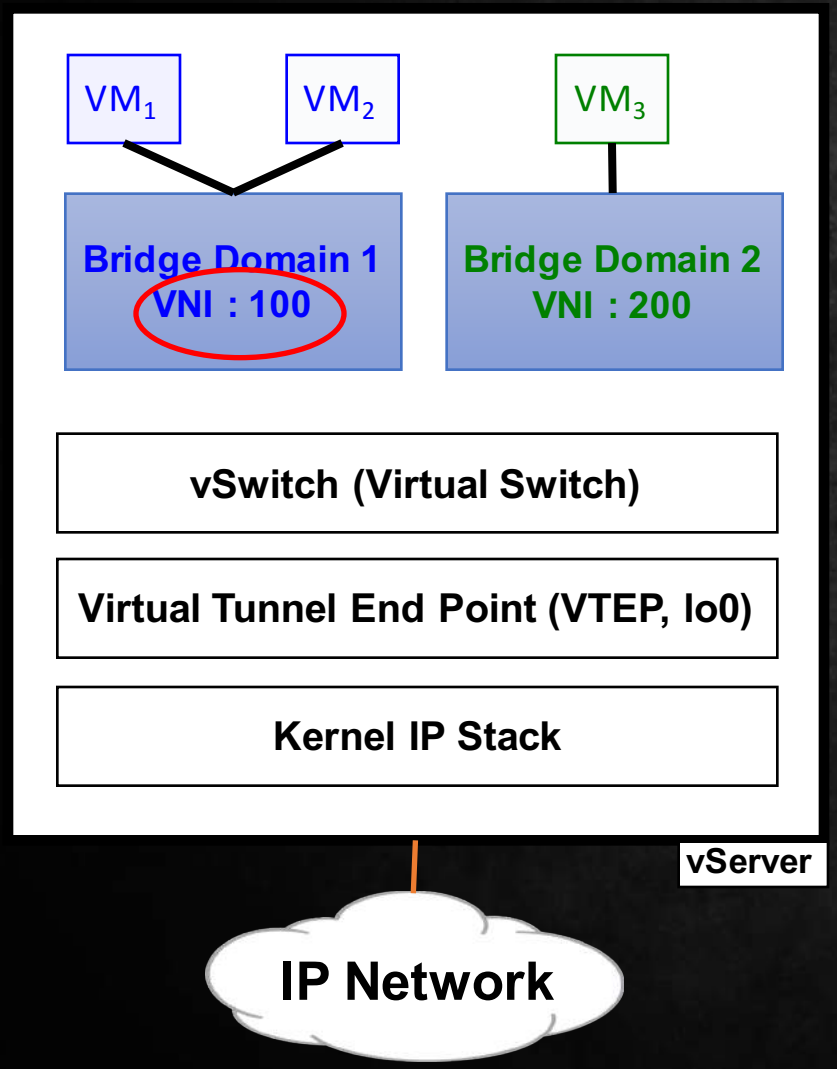
draft-ietf-bess-dci-evpn-overlay-02

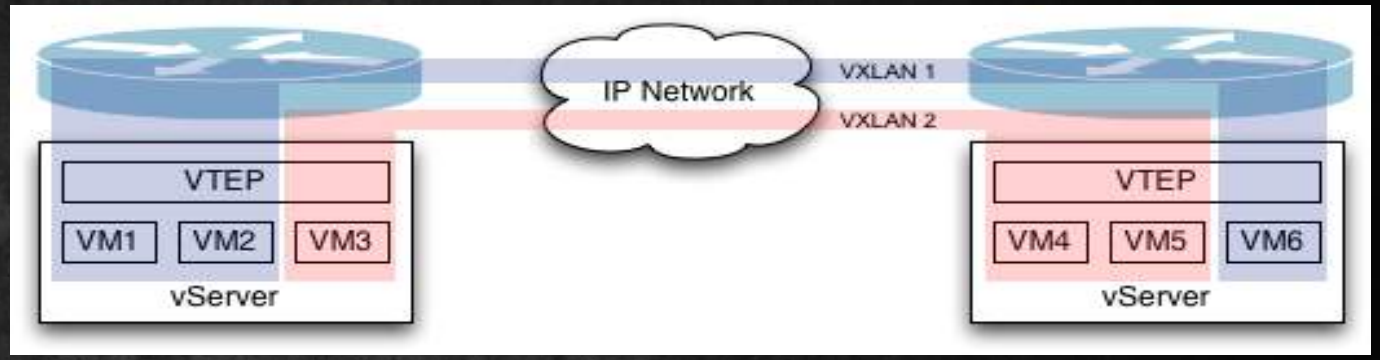
Document | IESG evaluation record | IESG writeups | Email expansions | History

