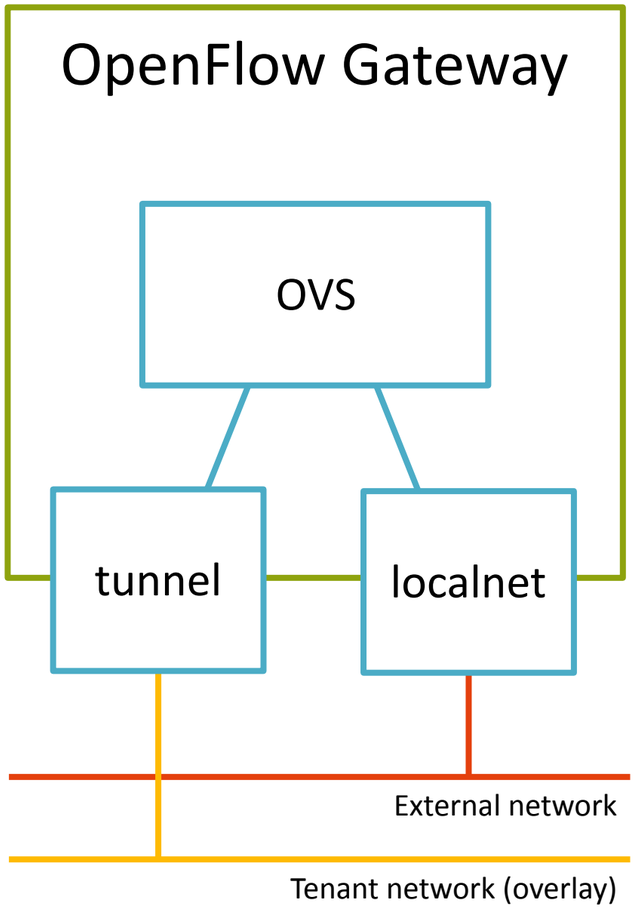
NUTANIX™

External Networks on the Overlay

December 6, 2018

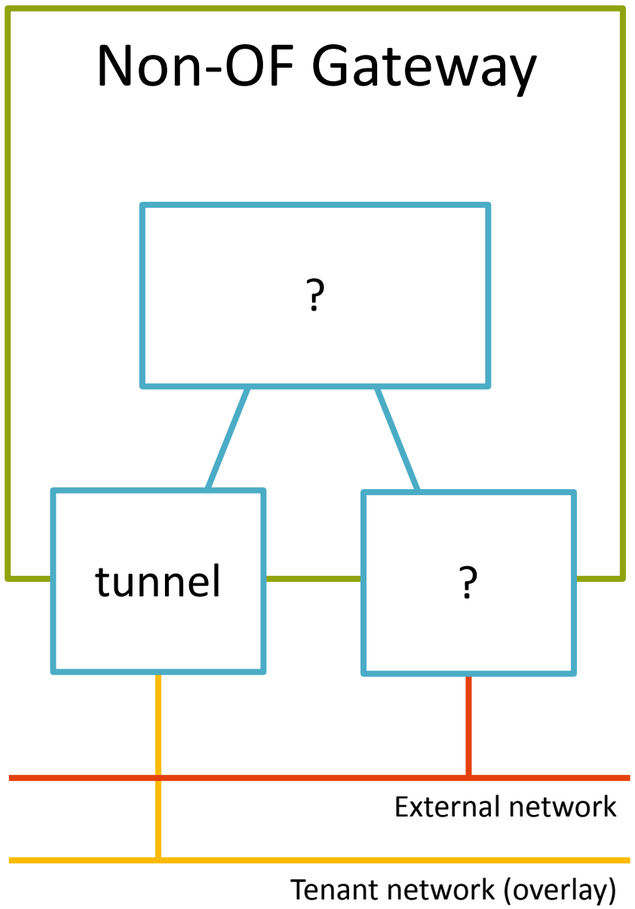
Gregory Smith

Typical External Connectivity



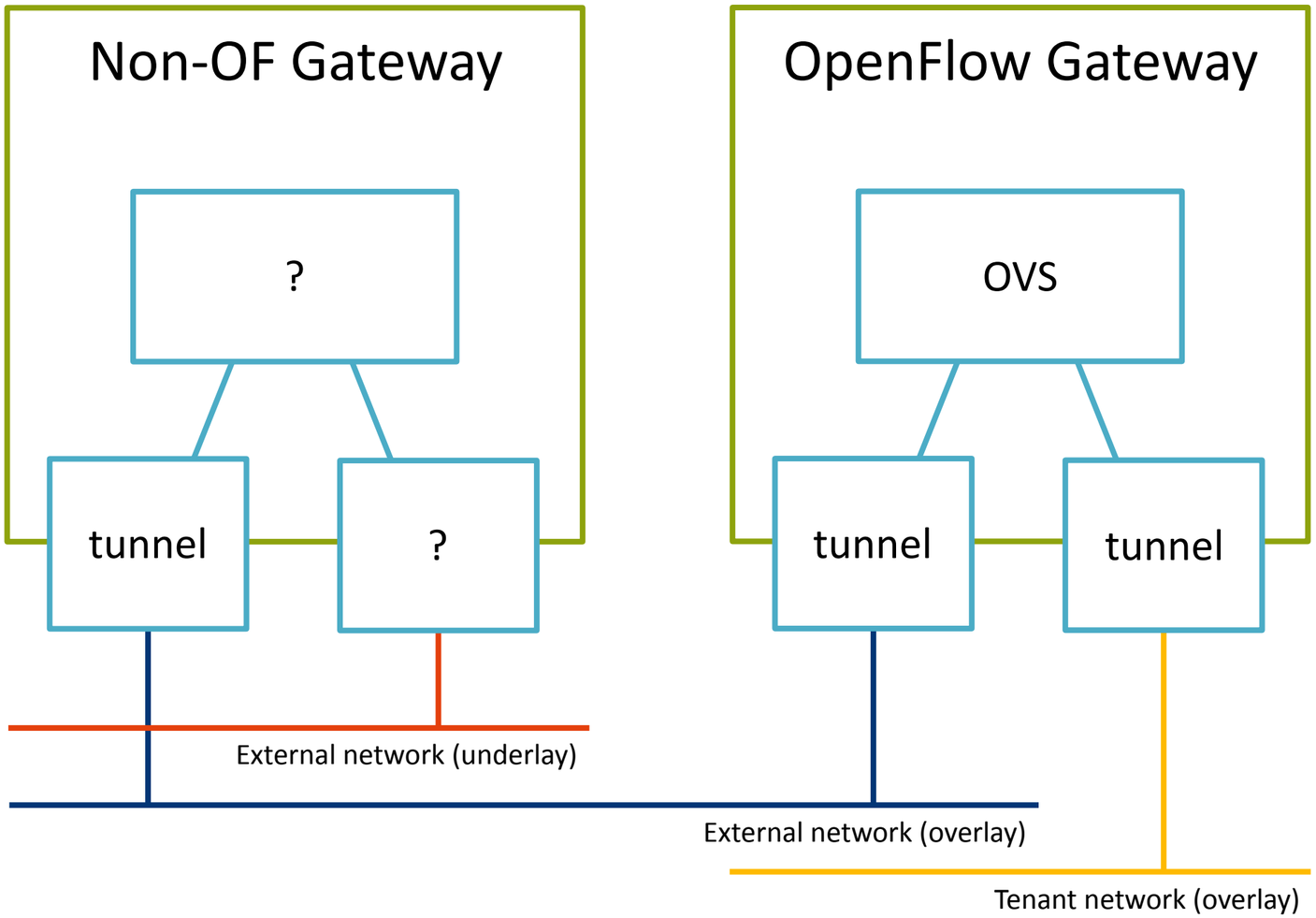
- Gateway chassis physically connected to external network via localnet port.
