



INTEGRATION OF MULTIPLE OPENSTACK CLOUDS
WITH A CORE MPLS NETWORK

Curtis Collicutt

- OpenStack Architect
- Interdynamix
- curtis@interdynamix.com

Rahul Kumar

- Sr. Network Engineer
- Interdynamix
- rahul@interdynamix.com

- Standing on shoulders of giants
 - All the people work on OpenStack other open source systems
 - BGP-VPN/Bagpipe presentations and discussions
 - People who create networking labs and publish them
 - People who write books on MPLS, eg. “MPLS in the SDN Era”
- Perspective - From an **OpenStack Operators** perspective, how do we:
 - Integrate a large number of OpenStack deployments with a SP-like MPLS WAN
 - Automate networking
 - Participate in modern networking paradigms, ie. NFV, network slicing, SDN...
 - Simulate these systems in some capacity to understand them better
 - Investigate and observe their operation

- **Omnipresence** - Many devices understand MPLS
 - Including vswitches
- **Has aged gracefully** - Well understood, interoperable implementations, standards based
- **Scales reasonably**
- **Large % of mobile traffic** crosses an MPLS network at some point
 - Also used by non-telecoms
- **Protocol neutral**
- **Supports Traffic Engineering**
 - RSVP-TE - Resource Reservation Protocol with Traffic Engineering
- **Fast reroute**

- MPLS as a paradigm vs MPLS as an encapsulation
 - “MPLS is not an encapsulation. It is an architectural framework...”
- Decoupling:
 - control plane from data/forwarding plan
 - service from transport
 - overlay from underlay
- Layered architecture with a feature-rich edge and a fast transport core. This approach can be applied to the WAN, to data centers, and so on.
- Building overlay networks at the edge in order to support multitenancy and multiservice
- Minimizing the forwarding state in the core
- Advanced packet steering by either signaling forwarding paths and/or by stacking instructions on packet headers



Daniel Dib @danieldibswe · Apr 21

"MPLS is dead" - I can guarantee my packets from me to Twitter were MPLS switched at some point in the path. Where's your argument now? What you really meant was L3 VPNs are dead, that's not true either. If you are going to sell SD-WAN, learn technologies, learn the history.

22 35 117



Tim Hoffman
@hoffnz

Following

Replying to @danieldibswe

As the guy who led twitter's backbone team I can categorically tell you that your Twitter traffic is all in labels for the majority of their journey