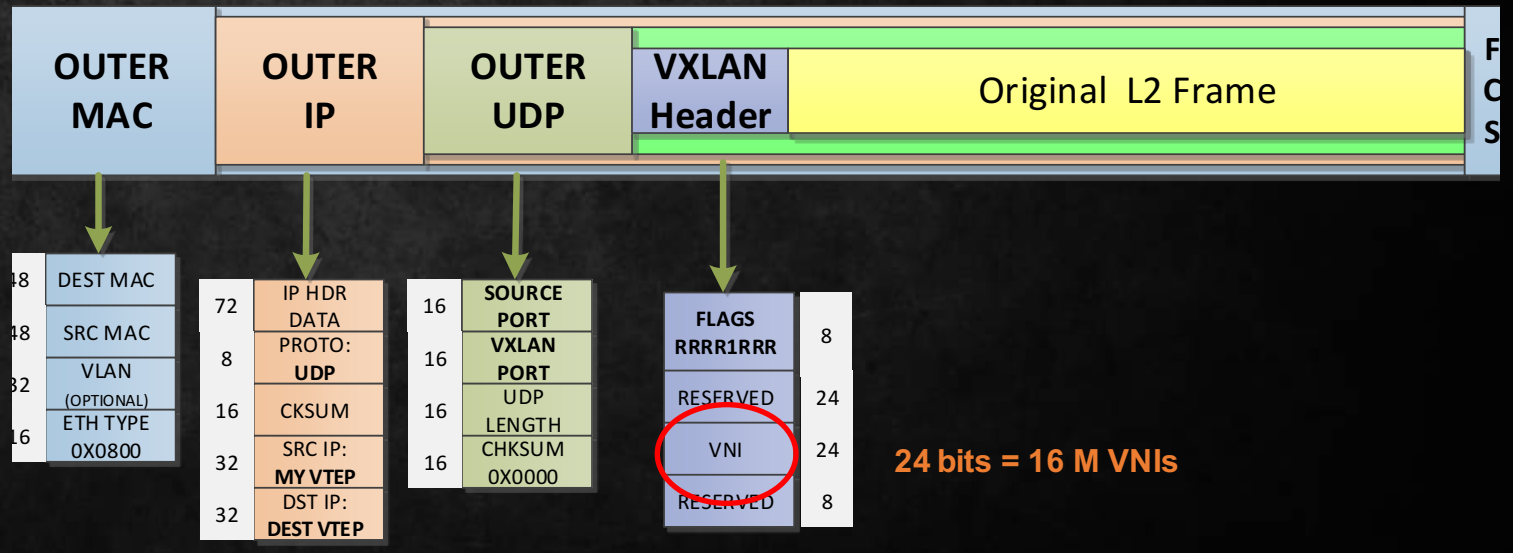
Versions 00 01 02



VXLAN : Building blocks





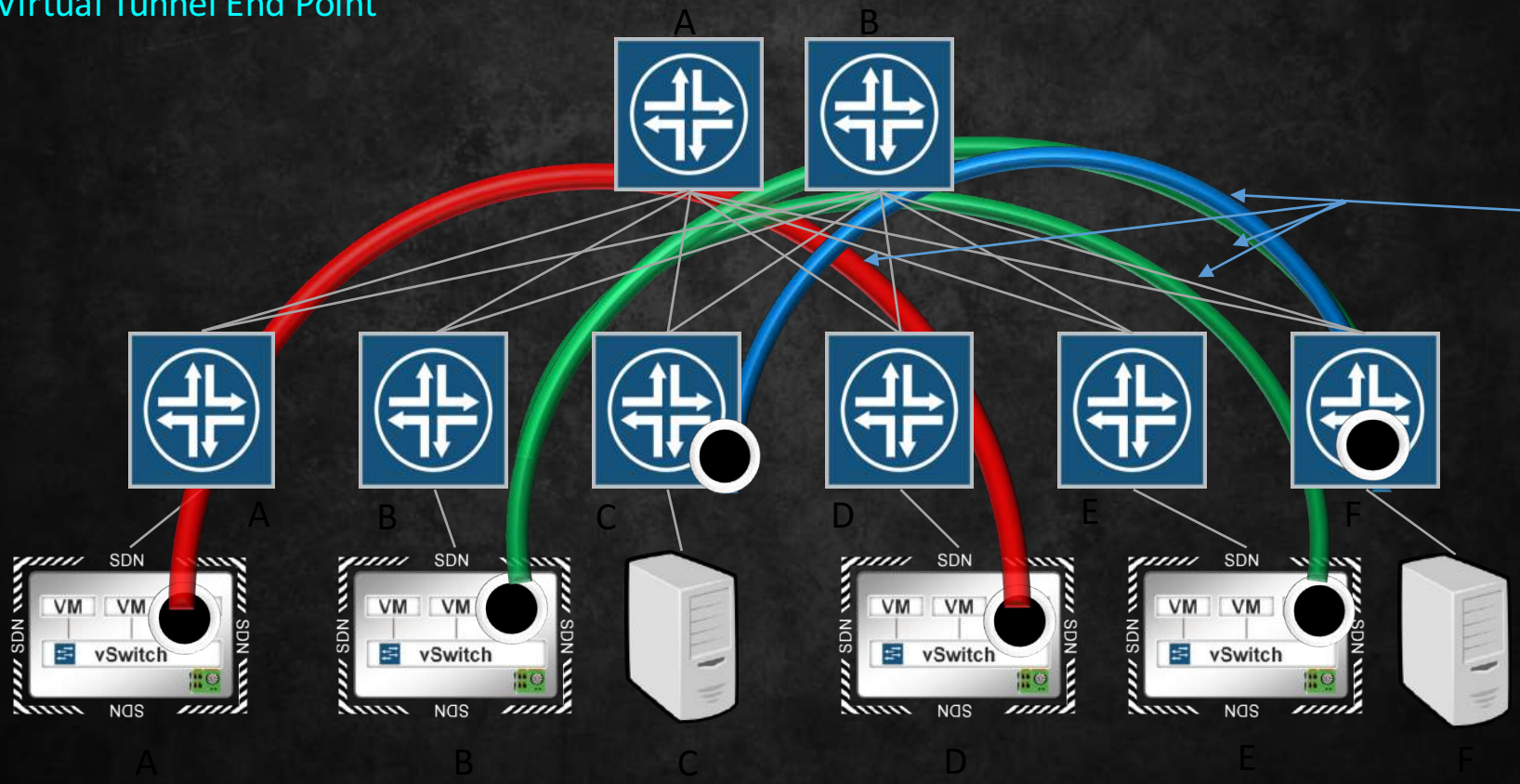


24 bits = 16 M VNIs


YOU GET A VNI! YOU GET A VNI!