- OVN programs NAT rules into OVS via openflow for floating IPs and SNAT.

Non-OpenFlow Gateway

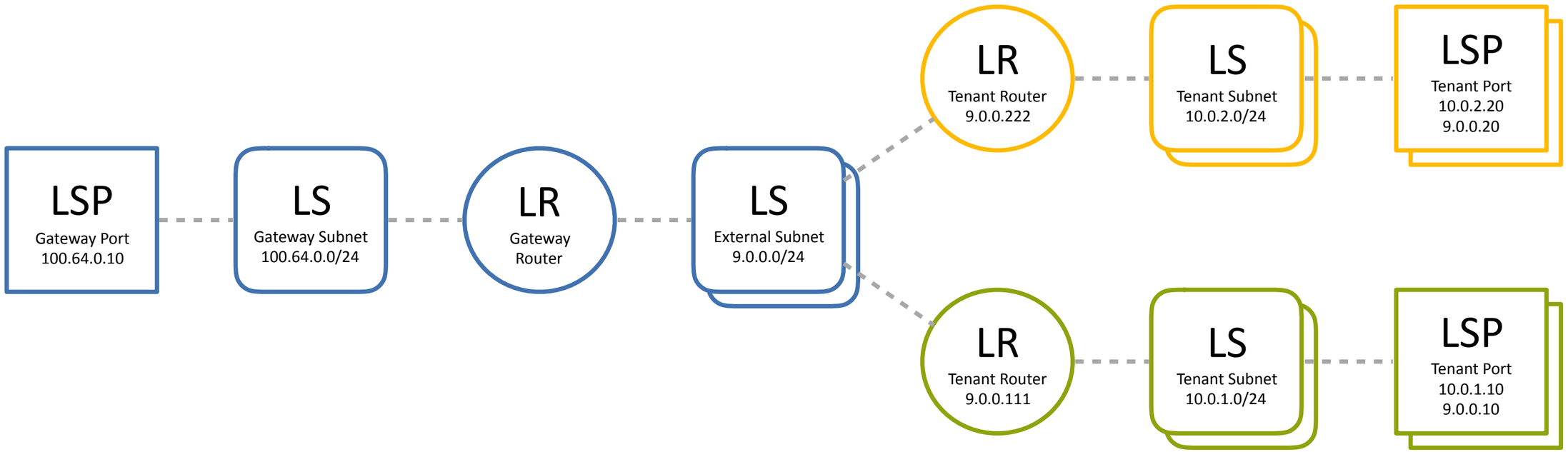


- Program gateway through netconf instead of OVN/OpenFlow.
- Keep configuration simple and static - no NAT.