10:35 AM - 22 Apr 2018 from San Francisco, CA

3 Retweets 16 Likes



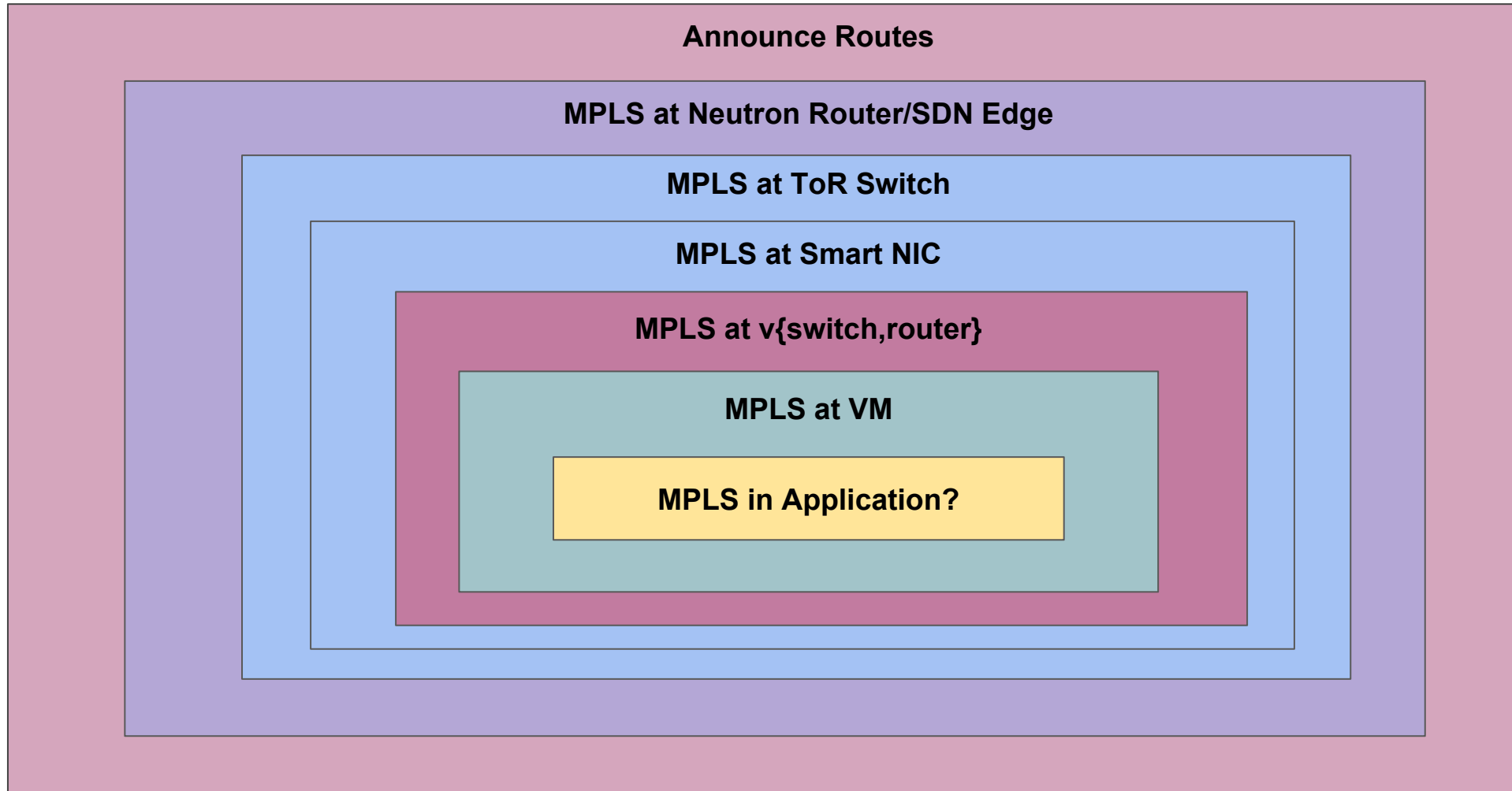
3 16



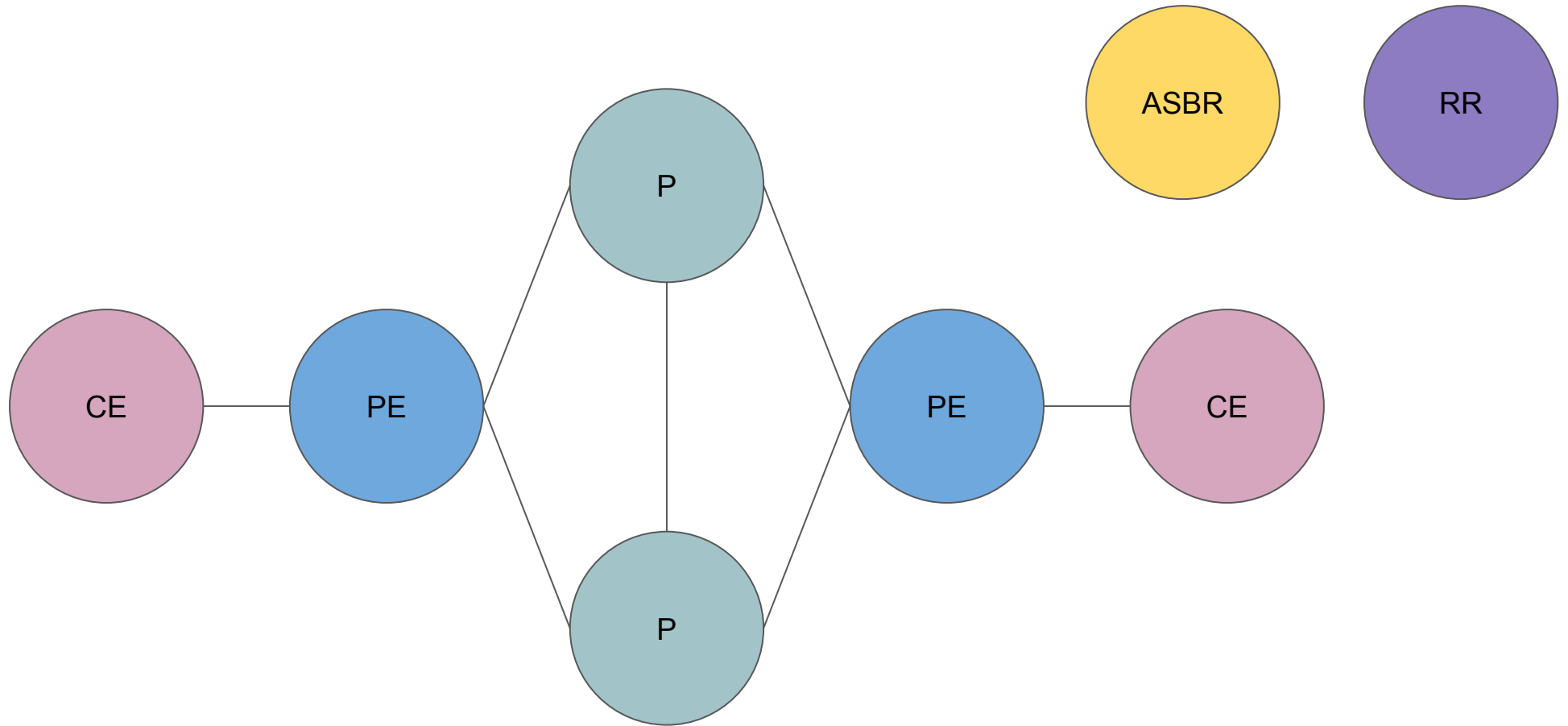
1. Do nothing - Standard use of OpenStack
 - a. Floating IPs
 - b. VLAN based, admin controlled provider networks