EVERYONE GETS A VNI!
imgflip.com

VXLAN – Putting it Together

VTEP: Virtual Tunnel End Point



Routers

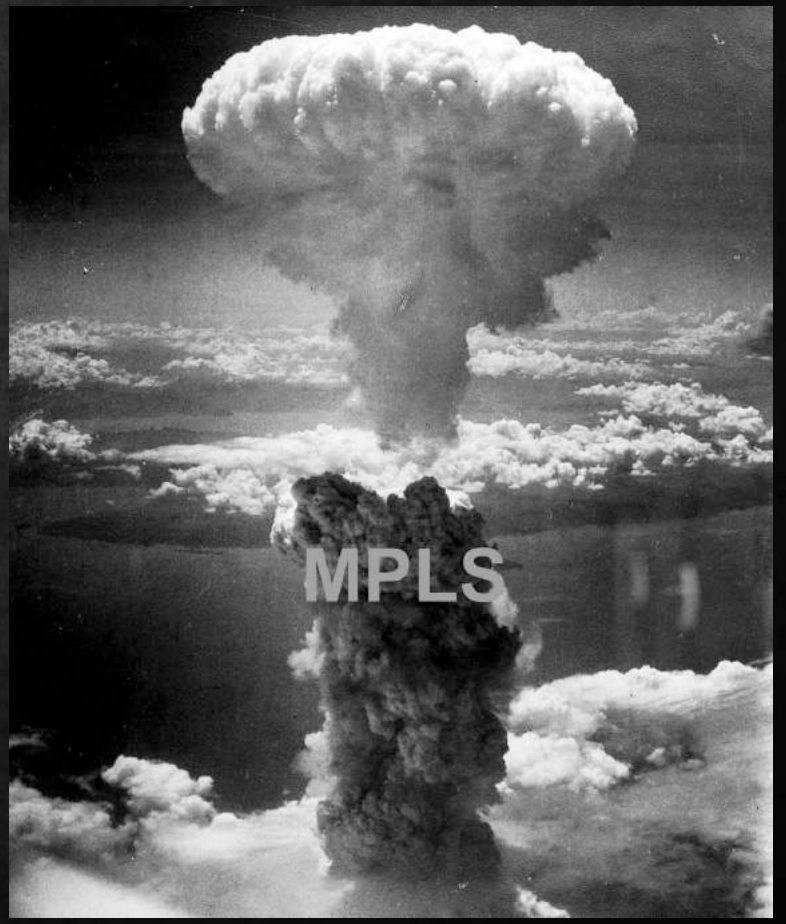
VXLAN tunnels

TOR Switches

Servers

Why VXLAN/EVPN?