Splitting The Gateway



Logical Topology



Legend

LR: Logical router

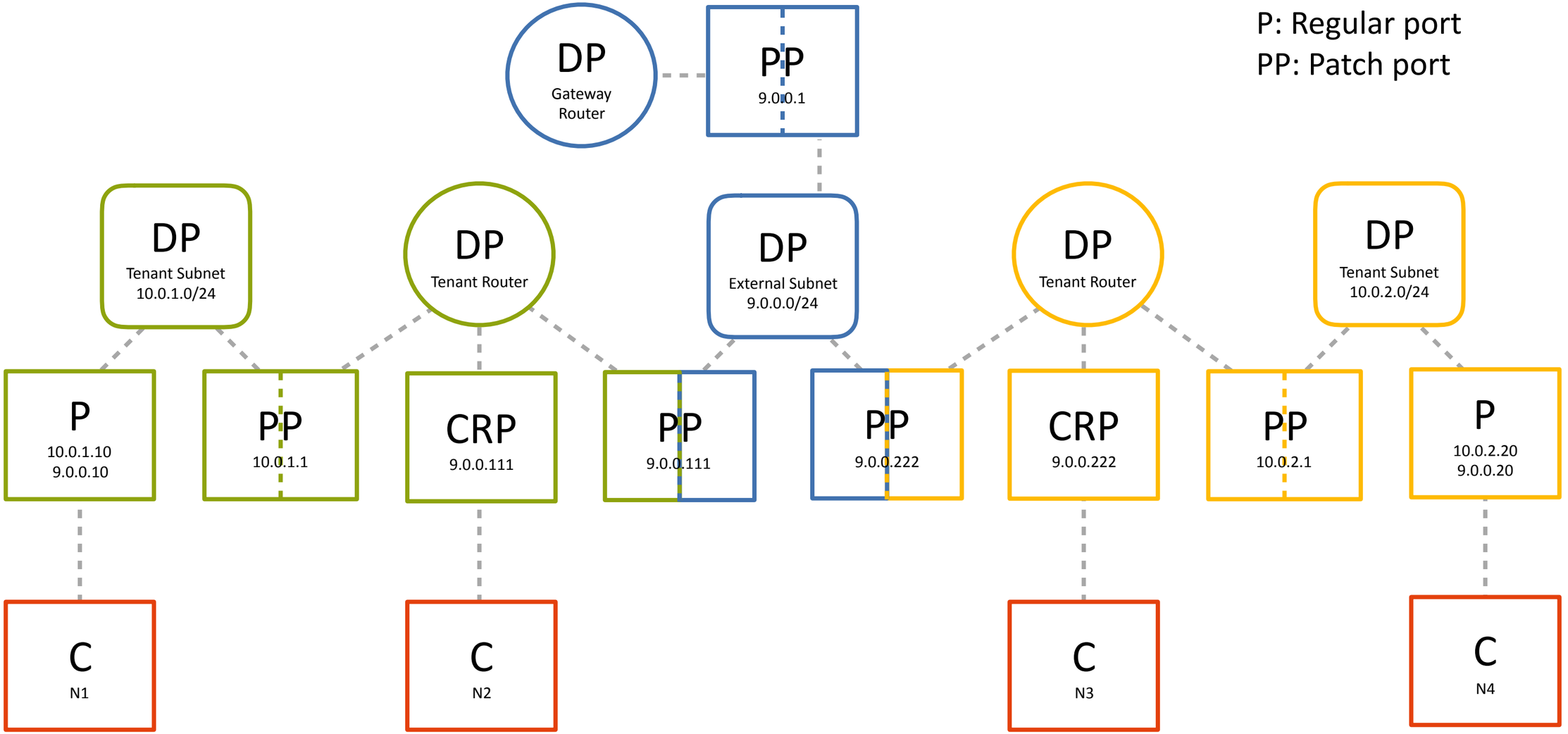
LS: Logical switch

LSP: Logical switch port

Datapath Topology

Legend

C: Chassis
CRP: Chassis redirect port
DP: Datapath
P: Regular port
PP: Patch port



