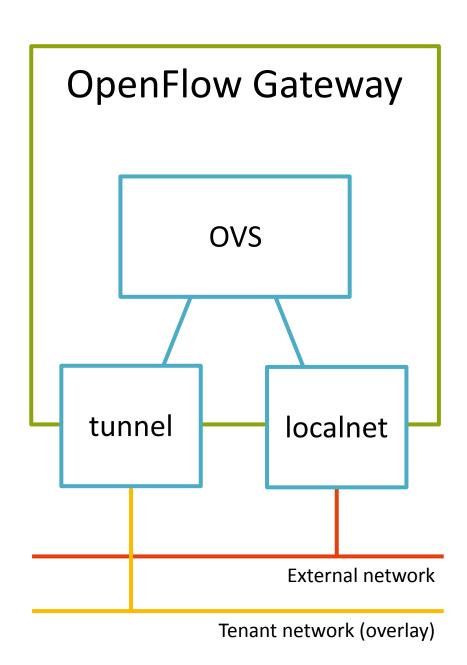
NUTANIX. External Networks on the Overlay

December 6, 2018 Gregory Smith



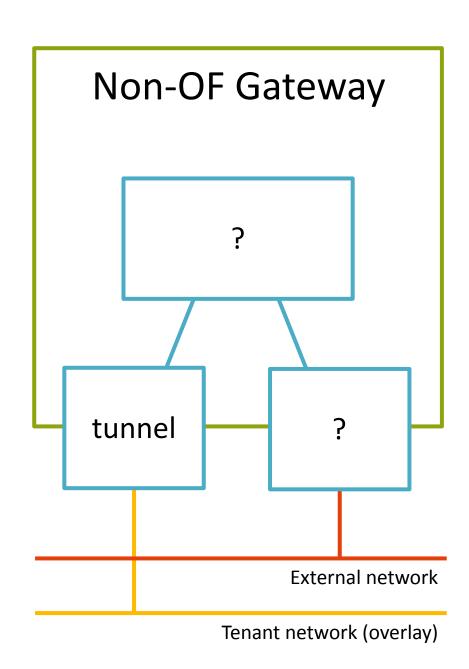
> Typical External Connectivity



- Gateway chassis physically connected to external network via localnet port.
- OVN programs NAT rules into OVS via openflow for floating IPs and SNAT.



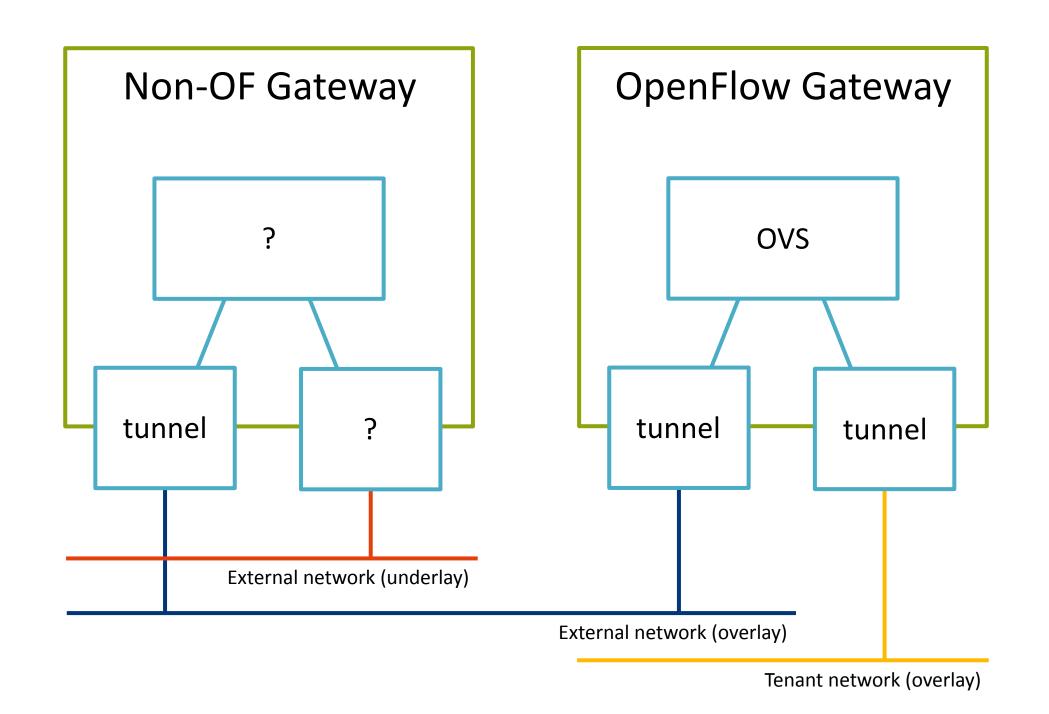
Non-OpenFlow Gateway



- Program gateway through netconf instead of OVN/OpenFlow.
- Keep configuration simple and static no NAT.