2. Neutron Dynamic Routing
3. Neutron BGP-VPN / BagPipe
4. Future - Neutron to Neutron Interconnection, others
5. SDN
 - a. Tungsten Fabric
 - b. OpenDaylight
 - c. Other...

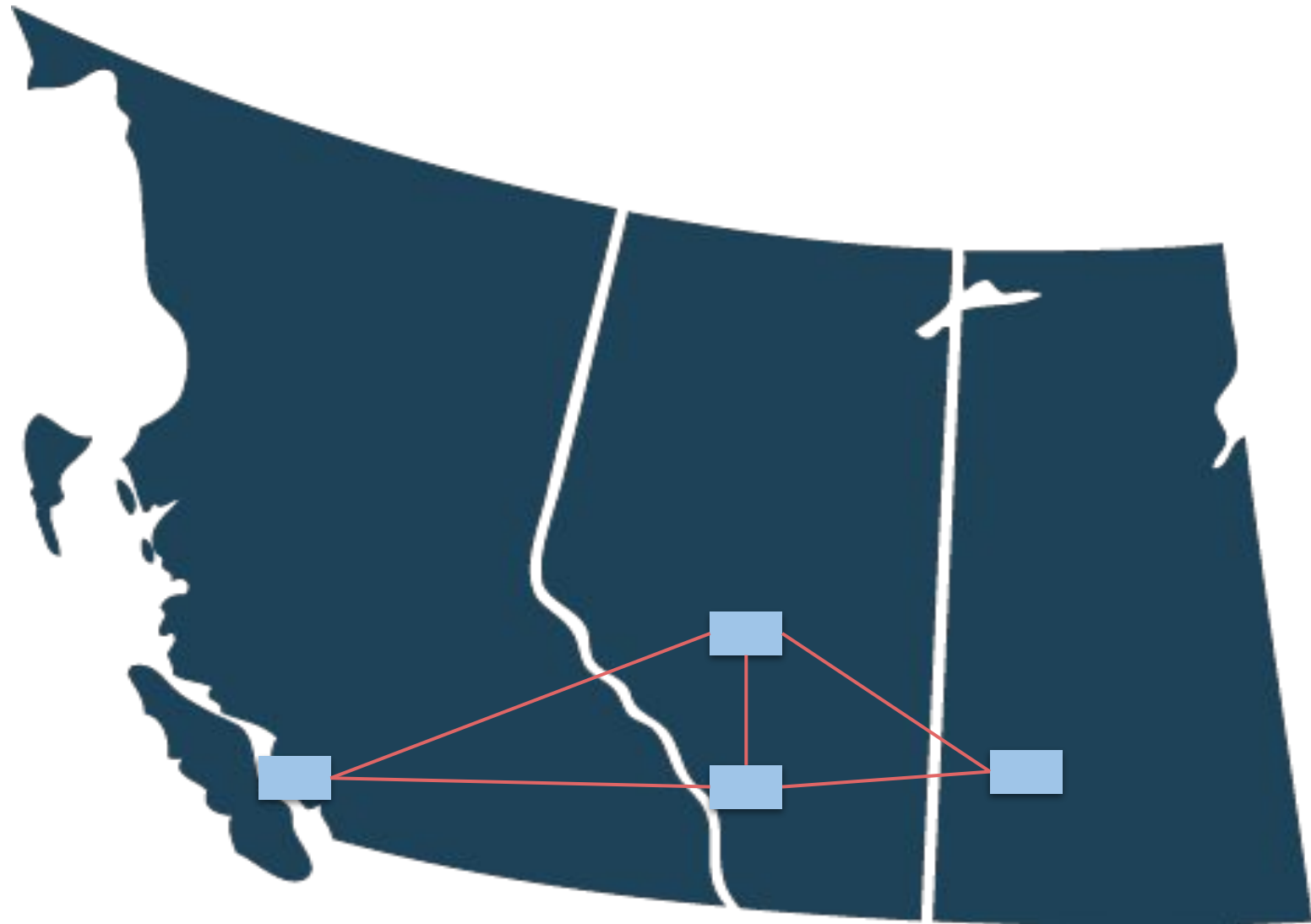
WHERE DOES MPLS LABELS GET APPLIED?