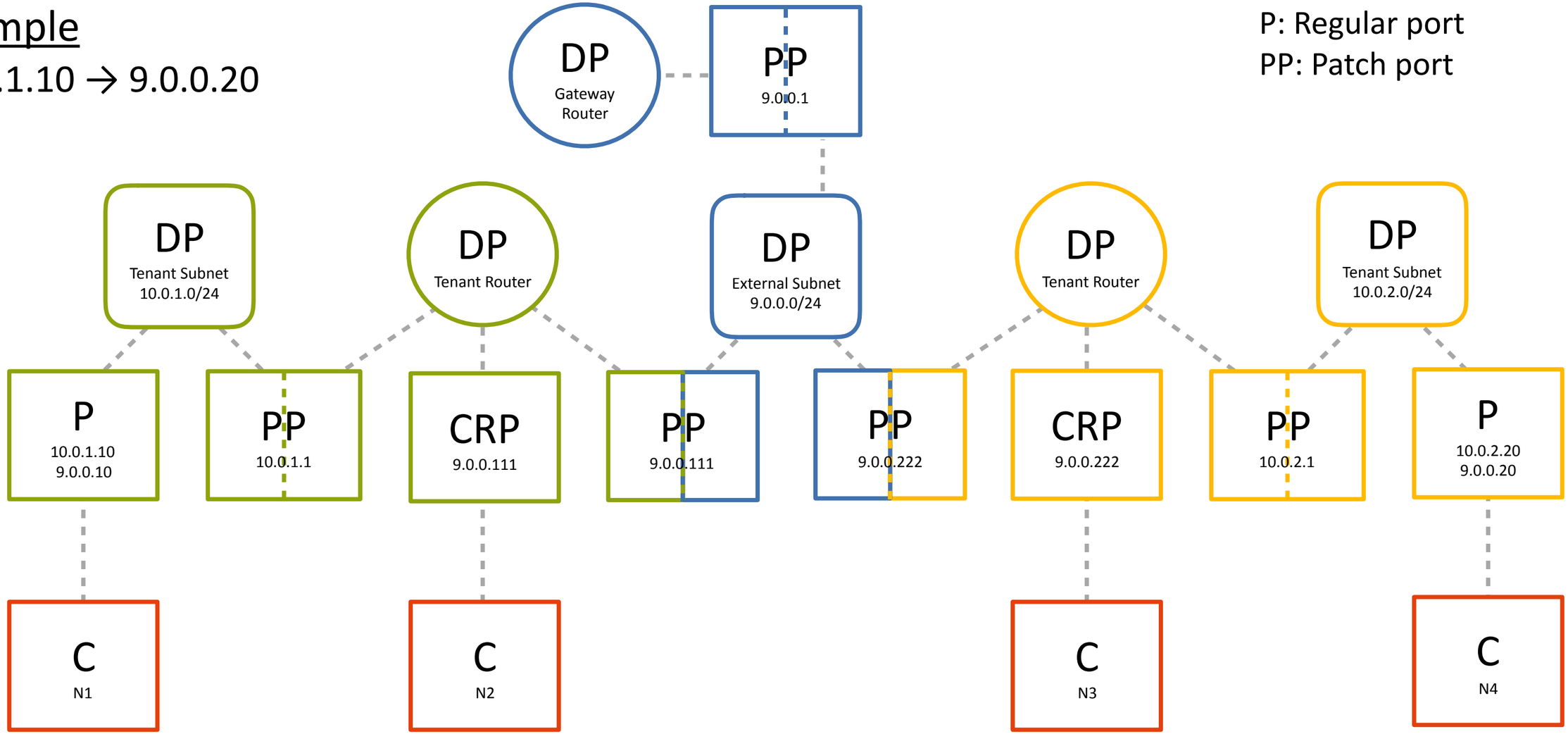
Datapath Example

Legend

C: Chassis
CRP: Chassis redirect port
DP: Datapath
P: Regular port
PP: Patch port

Example

10.0.1.10 → 9.0.0.20



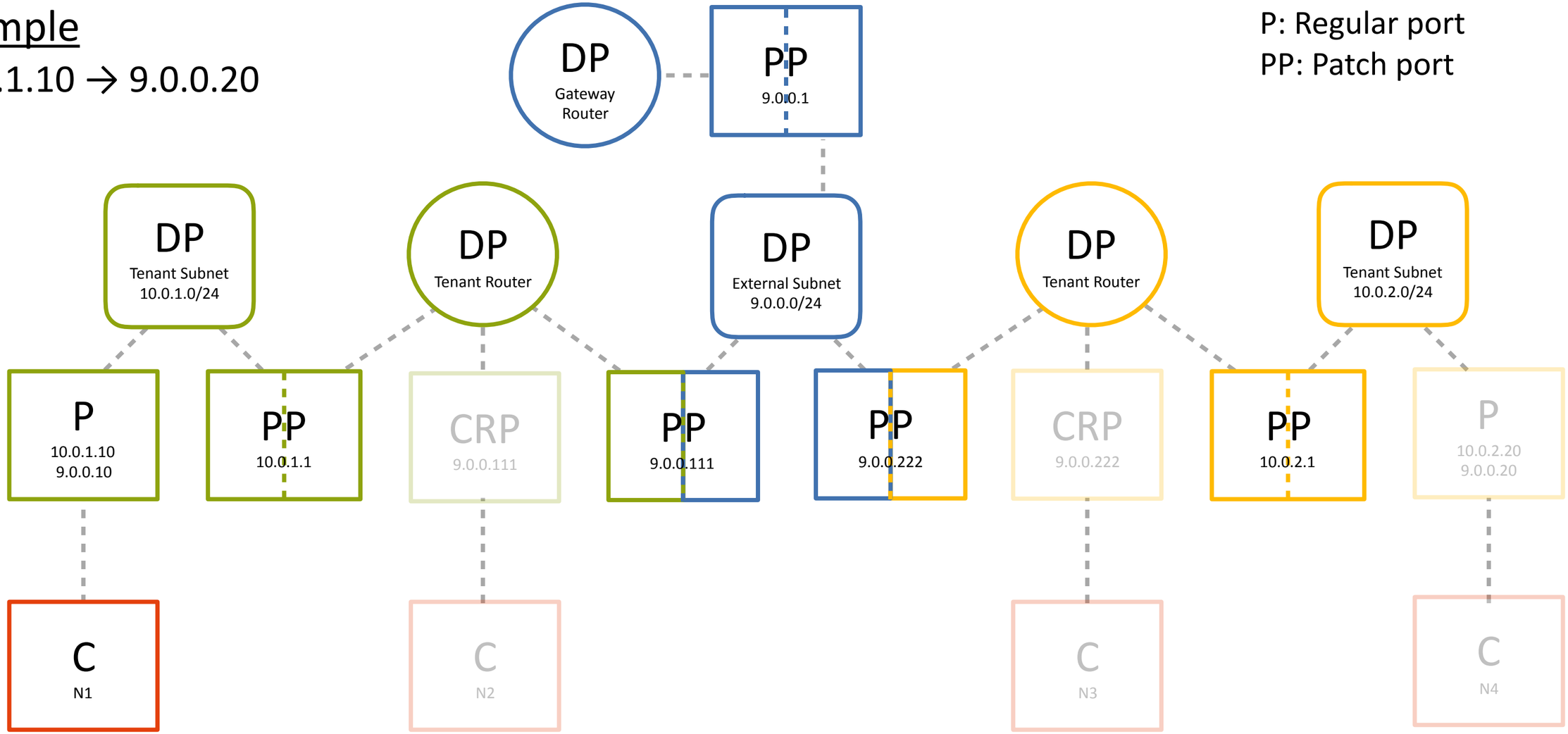
Datapath Example

Legend

C: Chassis
CRP: Chassis redirect port
DP: Datapath
P: Regular port
PP: Patch port