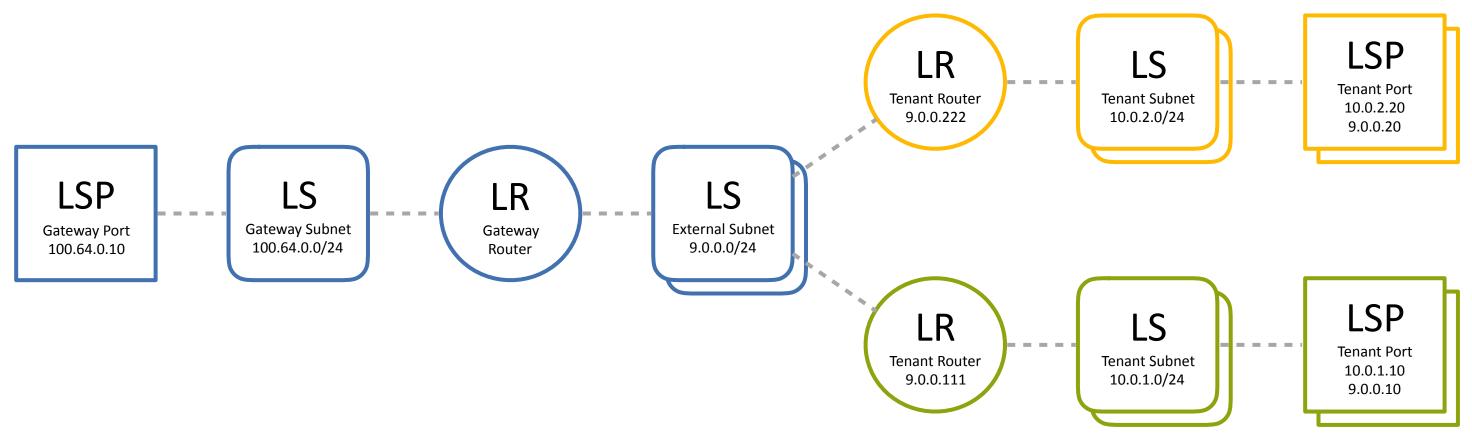


> Splitting The Gateway





Logical Topology



Legend

LR: Logical router

LS: Logical switch

LSP: Logical switch port

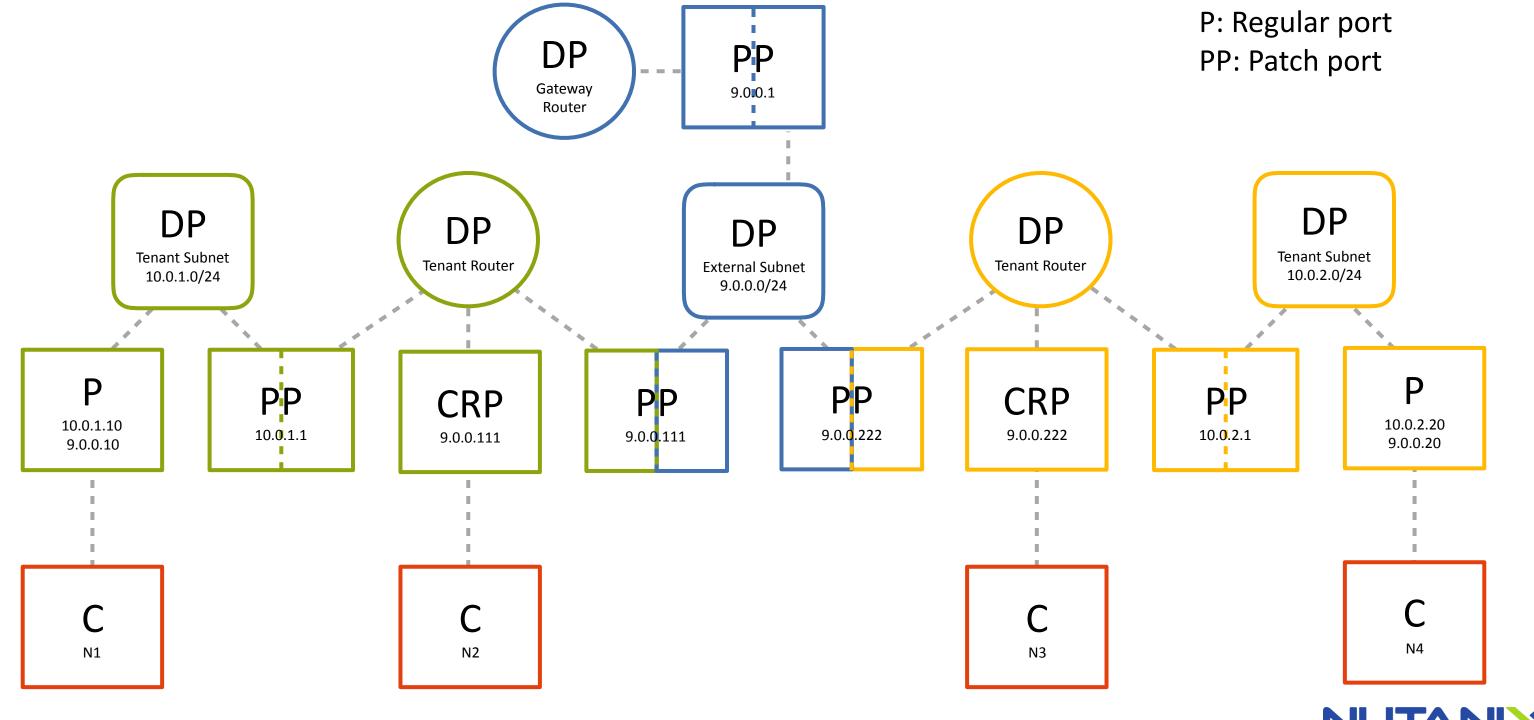


> Datapath Topology

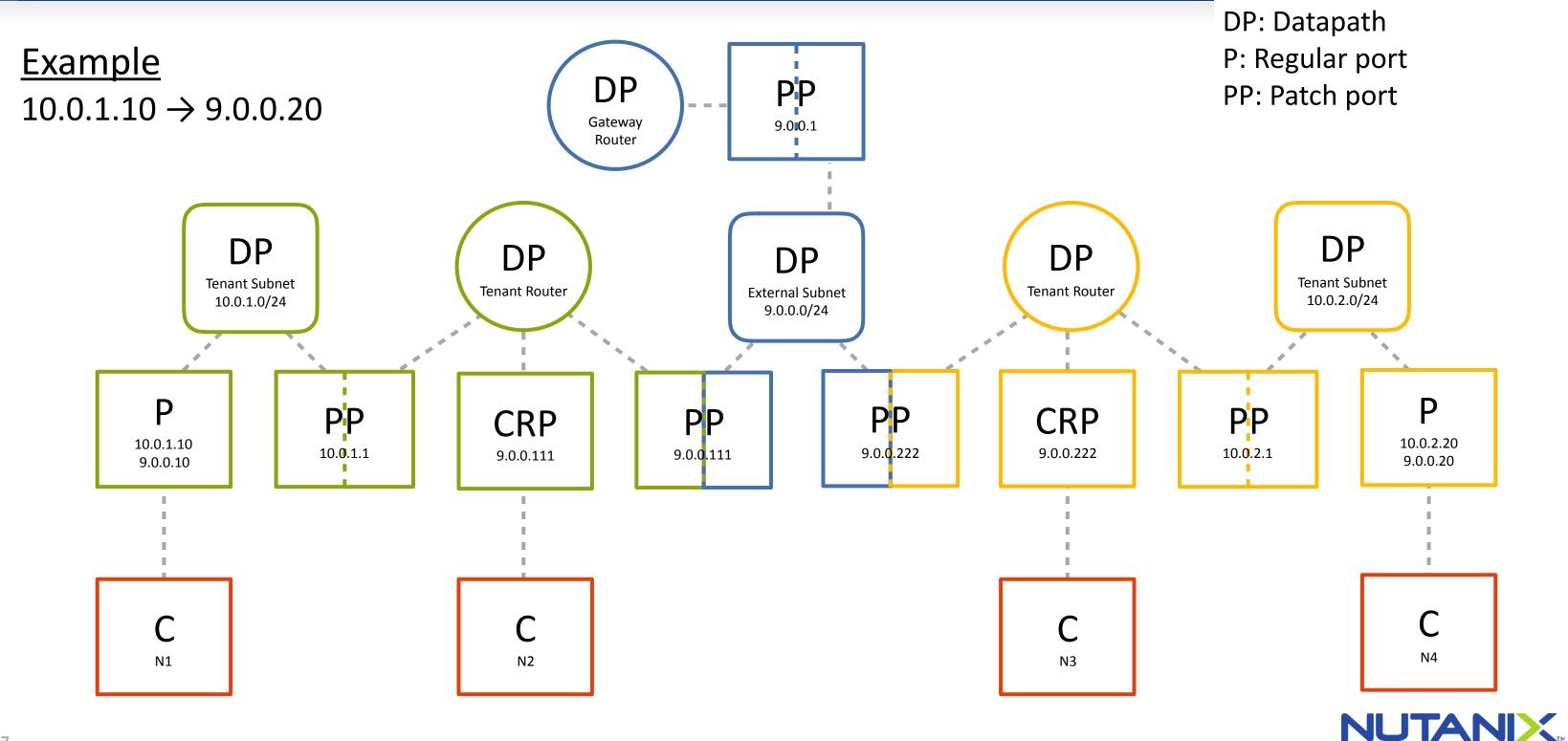