- Limited hardware specs
- GRE hashing across WAN limits
- IP Fabrics are becoming more popular
- In enterprise, MPLS is *really HARD!* …Or so they say

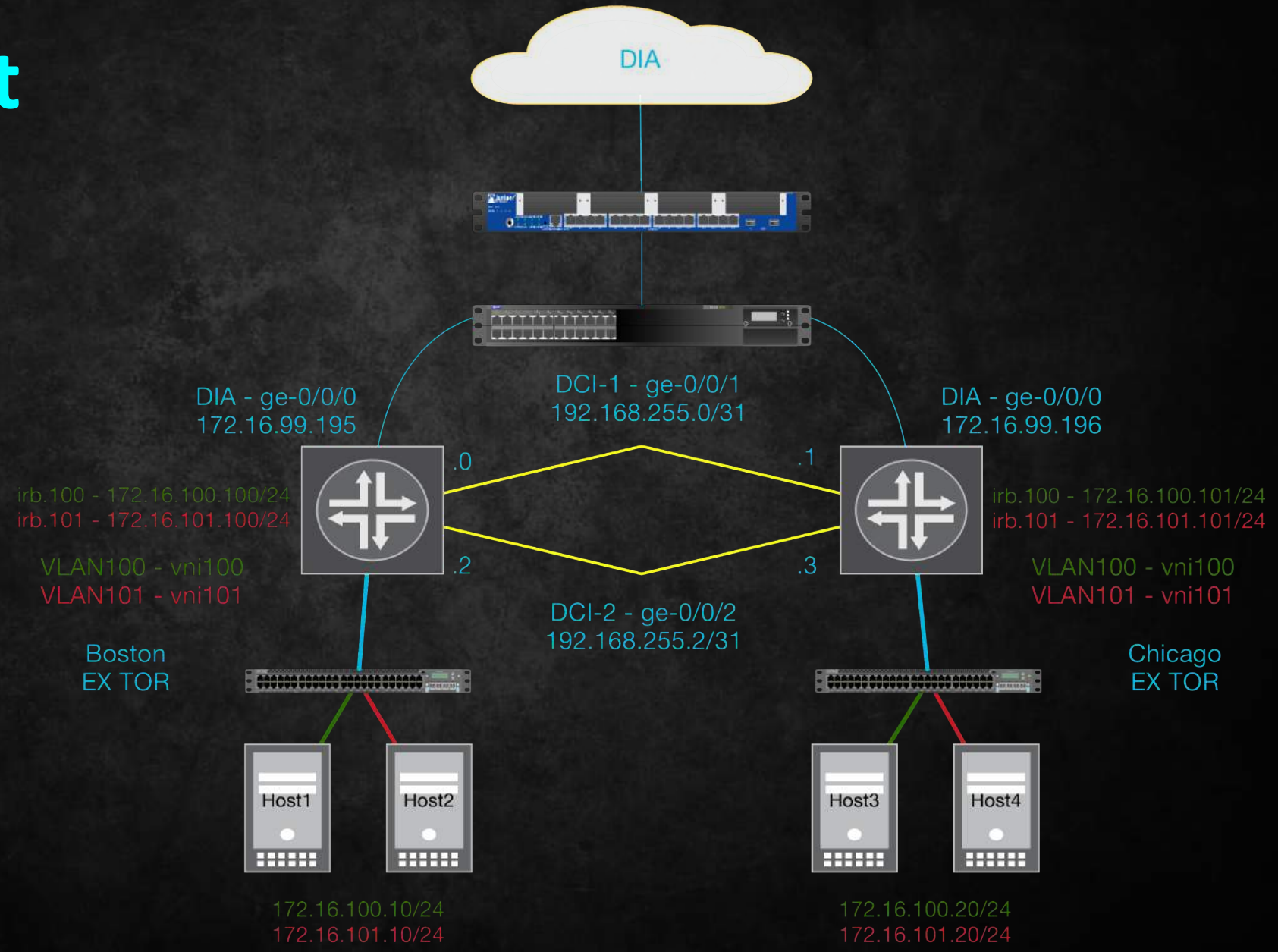


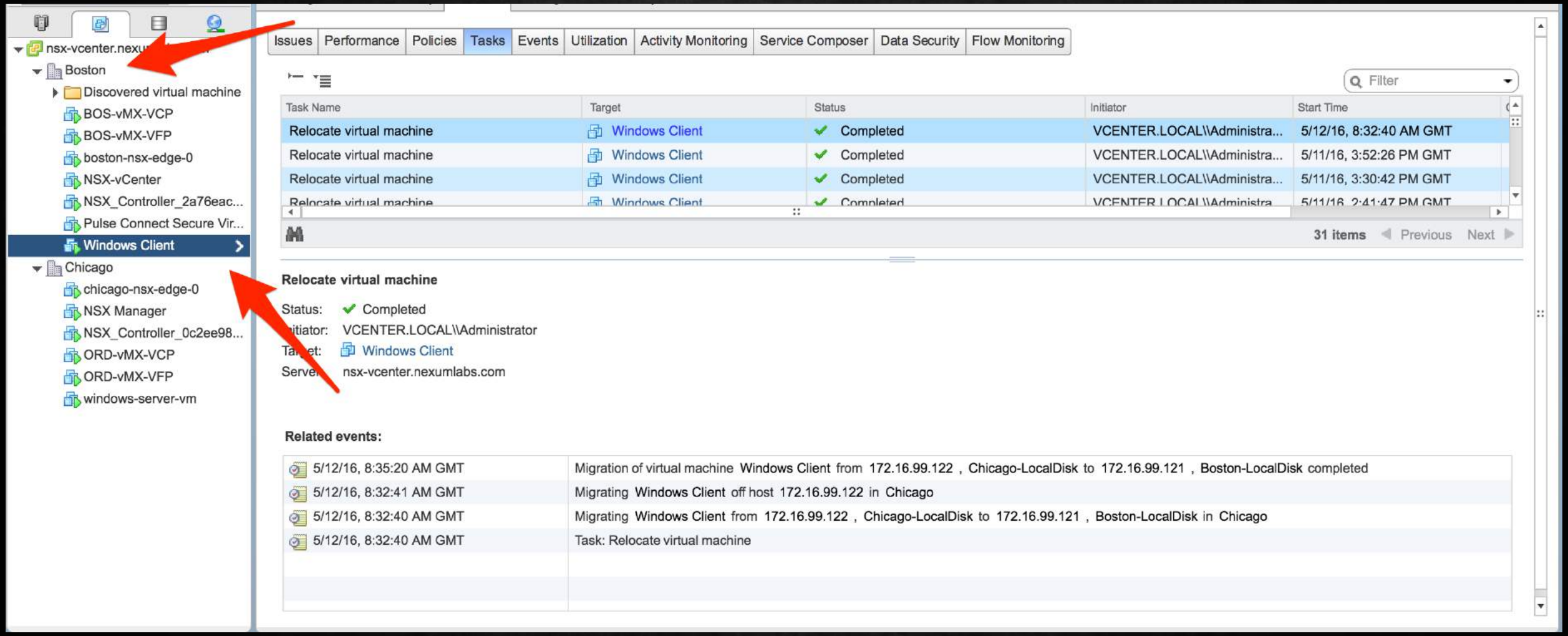
National Archives image (208-N-43888)

VXLAN Deployment Options

Data plane Based	Control Plane Based
Virtual Networks created using Multicast (PIM) groups. Susceptible to data trombone effects across DC's	Virtual Networks created using **3rd party controllers** Virtual Networks with benefits such as VM traffic optimization
PIM creates fully meshed P2P tunnels for known unicast PIM creates multicast tunnels for L2 BUM	Virtual Network IDs (VNID) communicated using EVPN Fully meshed VXLAN tunnels forward traffic

Lab Layout


DIA
DIA - ge-0/0/0
172.16.99.195
DCI-1 - ge-0/0/1
192.168.255.0/31
DIA - ge-0/0/0
172.16.99.196
.0
.1
.2
.3
irb.100 - 172.16.100.100/24
irb.101 - 172.16.101.100/24
VLAN100 - vni100
VLAN101 - vni101
irb.100 - 172.16.100.101/24
irb.101 - 172.16.101.101/24
VLAN100 - vni100
VLAN101 - vni101
DCI-2 - ge-0/0/2
192.168.255.2/31
Boston
EX TOR
Chicago
EX TOR
Host1
Host2
Host3
Host4
172.16.100.10/24
172.16.101.10/24
172.16.100.20/24
172.16.101.20/24