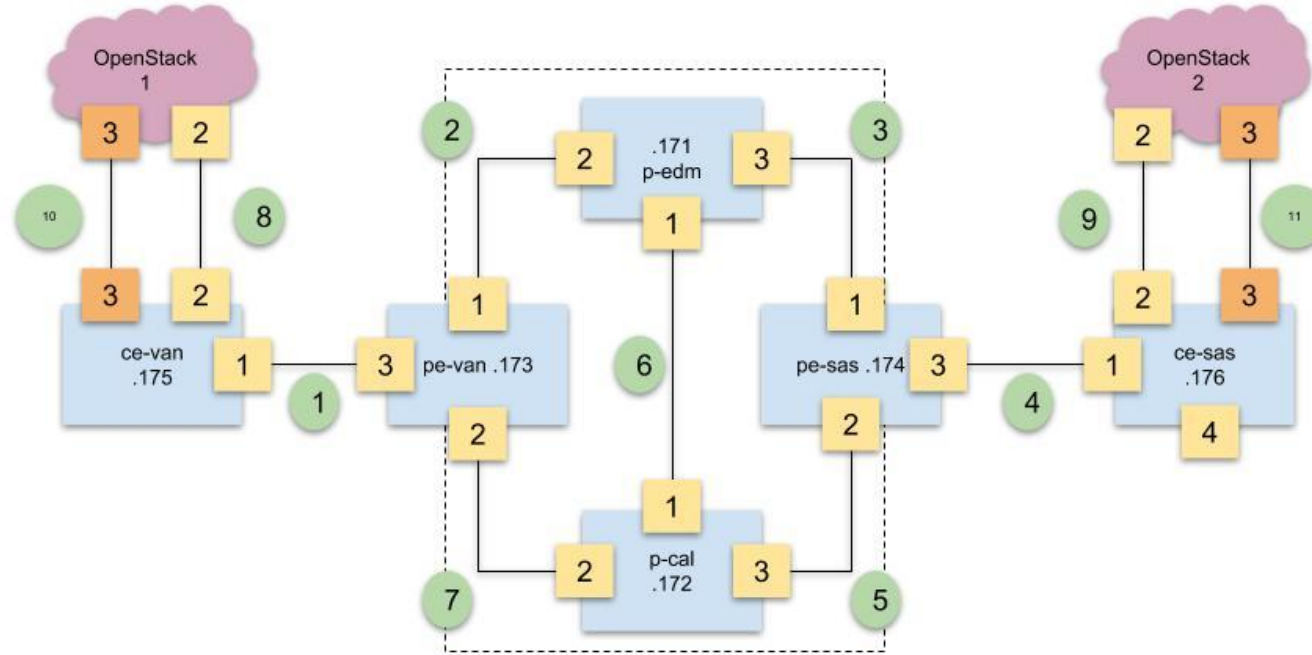
No Integration or Static



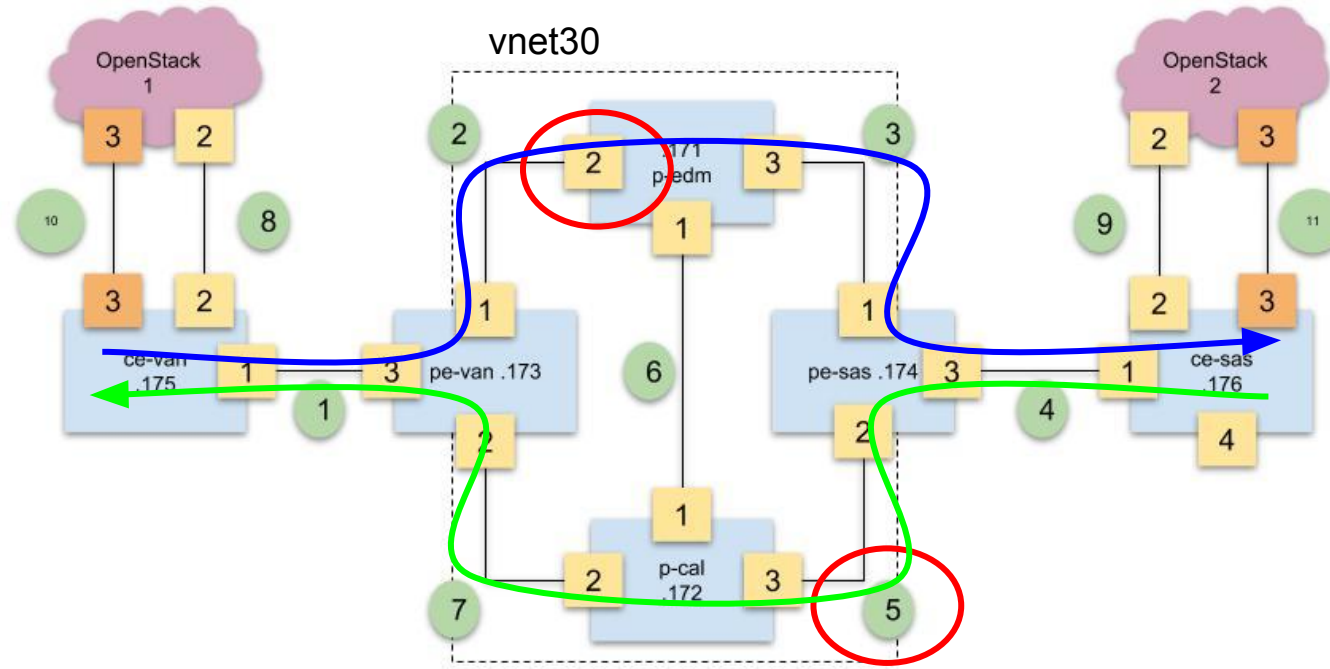
- “Just enough” MPLS to run our lab
- Based on [Simple Juniper MPLS Core with L3VPN blog post](#)
 - 2 provider routers
 - 2 provider edge routers
 - 2 customer edge routers
- OpenStack clouds connected at various points
- Meant to simulate a geographically distributed MPLS network
- But not 100% replicate a Service Provider network
- Added latency to interfaces with tc to simulate large distances
- Completely automated using Ansible, Netconf