Example

10.0.1.10 → 9.0.0.20



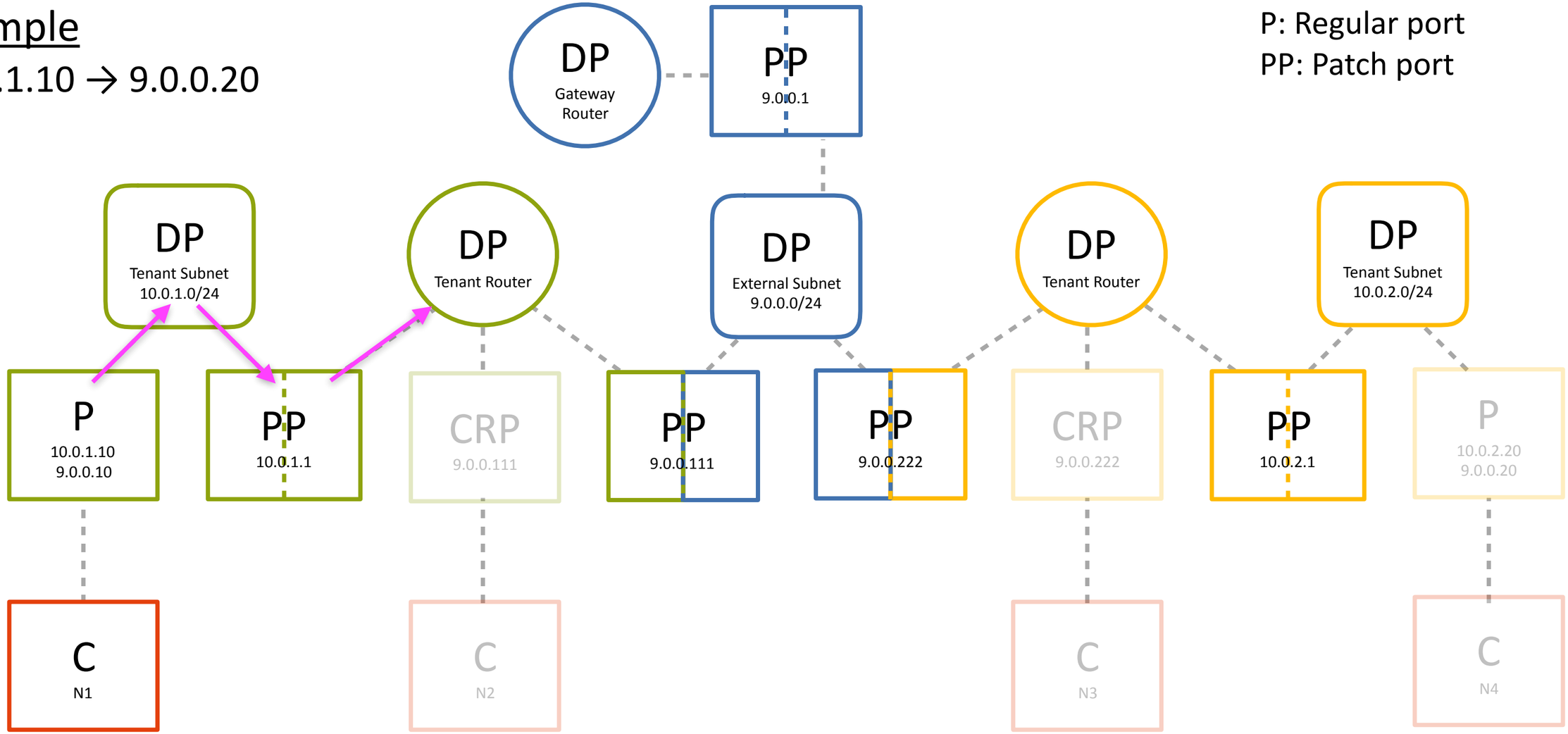
Datapath Example

Legend

C: Chassis
CRP: Chassis redirect port
DP: Datapath
P: Regular port
PP: Patch port

