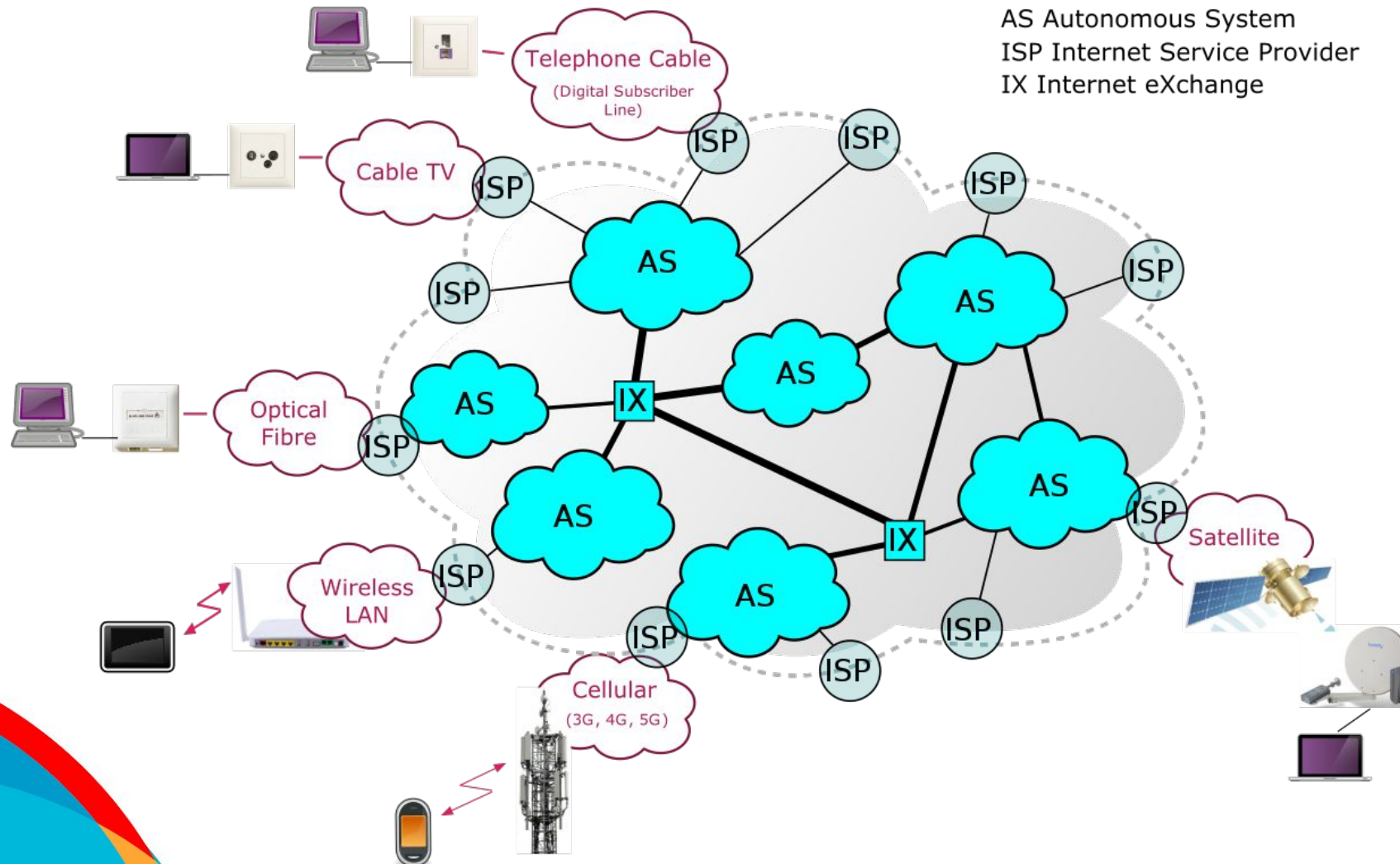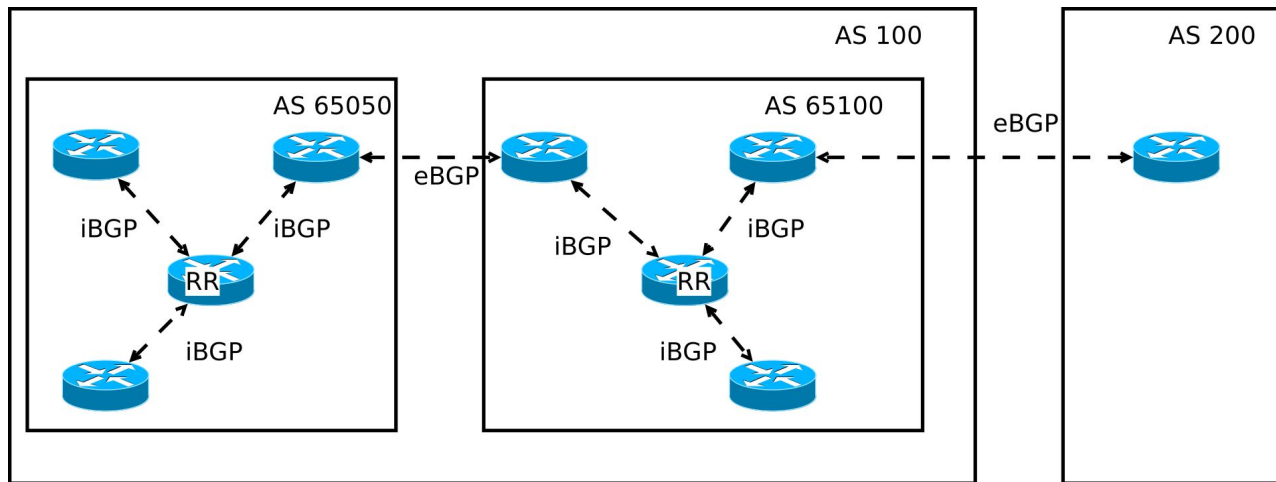# OVN BGP Agent