- z = Control Plane BGP
- x = Link #
- y = Interface #
- = Router
- = OSPF Area



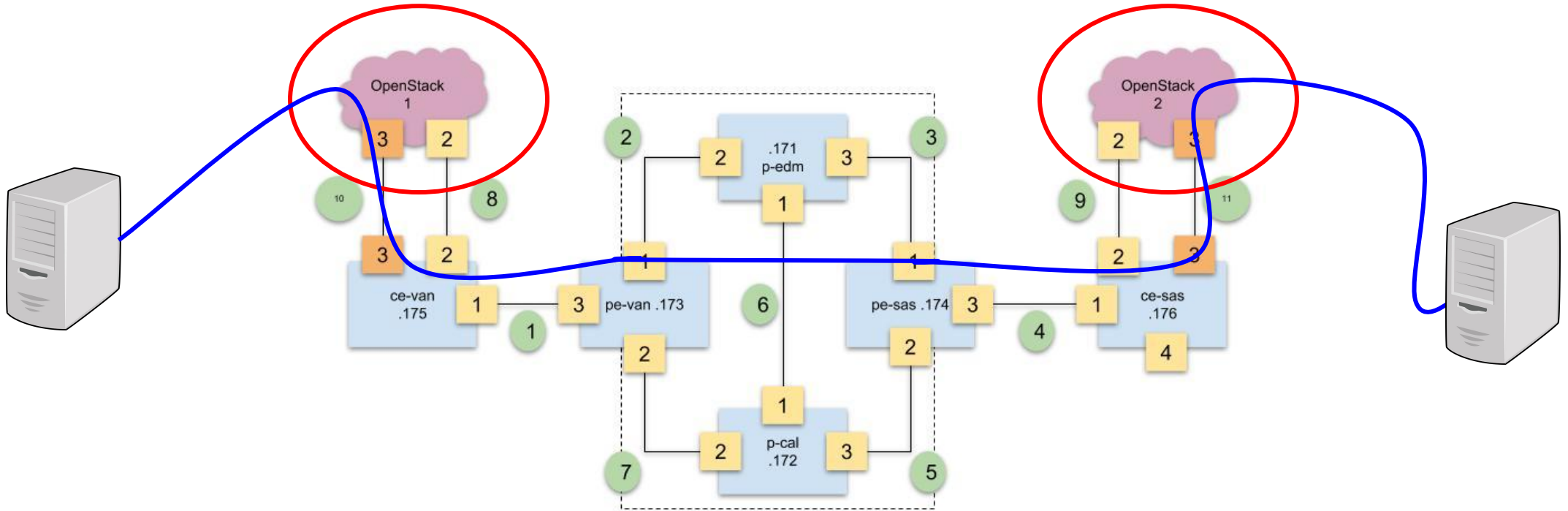
vnet18

- = Control Plane BGP
- = Link #
- = Interface #
- = Router
- = OSPF Area

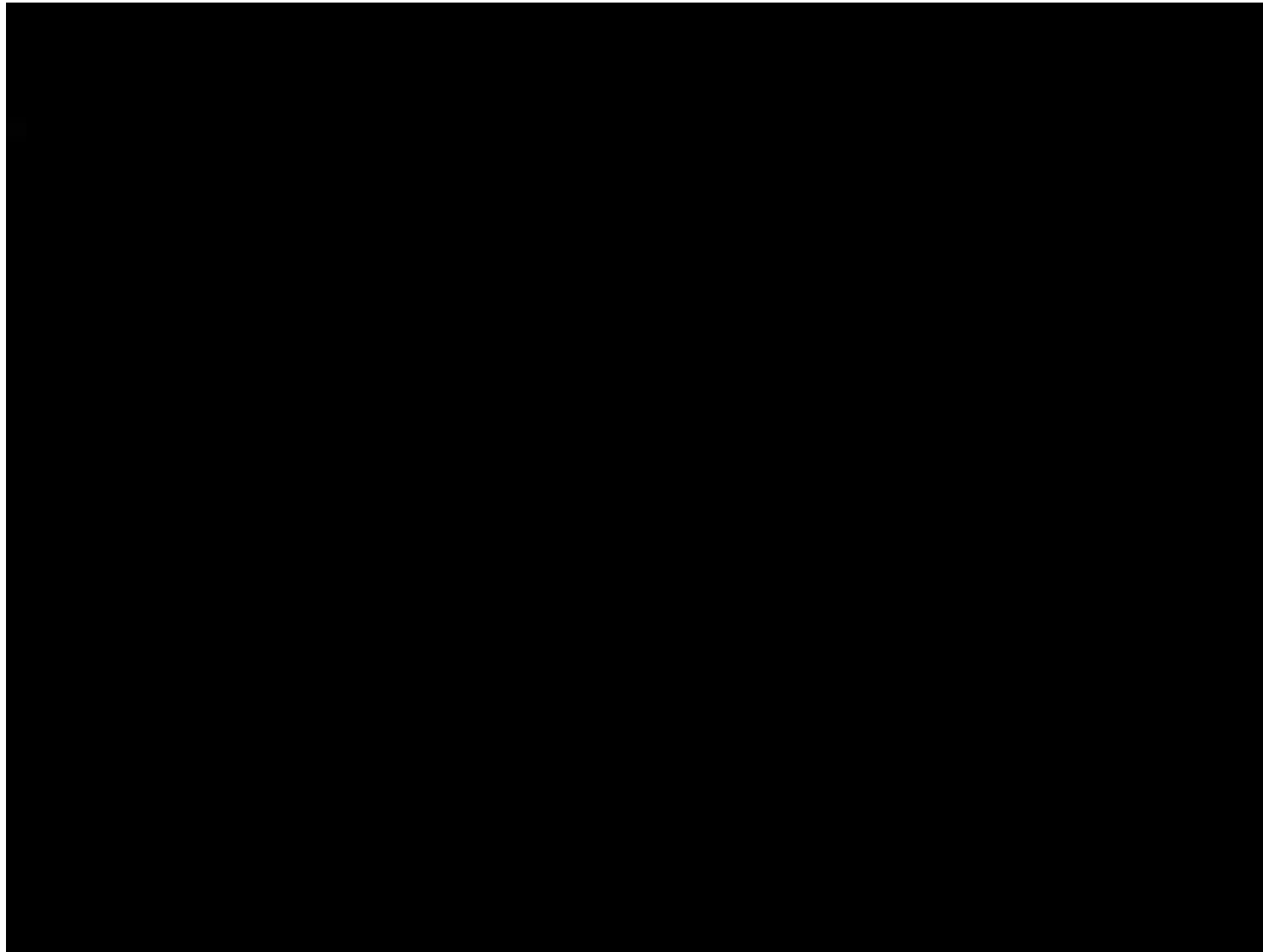
NEUTRON BGP DYNAMIC ROUTING

- Want ingress L3 access to tenant VXLANs
 - Next hop for inbound is not known by rest of infrastructure
- If we do static routing, all kinds of manual intervention required
- How to automate access to tenant networks and participate in a larger network?

- OpenStack deployments run Neutron Dynamic Routing
 - Basically Neutron gets a BGP speaker
- OpenStack announces routes to CE router, CE router peers with PE router
 - Or could be CE router...
- Value is that self service networks are accessible at L3 via Neutron routers immediately and automatically
- Clouds can be given particular networks they are allowed to announce routes in
- Not the same as BGP-VPN (will cover that later)
- Not directly integrated with MPLS (not meant to be at this time)



- z = Control Plane BGP
- x = Link #
- y = Interface #
- = Router
- = OSPF Area

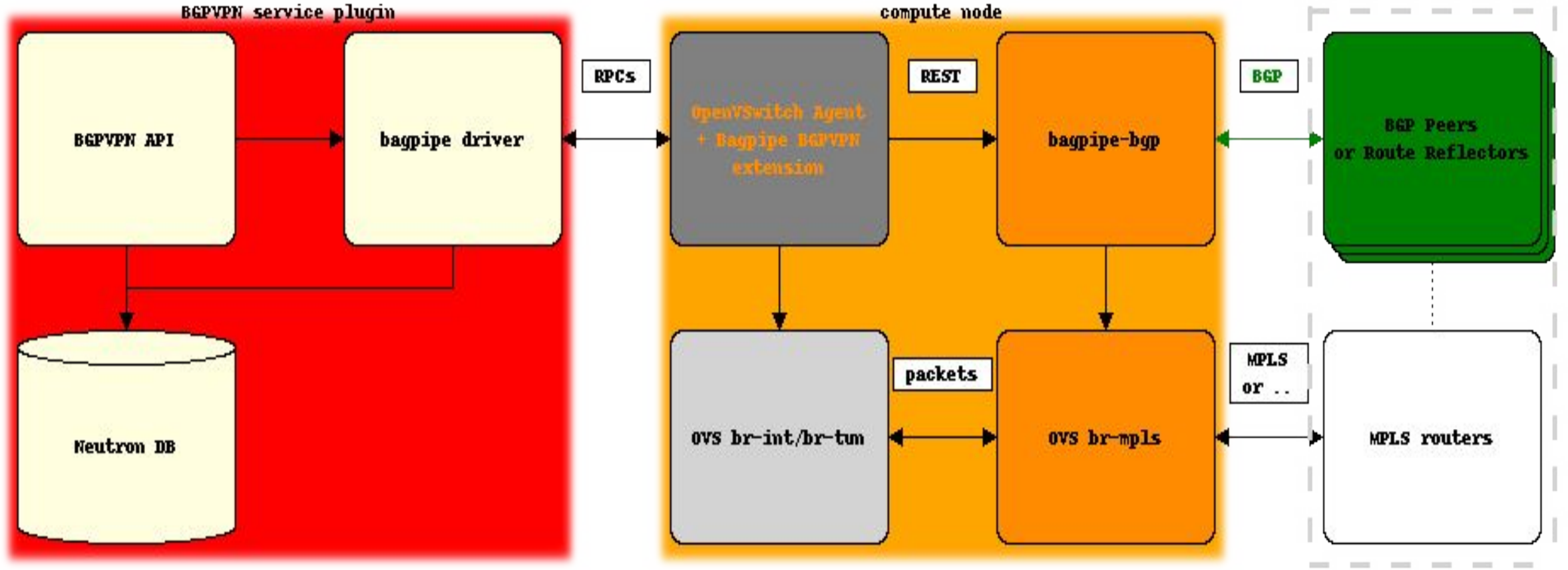


NEUTRON BGPVPN INTERCONNECTION

- “the goal...is to make it possible to **create the interconnect we need when the technology of BGP-based VPNs is already used outside an OpenStack cloud.**”
- “... the PE and VRF functions are implemented on compute nodes and the VMs are acting as CEs. This PE function will BGP-peer with edge IP/MPLS routers, BGP Route Reflectors or other PEs.”
- “Bagpipe BGP...**could also** be instantiated in network nodes, at the I3agent level, with a BGP speaker on each I3agent; **router namespaces could then be considered as CEs.**”

- [BGP VPN Interconnection Overview](#)

* Diagram borrowed from Openstack.org



- Related talks at this summit:
 - Flexible NFV WAN interconnections with Neutron BGP VPN
 - Thursday, May 24, 3:30pm-4:10pm, presented by Thomas Morin
 - Using Neutron BGP VPN for edge networking
- Previous talks/slides
 - OpenStack Neutron Interconnections with BGP-MPLS VPN
 - Neutron Software Define Interconnects with WAN BGP VPNs
 - Interconnecting Neutron and Network Operators BGP VPNs
- Standard BGP extensions to build multi-site L2 or L3 virtual networks

NEUTRON TO NEUTRON INTERCONNECT

How to interconnect two OpenStack deployments ? (..... two or more OpenStack ? two regions ?)