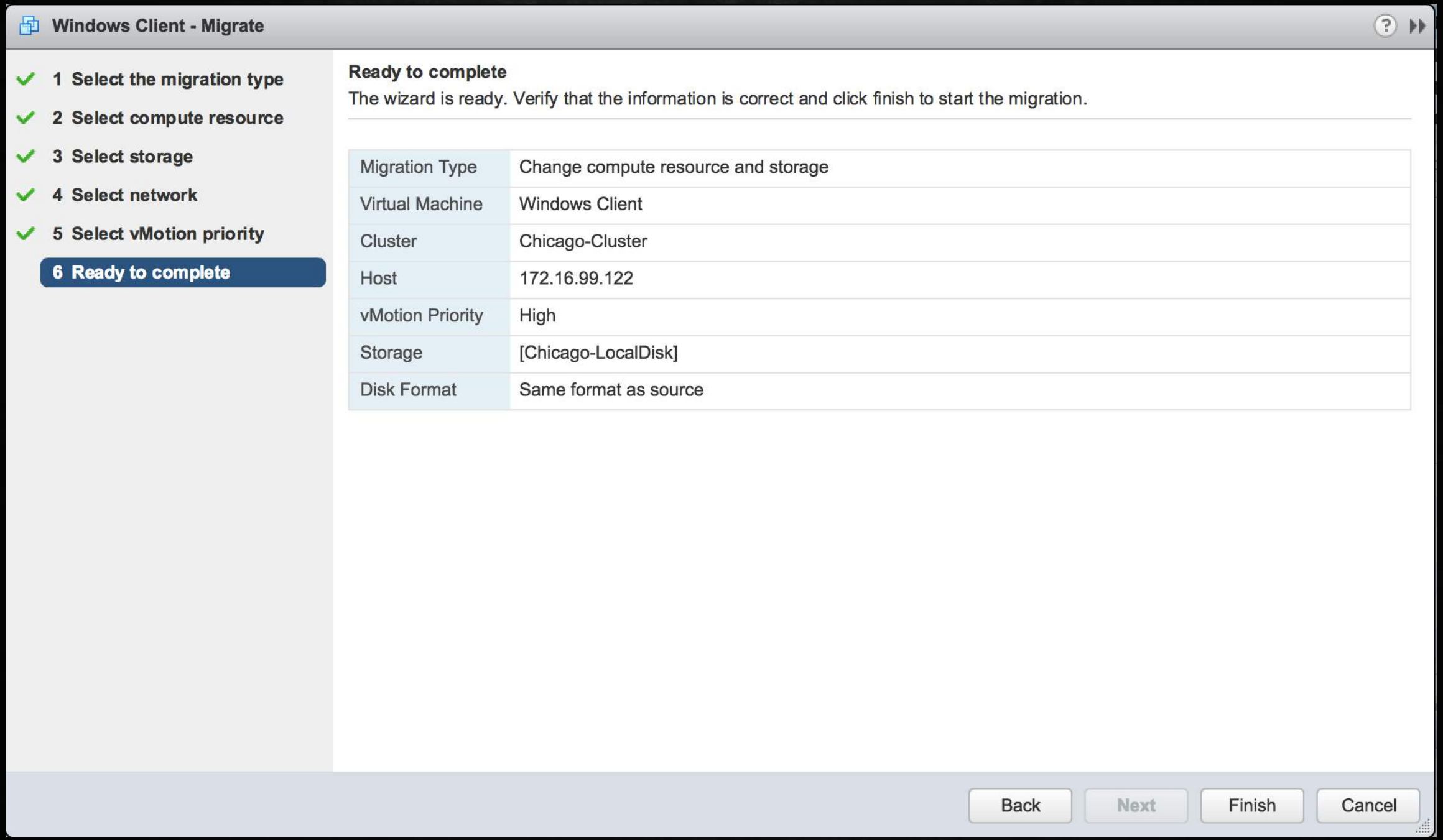

nsx-vcenter.nexu
Boston
Discovered virtual machine
BOS-vMX-VCP
BOS-vMX-VFP
boston-nsx-edge-0
NSX-vCenter
NSX_Controller_2a76eac...
Pulse Connect Secure Vir...
Windows Client
Chicago
chicago-nsx-edge-0
NSX Manager
NSX_Controller_0c2ee98...
ORD-vMX-VCP
ORD-vMX-VFP
windows-server-vm
Issues
Performance
Policies
Tasks
Events
Utilization
Activity Monitoring
Service Composer
Data Security
Flow Monitoring
Filter
Task Name
Target
Status
Initiator
Start Time
Relocate virtual machine
Windows Client
Completed
VCENTER.LOCAL\\Administra...
5/12/16, 8:32:40 AM GMT
Relocate virtual machine
Windows Client
Completed
VCENTER.LOCAL\\Administra...
5/11/16, 3:52:26 PM GMT
Relocate virtual machine
Windows Client
Completed
VCENTER.LOCAL\\Administra...
5/11/16, 3:30:42 PM GMT
31 items
Previous
Next
Relocate virtual machine
Status: Completed
Initiator: VCENTER.LOCAL\\Administrator
Target: Windows Client