Bring BGP in your OpenStack cloud

Bernard Cafarelli <bcafarel@redhat.com>

What is BGP?

What is BGP?

- Allows exchange of routing information (tcp/179)
- Autonomous System (AS) is a collection of IP routing prefixes under the control of a single administrative entity.
- Exterior (eBGP) and Interior (iBGP).
- Standard for Internet routing, most ISPs require it to establish routing between one another.
- Widely used in large private IP networks

BGP for OpenStack

Layer 2 deployments

Typically data centers route packets at upper layers of the topology, with the lower layers operating as a L2 network.

Issues
- Large failure domains
- Large broadcast domains (broadcast, multicast, anycast)
- More static configuration
- Scalability issues
- Slow convergence time (Gratuitous ARPs, MAC learning, STP)
- TCAM exhaustion

Network Protocols
- ARP
- STP
- VLAN
- LACP Bonding
- M-LAG
- vPC

Red Hat

BGP for OpenStack

Going to Layer 3

L3 data centers route packets at the lowest levels of the hierarchy (ToR and on the worker nodes).

Resolves Issues
- Smaller L2 failure domains
- Smaller L2 broadcast domains
- Dynamic configuration
- Faster convergence (<1 ms)

Network Protocols
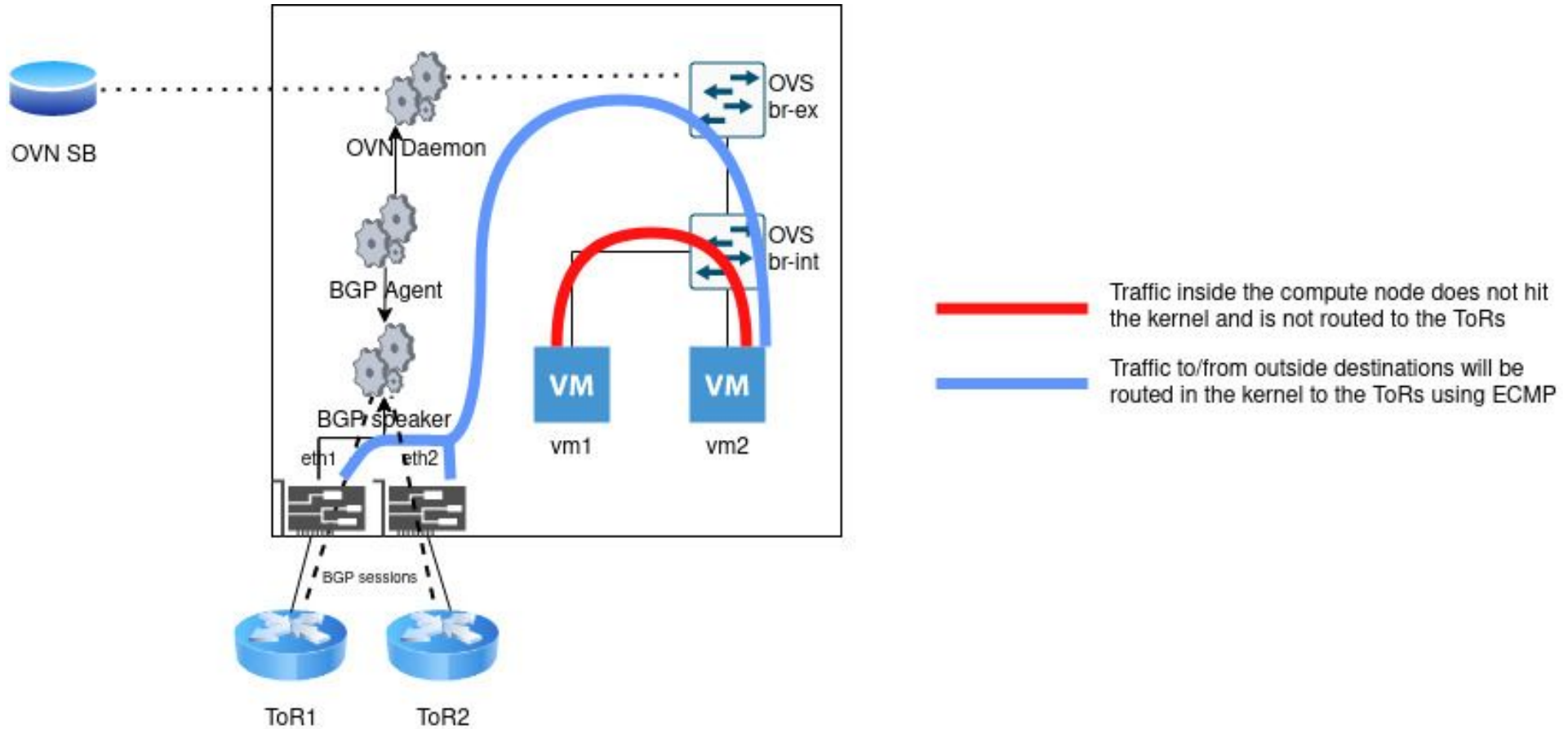- Routing Protocol (e.g. BGP)
- ECMP
- BFD