Between datacenters or between NFV POPs,
you may want network interconnections with
the following properties:

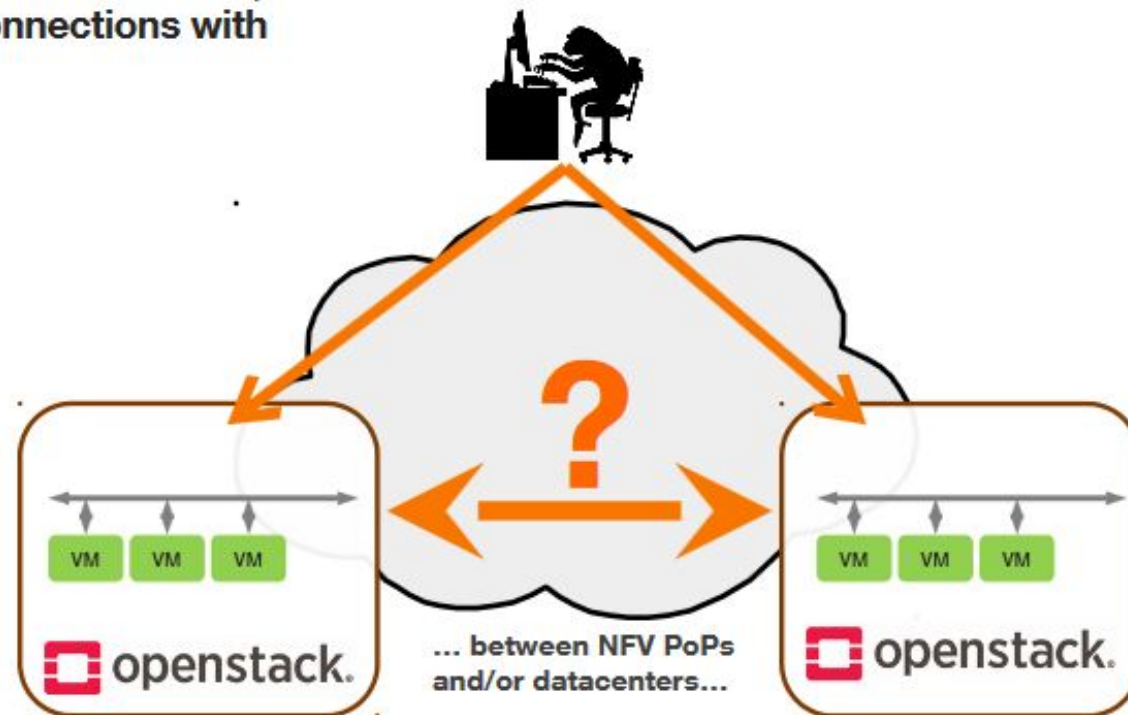
on demand



private addressing
& isolation



avoid the overhead of
packet encryption



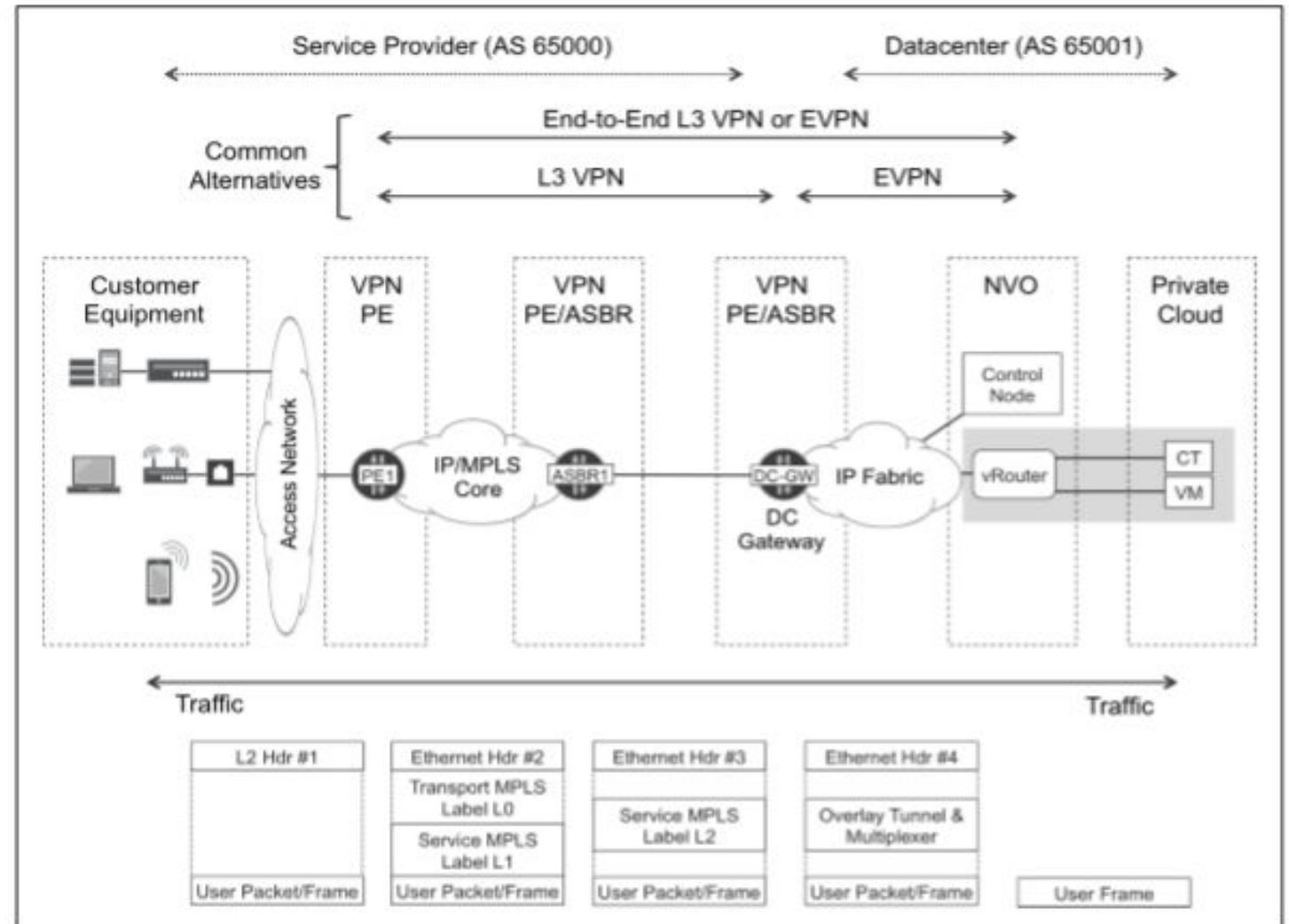
Source: <https://www.openstack.org/assets/presentation-media/20171107-OpenStackSummit-Sydney-NeutronNeutronX-02.pdf>

- Neutron to Neutron Interconnect Spec
 - <http://logs.openstack.org/26/545826/15/check/build-openstack-sphinx-docs/1107e34/html/specs/rocky/neutron-inter.html>
- With existing tools we can connect routers and networks to existing BGP VPNs, but requires some external orchestration and configuration
- What if we want to connect isolated networks on-demand?
 - Isolated
 - Private addressing
 - Not using encrypted VPNs
- Neutron to Neutron Interconnect Spec