Server: nsx-vcenter.nexumlabs.com
Related events:
5/12/16, 8:35:20 AM GMT
Migration of virtual machine Windows Client from 172.16.99.122 , Chicago-LocalDisk to 172.16.99.121 , Boston-LocalDisk completed
5/12/16, 8:32:41 AM GMT
Migrating Windows Client off host 172.16.99.122 in Chicago
5/12/16, 8:32:40 AM GMT
Migrating Windows Client from 172.16.99.122 , Chicago-LocalDisk to 172.16.99.121 , Boston-LocalDisk in Chicago
5/12/16, 8:32:40 AM GMT
Task: Relocate virtual machine

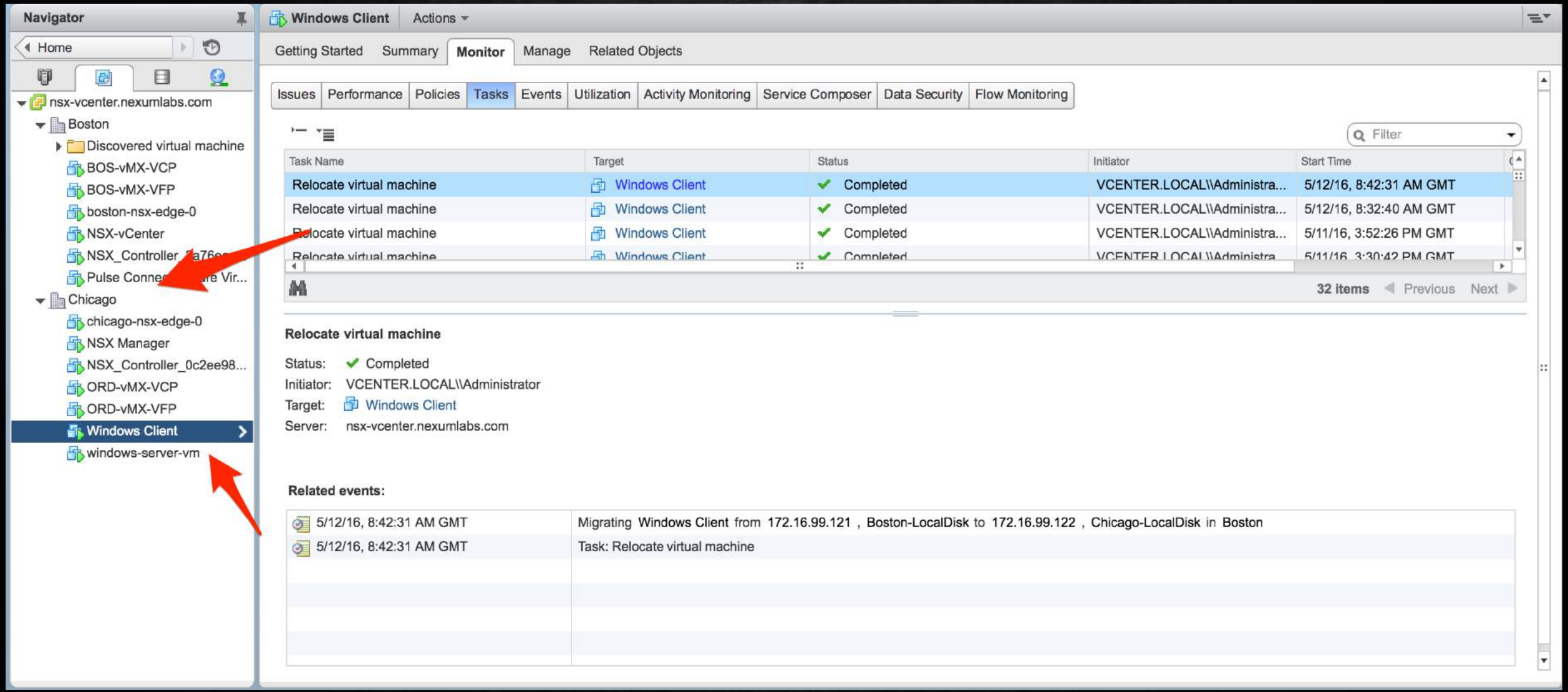
```
bgp.evpn.0: 8 destinations, 8 routes (8 active, 0 holddown, 0 hidden)

EVPN100.evpn.0: 4 destinations, 4 routes (4 active, 0 holddown, 0 hidden)

EVPN101.evpn.0: 10 destinations, 10 routes (10 active, 0 holddown, 0 hidden
)
+ = Active Route, - = Last Active, * = Both

2:172.16.31.1:2::101::00:0c:29:1e:3f:a8/304
                   *[EVPN/170] 00:05:02
                      Indirect

root@vmx-boston>
```