Example

10.0.1.10 → 9.0.0.20



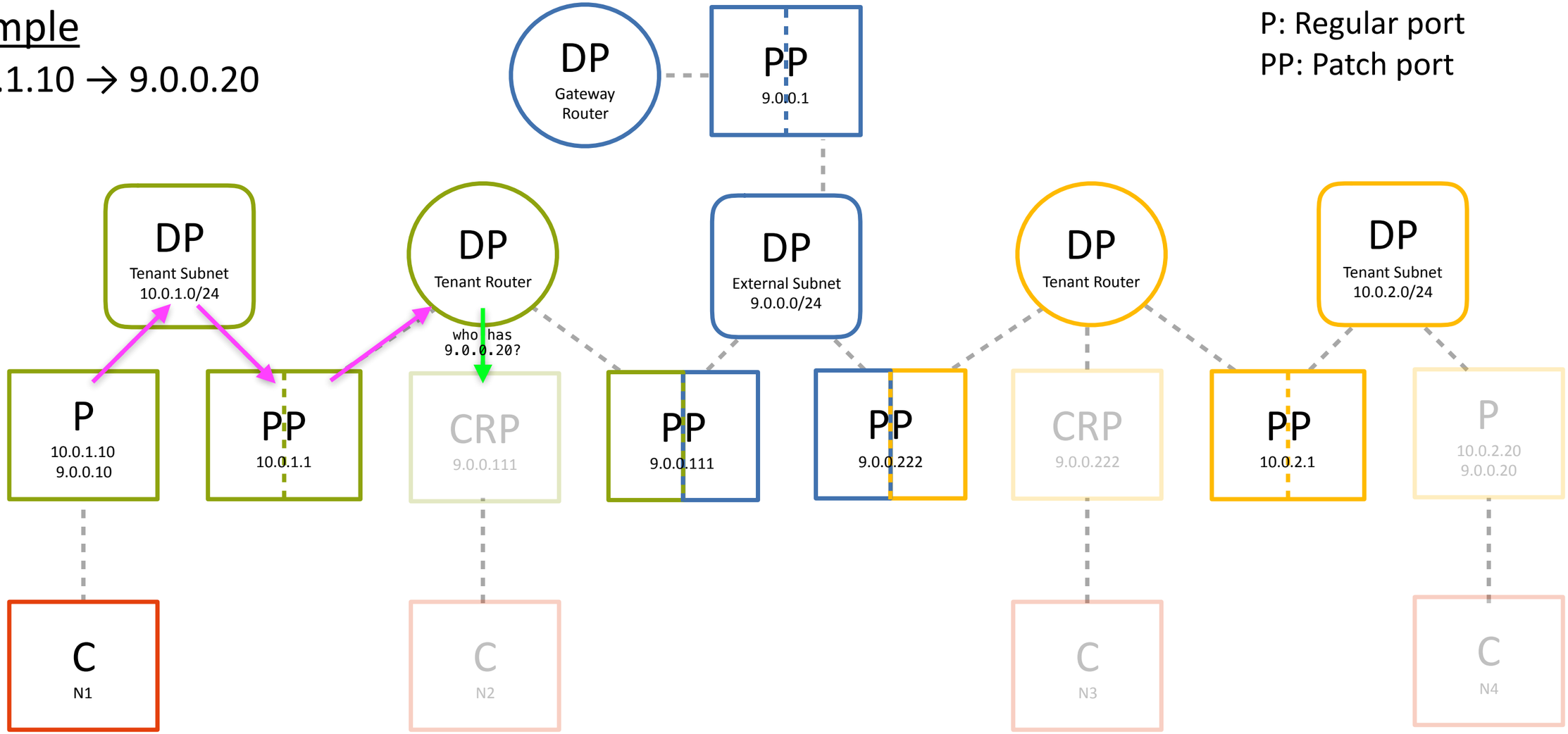
Datapath Example

Legend

C: Chassis
CRP: Chassis redirect port
DP: Datapath
P: Regular port
PP: Patch port