SDN RELATED SOLUTIONS

- Many different SDN solutions
- SDN by definition means networking is implemented in software, and thus architectural and implementation decisions are made in that software
- To discuss all solutions would be to discuss...many implementations
- Examples could be:
 - Tungsten Fabric
 - OpenDaylight
 - Calico
 - Many others...
- If Neutron does not have a particular API available, then must use SDN's API

- TF heavily leverages MPLS and XMPP and BGP
- “VMs and containers are like CEs”
- “vRouters are like PEs”
- “Like ASBR1, the data center gateway is an **Option B** Autonomous System Border Router (ASBR)”



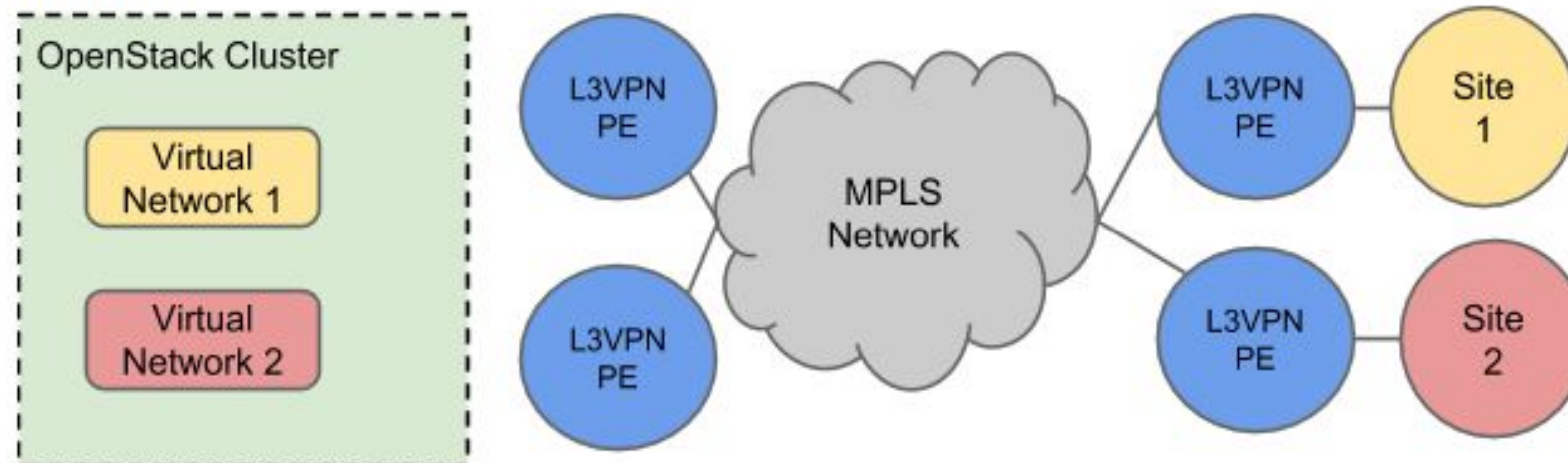
Source: MPLS in the SDN Era

Figure 11-2. Subscriber access to a private cloud

OTHER NEUTRON SPECS AND PROJECTS

- Spec - “Connect Neutron Virtual Network with Existing VPN with routes being dynamically exchanged via BGP. This scenario assumes the use of inter-AS option B) between the OpenStack cluster and the external BGP/MPLS L3VPN network.” - [Neutron BGP MPLS](#)
- RFC 4364 (same as neutron-bgpvpn)

Use case 1:

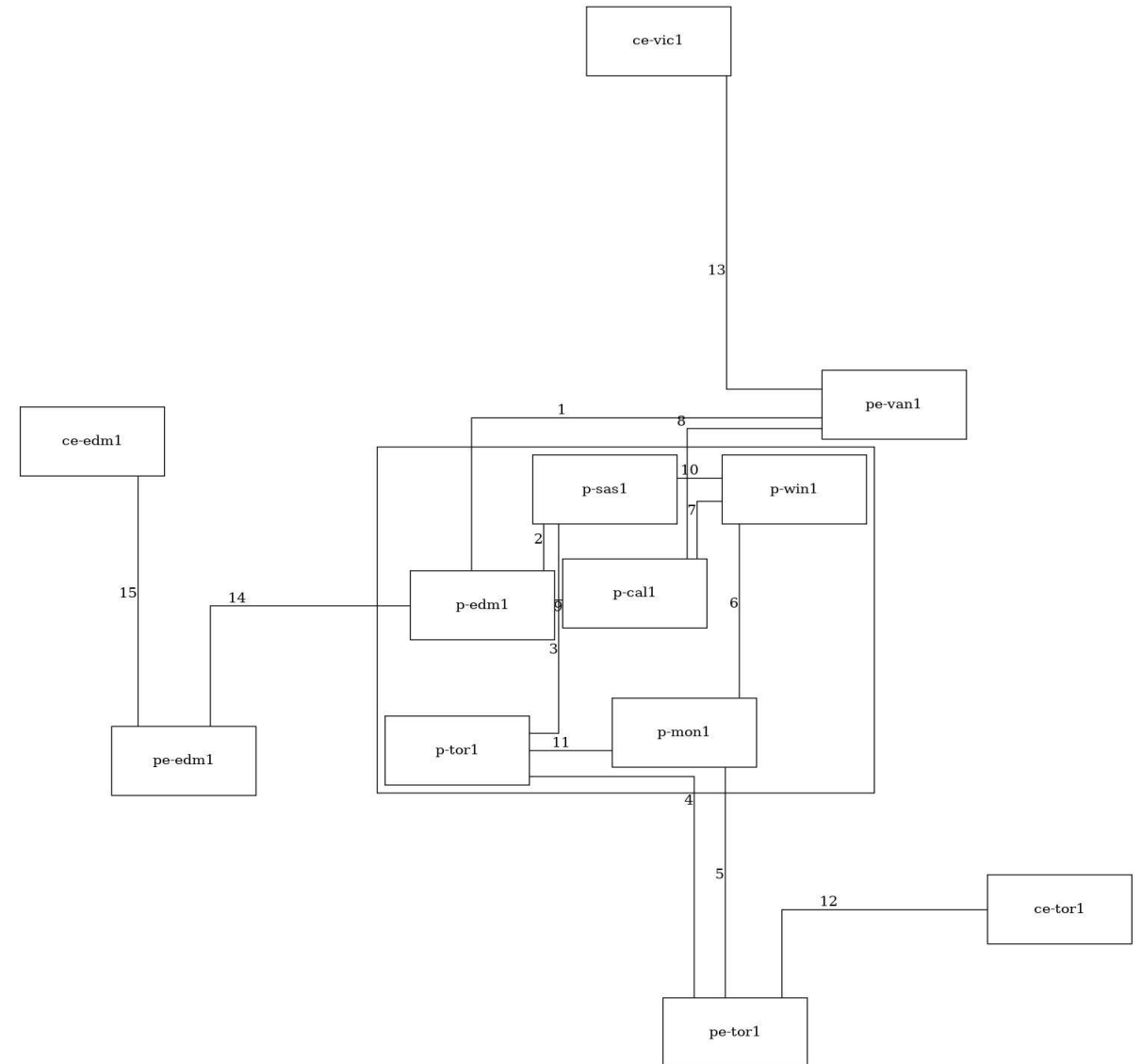


- Spec - Neutron Edge VPN
 - “This blueprint defines APIs to provision edge VPNs for MPLS VPN use case. MPLS enables the deployment of L2 and L3 VPNs between data centers, with QoS guarantees, to provide inter-datacenter connectivity. With these capabilities, **we can achieve WAN orchestration where tenant VMs located in different datacenters can communicate transparently over an MPLS network.**” - [Neutron BGP MPLS](#)
- Project - Tricircle
 - “The purpose of the Tricircle project is to provide networking automation across Neutron servers in multi-region OpenStack clouds deployment.”

LESSONS LEARNED

- vSRX can be up to 30 minutes before completely available
 - Could be easy wins to reduce time; haven't investigated fully
- Getting all in one deployments of DevStack, Tungsten Fabric, etc, not easy to maintain over time
- Automating everything allows us to move to new hardware easily without having to do manual work, rename things, deploy different topologies, etc
- Quite a few todos:
 - Schedule networks instead of hardcoding via Ansible inventory
 - Move to YANG

- Link based inventory -- First we define all links in YAML, then custom Ansible inventory is based off of that
- Also working on automatic generation of network diagrams based on links and pygraphviz

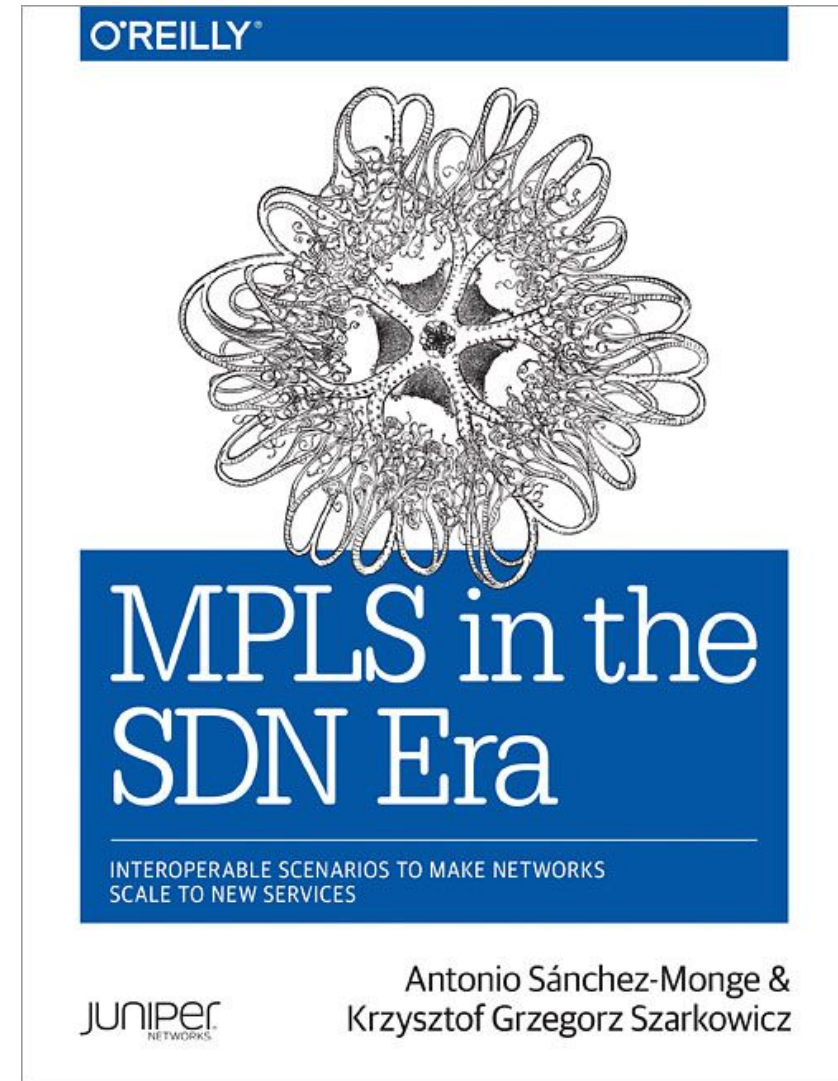


- Had some issues keeping bgp connection alive, occasionally had to recreate speakers
 - If the bgp dragent was restarted, bgp connectivity was lost
 - Not much investigation done here yet, but are wondering how to keep this highly available outside of a PoC
- Still need some manual architecture work done in terms of what networks clouds are allowed to announce
 - This would be determined by a network ops team outside of OpenStack
 - Eg. specified subnet pools: openstack-1: 10.0.1.0/16, openstack-2: 10.0.2.0/16

- Based on iBGP
- Major use case is for connecting Neutron routers and networks with existing VPNs
- When connecting networks/ports, each compute node essentially be a PE router
- Need to test multiple compute based infrastructure; deployments were “all in one” instances

END MATTER

- “MPLS in the SDN Era” is quite an astounding piece of work
 - Antonio Sanchez-Monge and Krzysztof Grzegorz Szarkowicz



- Networking has a long history of standards and interop; lots to learn
- CE, P, and PE router definitions may not be granular enough for MPLS in NFV/SDN era
- MPLS integration is possible at many points
- Inter-AS vs single operator
- Options for interconnecting multiple OpenStack clouds are currently few
- Stepping outside of Neutron APIs is a slippery slope for interop

- Challenging to keep a dev lab up and running (complex systems moving quickly)
- Intersection of high-end networking and OpenStack: large technical domain, especially when including SDN
- Ultimately, network integration of multiple clouds is the act of automating (ie. providing OpenStack APIs for) existing, well known technologies like MPLS and BGP VPNs





- Thanks to:
 - Thomas Morin for help with BGP-VPN/Bagpipe
- Contact:
 - curtis@interdynamix.com
 - rahul@interdynamix.com
- Questions?

Interdynamix Edmonton
620 Manulife Place
10180 – 101 Street NW
Edmonton, AB T5J 3S4
780.423.7005

Interdynamix Toronto
140 Yonge St.
Suite 200
Toronto, ON M5C 1X6