C: Chassis

CRP: Chassis redirect port



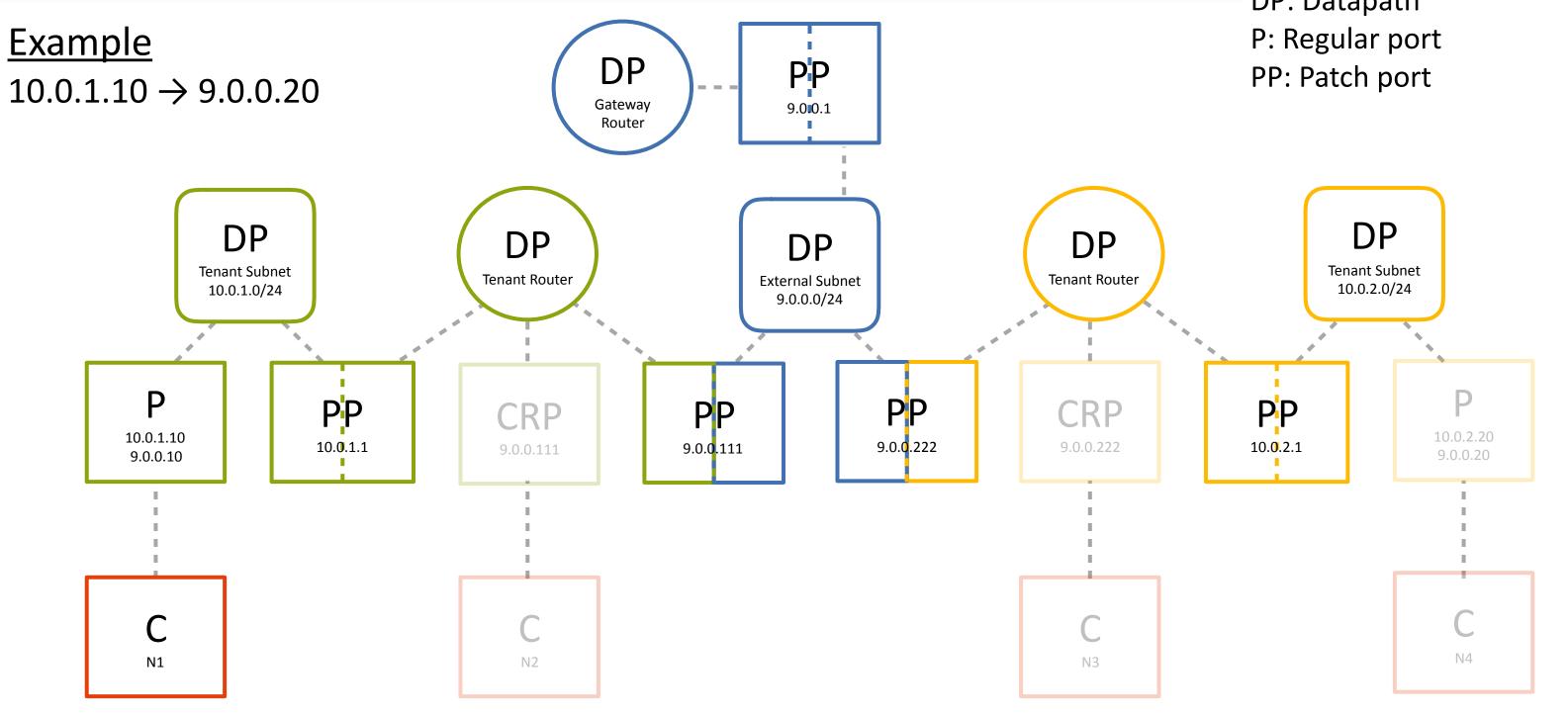




Legend

C: Chassis

CRP: Chassis redirect port



C: Chassis

CRP: Chassis redirect port



Example P: Regular port DP PP PP: Patch port $10.0.1.10 \rightarrow 9.0.0.20$ Gateway 9.0.0.1 Router DP DP DP DP DP **Tenant Subnet** Tenant Subnet Tenant Router **External Subnet** Tenant Router 10.0.1.0/24 10.0.2.0/24 9.0.0.0/24 P PP PP **CRP** PP PP CRP 10.0.2.20 10.0.1.10 9.0.0.222 10.0.1.1 9.0.0.111 9.0.0.222 10.0.2.1 9.0.0.111 9.0.0.20 9.0.0.10 N1 N2 N3