Boston

```
EVPN100.evpn.0: 4 destinations, 4 routes (4 active, 0 holddown, 0 hidd

EVPN101.evpn.0: 10 destinations, 10 routes (10 active, 0 holddown, 0 h
+ = Active Route, - = Last Active, * = Both

2:172.16.31.1:2::101::00:0c:29:1e:3f:a8/304
                   *[BGP/170] 00:05:08, localpref 100, from 172.16.31.
                      AS path: I, validation-state: unverified
                      to 192.168.255.0 via ge-0/0/1.0
                    > to 192.168.255.2 via ge-0/0/2.0

root@vmx-chicago>
```

Chicago


Windows Client - Migrate
1 Select the migration type
2 Select compute resource
3 Select storage
4 Select network
5 Select vMotion priority
6 Ready to complete
Ready to complete
The wizard is ready. Verify that the information is correct and click finish to start the migration.
Migration Type
Change compute resource and storage
Virtual Machine
Windows Client
Cluster
Chicago-Cluster
Host
172.16.99.122
vMotion Priority
High
Storage
[Chicago-LocalDisk]
Disk Format
Same format as source
Back
Next
Finish
Cancel


Navigator
Home
nsx-vcenter.nexumlabs.com
Boston
Discovered virtual machine
BOS-vMX-VCP
BOS-vMX-VFP
boston-nsx-edge-0
NSX-vCenter
Chicago
chicago-nsx-edge-0
NSX Manager
ORD-vMX-VCP
ORD-vMX-VFP
Windows Client
windows-server-vm
Windows Client
Actions
Getting Started
Summary
Monitor
Manage
Related Objects
Issues
Performance
Policies
Tasks
Events
Utilization
Activity Monitoring
Service Composer
Data Security
Flow Monitoring
Filter
Task Name
Target
Status
Initiator
Start Time
Relocate virtual machine
Windows Client
Completed
VCENTER.LOCAL\\Administra...
5/12/16, 8:42:31 AM GMT
Relocate virtual machine
Windows Client
Completed
VCENTER.LOCAL\\Administra...
5/12/16, 8:32:40 AM GMT
Relocate virtual machine
Windows Client
Completed
VCENTER.LOCAL\\Administra...
5/11/16, 3:52:26 PM GMT
32 items
Previous
Next
Relocate virtual machine
Status: Completed
Initiator: VCENTER.LOCAL\\Administrator
Target: Windows Client
Server: nsx-vcenter.nexumlabs.com
Related events:
5/12/16, 8:42:31 AM GMT
Migrating Windows Client from 172.16.99.121 , Boston-LocalDisk to 172.16.99.122 , Chicago-LocalDisk in Boston
5/12/16, 8:42:31 AM GMT
Task: Relocate virtual machine

```
EVPN101.evpn.0: 10 destinations, 10 routes (10 active, 0 holddown, 0 hidden
)
+ = Active Route, - = Last Active, * = Both

2:172.16.31.2:2::101::00:0c:29:1e:3f:a8/304
                   *[BGP/170] 00:04:45, localpref 100, from 172.16.31.2
                     AS path: I, validation-state: unverified
                     to 192.168.255.1 via ge-0/0/1.0
                   > to 192.168.255.3 via ge-0/0/2.0

root@vmx-boston>
```

Boston

```
EVPN101.evpn.0: 10 destinations, 10 routes (10 active, 0 holddown, 0 h
+ = Active Route, - = Last Active, * = Both

2:172.16.31.2:2::101::00:0c:29:1e:3f:a8/304
                   *[EVPN/170] 00:04:47
                      Indirect

root@vmx-chicago>
```

Chicago