Red Hat

Design

Components

- Python based daemon running on each node

  - Monitors the **OVN database** to catch the events to trigger the actions

  - Relies on **FRRouting (FRR)** for BGP protocol implementation

  - Multiple exposing methods to route the traffic between the OVN overlay and the physical network , leveraging either:

    - **kernel networking** capabilities

    - a node-local **OVN cluster**

- Exposed IP addresses

  - VMs and load balancers on provider networks (or associated Floating IPs)

  - Optionally, VMs on tenant networks (can be controlled by address scopes)

  - And more (IPv6 GUA, advertise tenant networks per VM or per subnet, …)

Design

Architecture

Traffic inside the compute node does not hit the kernel and is not routed to the ToRs

Traffic to/from outside destinations will be routed in the kernel to the ToRs using ECMP

Design

Traffic routing (kernel routing)

- Traffic between nodes/BGP Advertisement

  - Actions to trigger FRR to advertise the route

- Traffic within a node/Redirecting traffic to OVN overlay

  - Set of IP rules, IP routes (on separate routing tables) and ARP entries

- Traffic within a node/Redirecting traffic from OVN overlay

  - Kernel sysctl knobs for ARP/NDP

  - Extra flow on ovs provider bridges to rewrite destination MAC address

- No support for overlapping CIDRs (avoid with address scopes, subnet pools, …)

Design

Traffic routing (OVN cluster)