Legend

C: Chassis

CRP: Chassis redirect port



P: Regular port **Example** DP PP PP: Patch port $10.0.1.10 \rightarrow 9.0.0.20$ Gateway 9.0.0.1 Router DP DP DP DP DP **Tenant Subnet** Tenant Subnet Tenant Router **External Subnet** Tenant Router 10.0.1.0/24 10.0.2.0/24 9.0.0.0/24 who has 9.0.0.20? P PP PP **CRP** PP PP 10.0.2.20 10.0.1.10 9.0.0.222 10.0.1.1 9.0.0.222 10.0.2.1 9.0.0.111 9.0.0.111 9.0.0.20 9.0.0.10 N1 N2 N3

Legend

C: Chassis

CRP: Chassis redirect port



P: Regular port **Example** DP PP PP: Patch port $10.0.1.10 \rightarrow 9.0.0.20$ Gateway 9.0.0.1 Router DP DP DP DP DP **Tenant Subnet** Tenant Subnet Tenant Router **External Subnet** Tenant Router 10.0.1.0/24 10.0.2.0/24 9.0.0.0/24 who has 9.0.0.20? P PP PP CRP PP PP 10.0.2.20 10.0.1.10 9.0.0.222 10.0.1.1 9.0.0.222 10.0.2.1 9.0.0.111 9.0.0.111 9.0.0.20 9.0.0.10 N1 N2 N3

<u>Legend</u>

C: Chassis

CRP: Chassis redirect port



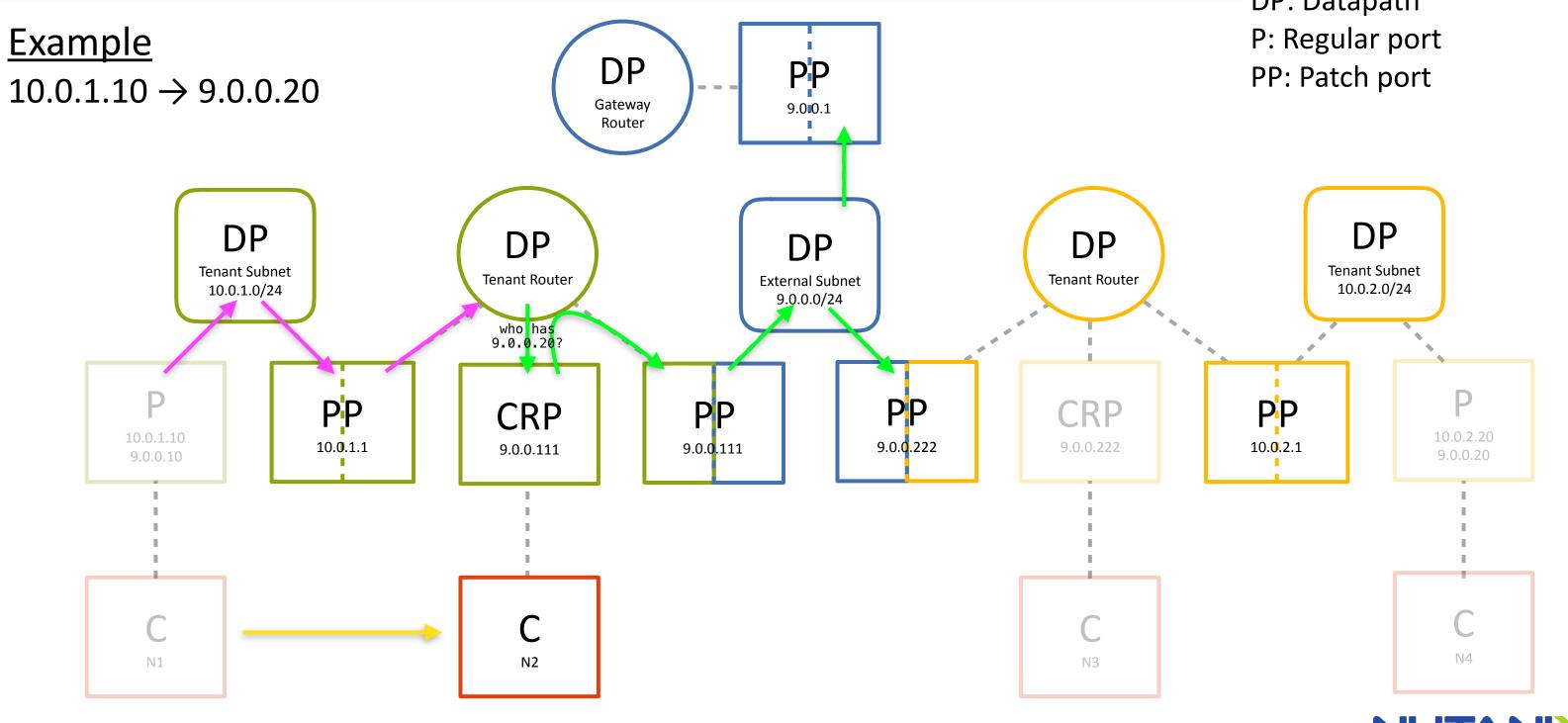
P: Regular port **Example** DP PP PP: Patch port $10.0.1.10 \rightarrow 9.0.0.20$ Gateway 9.0.0.1 Router DP DP DP DP DP **Tenant Subnet** Tenant Subnet Tenant Router **External Subnet** Tenant Router 10.0.1.0/24 10.0.2.0/24 9.0.0.0/24 who has 9.0.0.20? PP PP CRP PP PP **CRP** 10.0.2.20 10.0.1.10 9.0.0.222 10.0.1.1 9.0.0.111 9.0.0.222 10.0.2.1 9.0.0.111 9.0.0.10 9.0.0.20 N1 N2 N3