Example

10.0.1.10 → 9.0.0.20



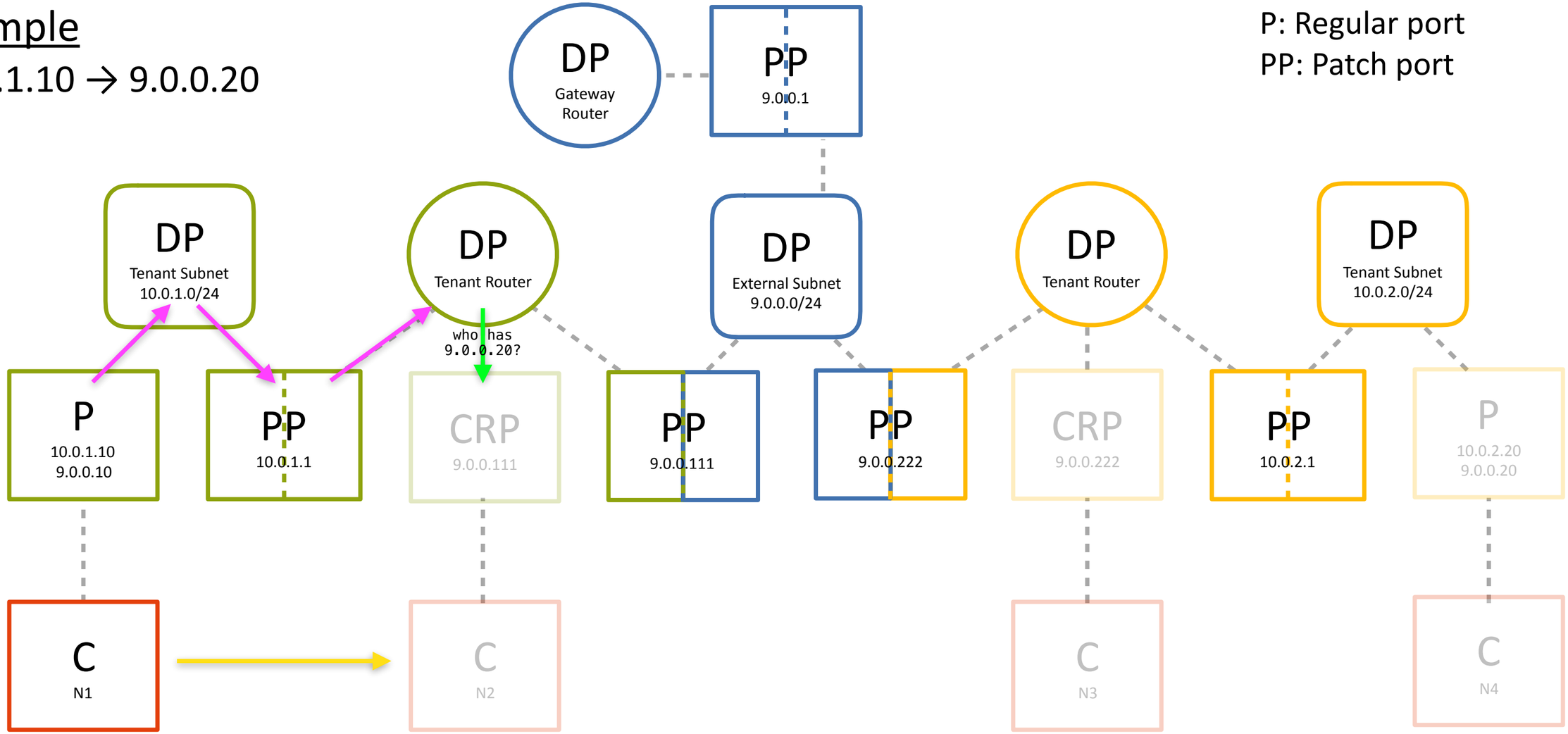
Datapath Example

Legend

C: Chassis
CRP: Chassis redirect port
DP: Datapath
P: Regular port
PP: Patch port

Example

10.0.1.10 → 9.0.0.20



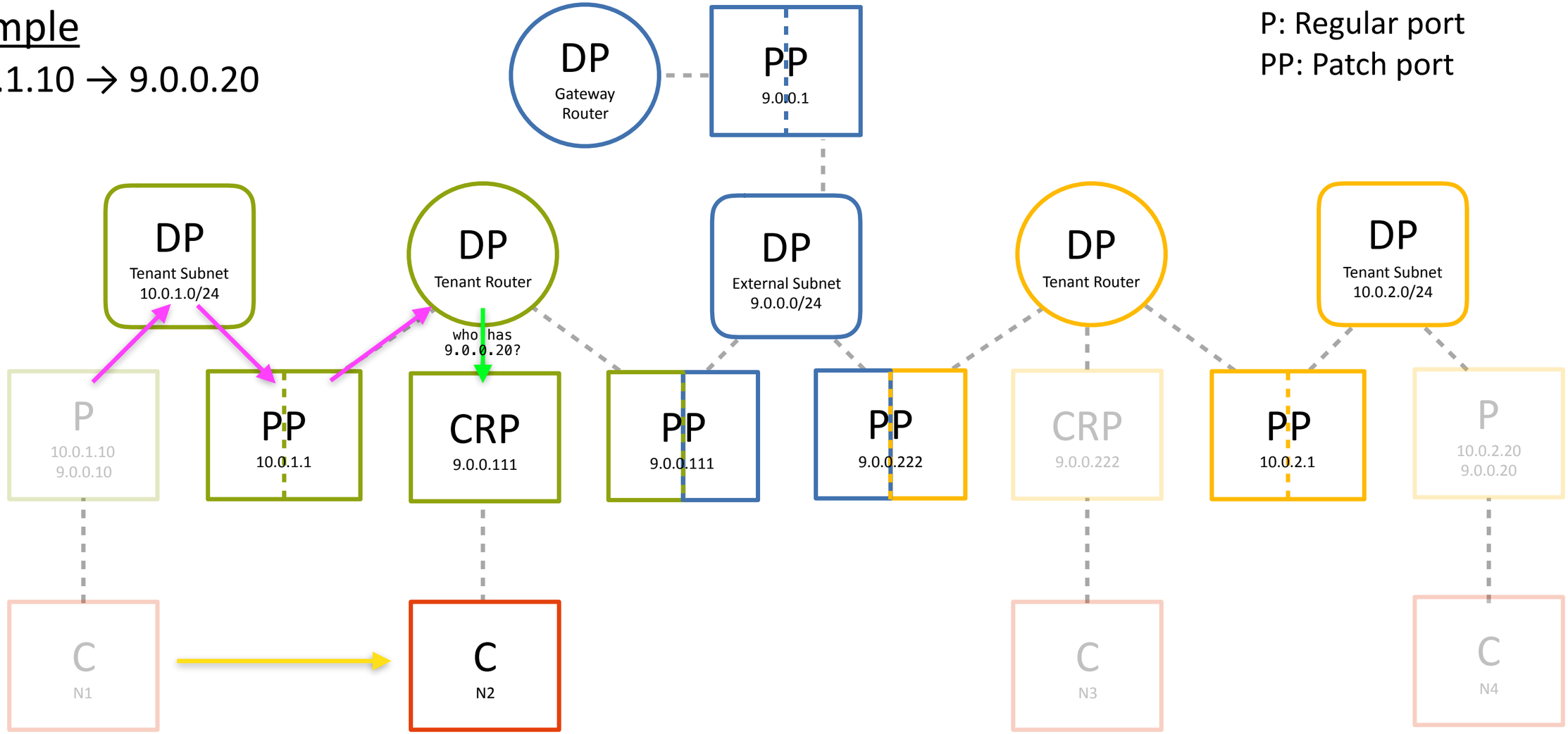
Datapath Example

Legend

C: Chassis
CRP: Chassis redirect port
DP: Datapath
P: Regular port
PP: Patch port