- Traffic between nodes/BGP Advertisement
  - Actions to trigger FRR to advertise the route
- Traffic within a node/Redirecting traffic to or from the OVN overlay
  - Local OVN cluster on node
- Limitations
  - Requires OVN>=23.09
  - No BFD support (for NIC failover)
  - IPv6 support
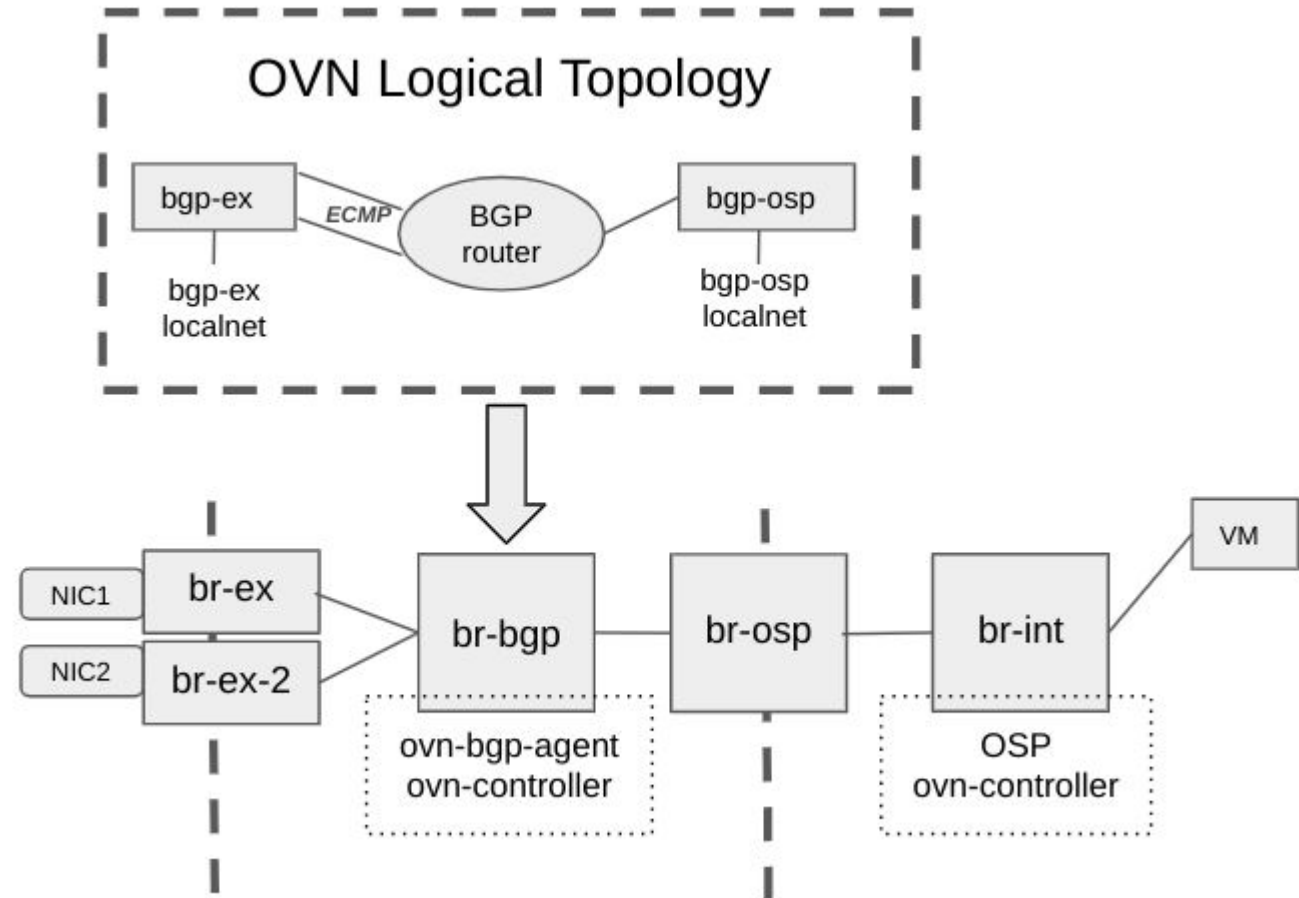
Improvements

- First implemented exposing method uses additional components and settings beyond OVN (IP rules, routes, ARP entries)

- Now supported natively in OVN via the local cluster per node
  - Less internal hops
  - This also allows to add support for hardware acceleration/DPDK

- Switch from OVN southbound database to northbound database (better scalability, more stable API)

- And more coming (initial support for L3 EVPN type-5 routes, …)

Project history so far

- Upstream milestones
  - August 2021: repository creation
  - February 2022: version 0.1
  - October 2022: new contributors and drivers
  - April 2023: project accepted in Neutron stadium
  - May 2023: version 0.4, integrated in OpenStack release cycle, initial support for local OVN cluster
- Red Hat OpenStack Platform
  - Became fully supported in version 17.1 (based on Wallaby, kernel routing)
  - Upcoming version 18 (agent based on 2023.2/Bobcat)

Red Hat

Upcoming steps

- Performance testing on NFV use-cases and exposing methods comparison

- L3 EVPN driver <u>recently merged</u>

- API efforts <u>in progress</u>

- OVN community exploring tighter integration between OVN and BGP, discussed in latest <u>community meeting</u>

Red Hat

Beyond the agent and dataplane

- Scope and design differences with neutron-dynamic-routing
  - Focus on OVN and pure L3 datacenter architecture
  - Full-feature FRR daemon vs dedicated specific BGP speaker
  - Support for provider network IP addresses
  - Support for EVPN
  - neutron-dynamic-routing caveats with DVR and ML2/OVN
- What about the control plane?
  - Desire for controllers in separate racks/datacenters
  - More static configuration, stretched L2 domain can work
  - Depends on your deployment tool - ask your favourite vendor

Red Hat

Resources

- [Project page](#) and Launchpad [bug reporting](#)
- Main [documentation](#) (including extensive contributor documentation)
- Luis Tomás Bolivar's [blog](#) has multiple detailed articles
- Red Hat OSP 17.1 [documentation](#) on dynamic routing (search "OSP 17.1 BGP")

Red Hat

Questions?

- [openstack-discuss](#) mailing list
- *#openstack-neutron* (OFTC network) IRC channel

Red Hat