C: Chassis

CRP: Chassis redirect port



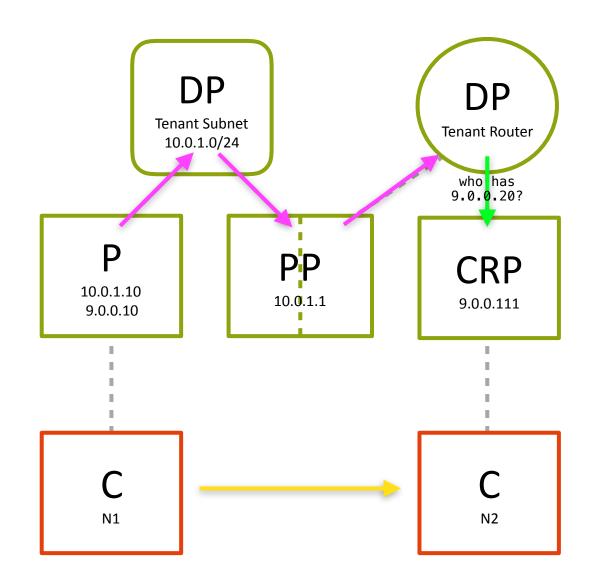


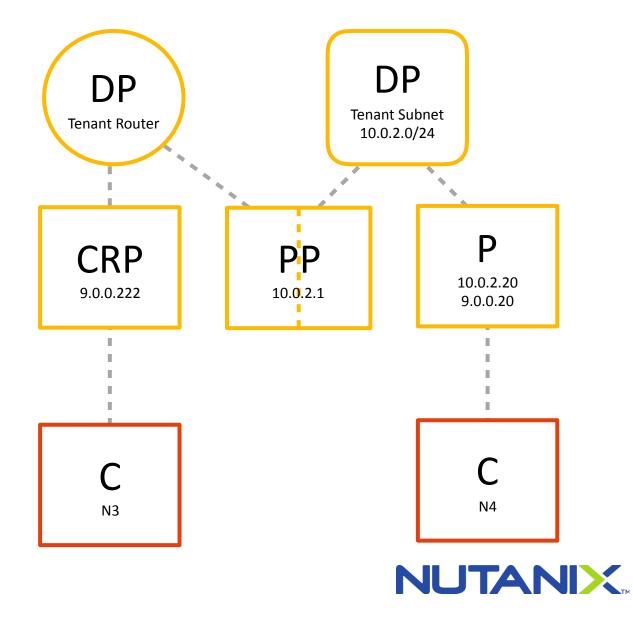
C: Chassis

CRP: Chassis redirect port

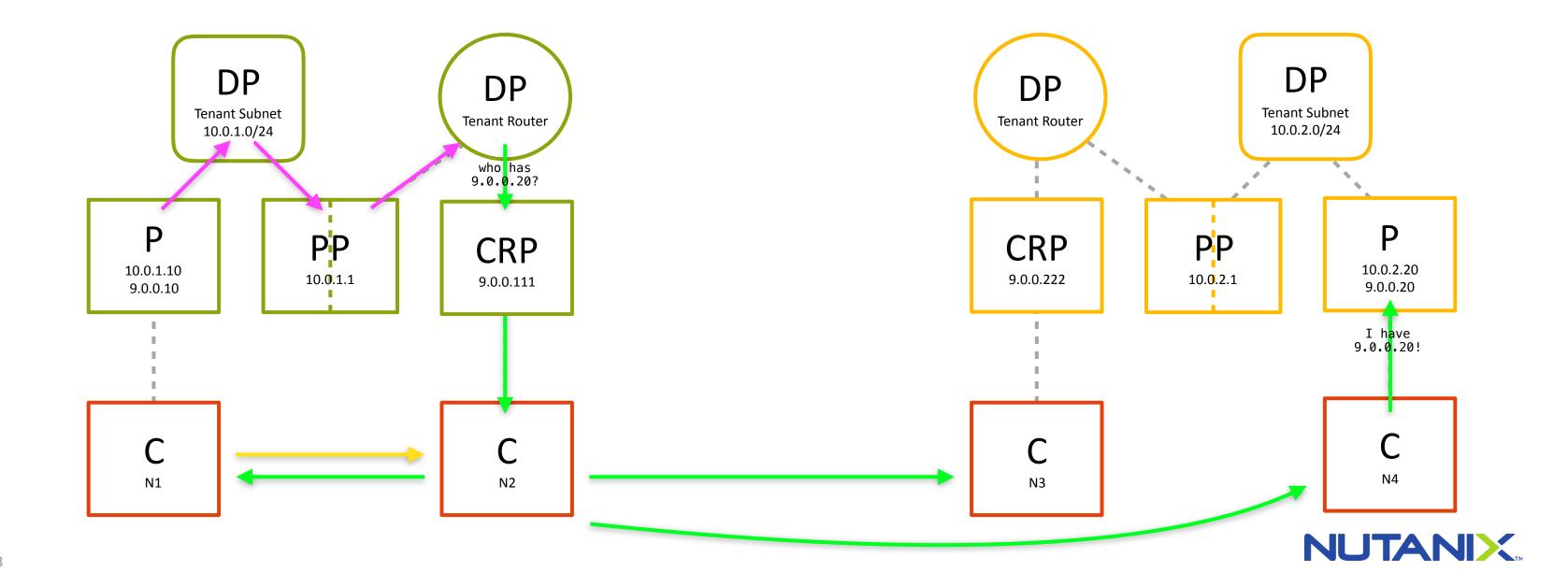


> Reference Architecture Example





> Reference Architecture Example

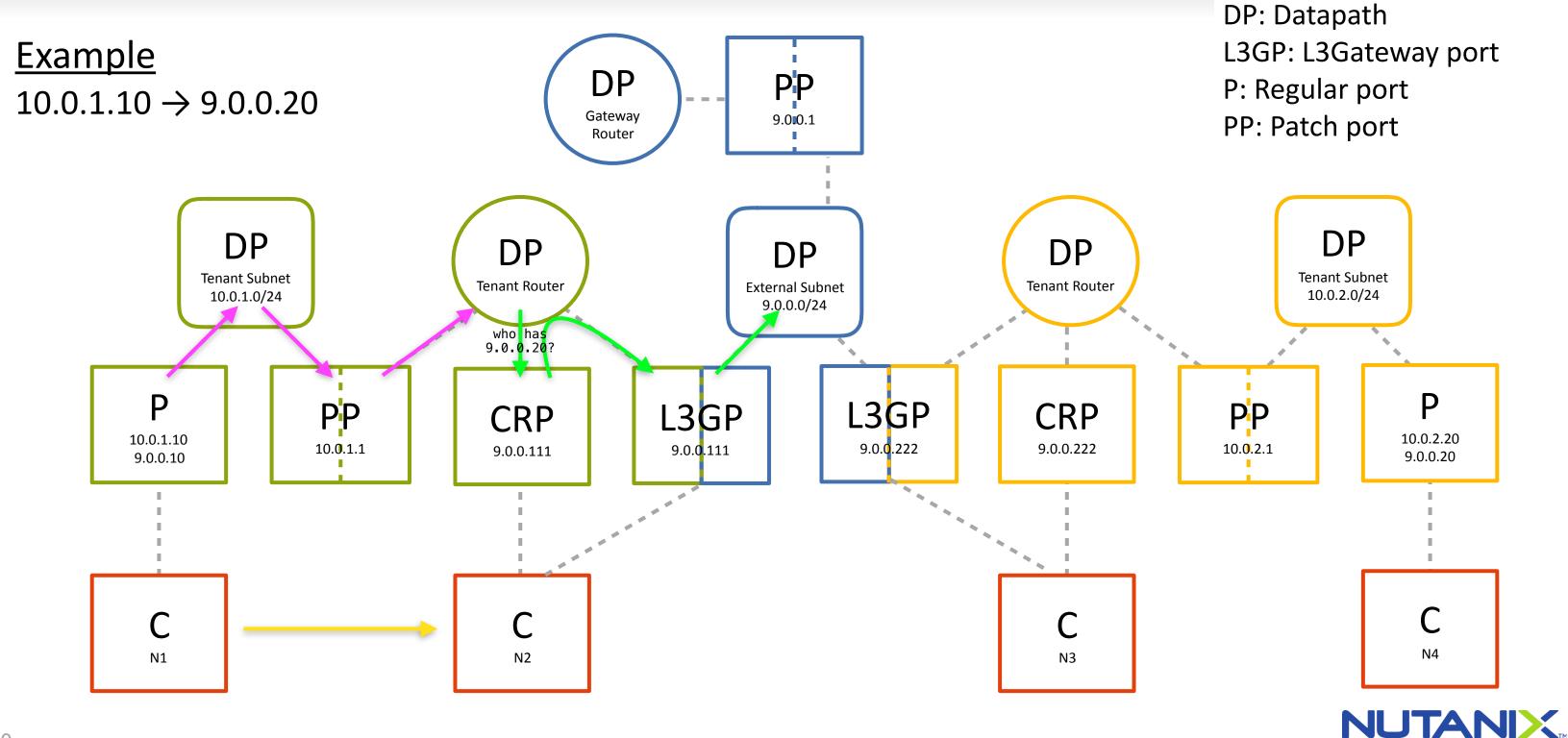


> (Ab)using l3gateway

Use "I3gateway" instead of "patch" ports.

- Bind the I3gateway and chassis-redirect ports to the same chassis.
- Configure floating IPs for "centralized NAT".
 - CRP chassis has ARP responders for both SNAT IP and FIPs.





Legend

C: Chassis

CRP: Chassis redirect port

Example L3GP: L3Gateway port DP PP P: Regular port $10.0.1.10 \rightarrow 9.0.0.20$ Gateway 9.0.0.1 PP: Patch port Router DP DP DP DP Tenant Subnet Tenant Subnet Tenant Router **External Subnet** Tenant Router 10.0.1.0/24 10.0.2.0/24 9.0.0.0/24 who has 9.0.0.20? PP _3GP CRP L3GP **CRP** 10.0.2.20 10.0.1.10 10.0.1.1 9.0.0.111 9.0.0.111 9.0.0.222 9.0.0.222 10.0.2.1 9.0.0.20 9.0.0.10 N1 N2 N3

Legend

C: Chassis

CRP: Chassis redirect port



Example L3GP: L3Gateway port DP PP P: Regular port $10.0.1.10 \rightarrow 9.0.0.20$ Gateway 9.0.0.1 PP: Patch port Router DP DP DP DP Tenant Subnet Tenant Subnet Tenant Router **External Subnet** Tenant Router 10.0.1.0/24 10.0.2.0/24 9.0.0.0/24 who has 9.0.0.20? PP _3GP **CRP** L3GP **CRP** 10.0.2.20 10.0.1.10 10.0.1.1 9.0.0.111 9.0.0.111 9.0.0.222 9.0.0.222 10.0.2.1 9.0.0.20 9.0.0.10 N1 N2 N3

Legend

C: Chassis

CRP: Chassis redirect port



L3GP: L3Gateway port **Example** DP PP P: Regular port $10.0.1.10 \rightarrow 9.0.0.20$ Gateway 9.0.0.1 PP: Patch port Router DP DP DP **Tenant Subnet Tenant Subnet External Subnet** Tenant Router Tenant Router 10.0.1.0/24 10.0.2.0/24 9.0.0.0/24 who has 9.0.0.20? PP L3GP **CRP** 3GP 10.0.2.20 10.0.1.10 9.0.0.222 10.0.1.1 9.0.0.111 9.0.0.222 10.0.2.1 9.0.0.111 9.0.0.20 9.0.0.10 N1 N2 N3

C: Chassis

CRP: Chassis redirect port



L3GP: L3Gateway port **Example** DP PP P: Regular port $10.0.1.10 \rightarrow 9.0.0.20$ Gateway 9.0.0.1 PP: Patch port Router DP DP DP **Tenant Subnet Tenant Subnet External Subnet** Tenant Router Tenant Router 10.0.1.0/24 10.0.2.0/24 9.0.0.0/24 who has 9.0.0.20? 9.0.0.20! PP L3GP **CRP** 3GP 10.0.2.20 10.0.1.10 9.0.0.222 10.0.1.1 9.0.0.111 9.0.0.222 10.0.2.1 9.0.0.111 9.0.0.20 9.0.0.10 N1 N2 N3

C: Chassis

CRP: Chassis redirect port



> Trade-offs

Pros:

- Works for FIP→FIP, SNAT→FIP.
- Provides isolation; each chassis subscribes only to relevant datapaths.
- Requires only minor changes to networking-ovn and ovn-northd.
- No hair-pinning of traffic to gateway appliance.

Cons:

- Doesn't work with distributed NAT.
- Doesn't work (nicely) with gateway HA.
- Still relies on ARP, even though we know the MAC binding.



Closing Questions

- Does anyone else see value in splitting the gateway?
- The I3gateway thing is obviously a hack. How can we do better?