Example

10.0.1.10 → 9.0.0.20



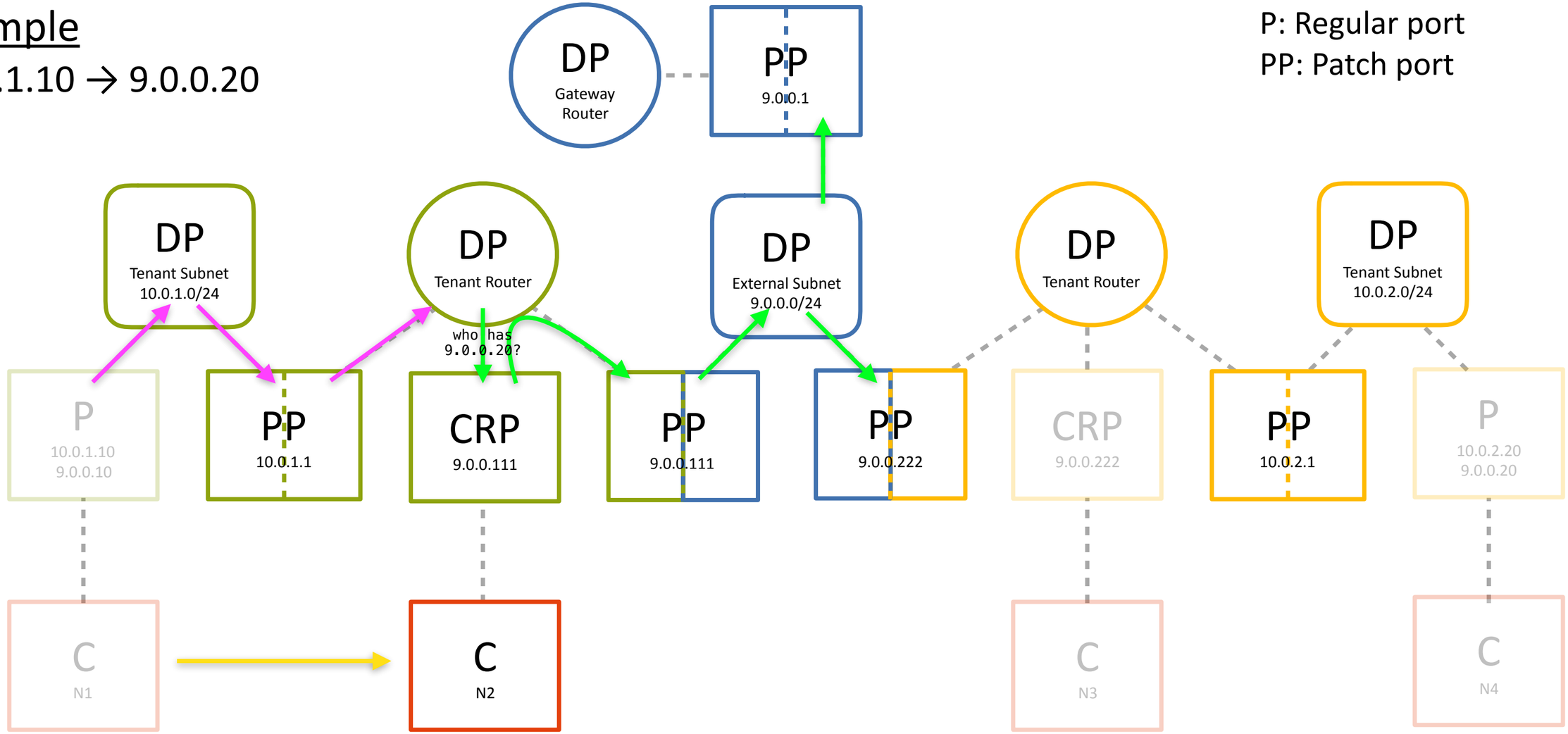
Datapath Example

Legend

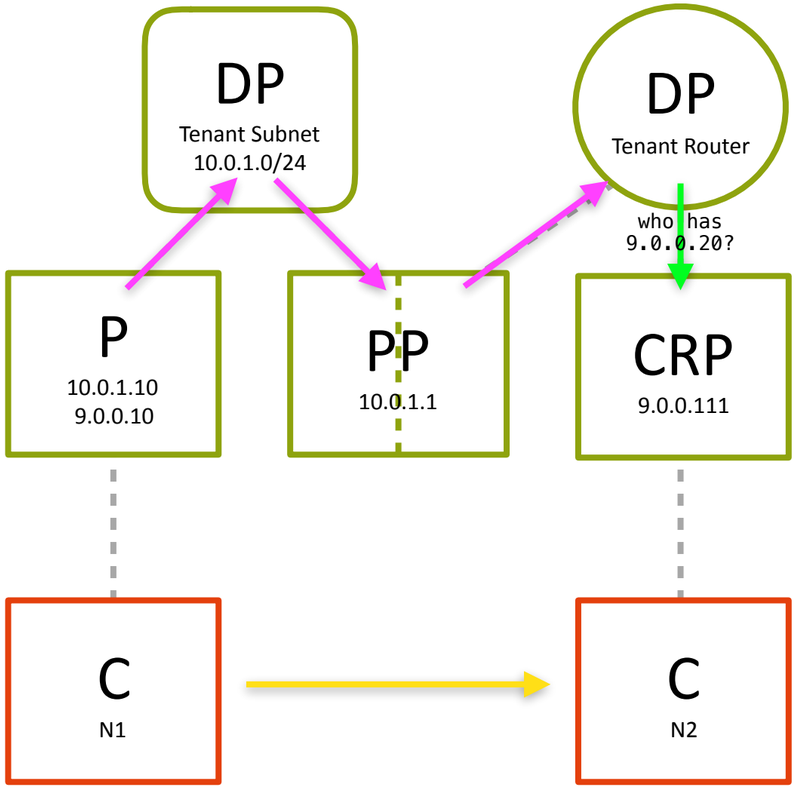
C: Chassis
CRP: Chassis redirect port
DP: Datapath
P: Regular port
PP: Patch port

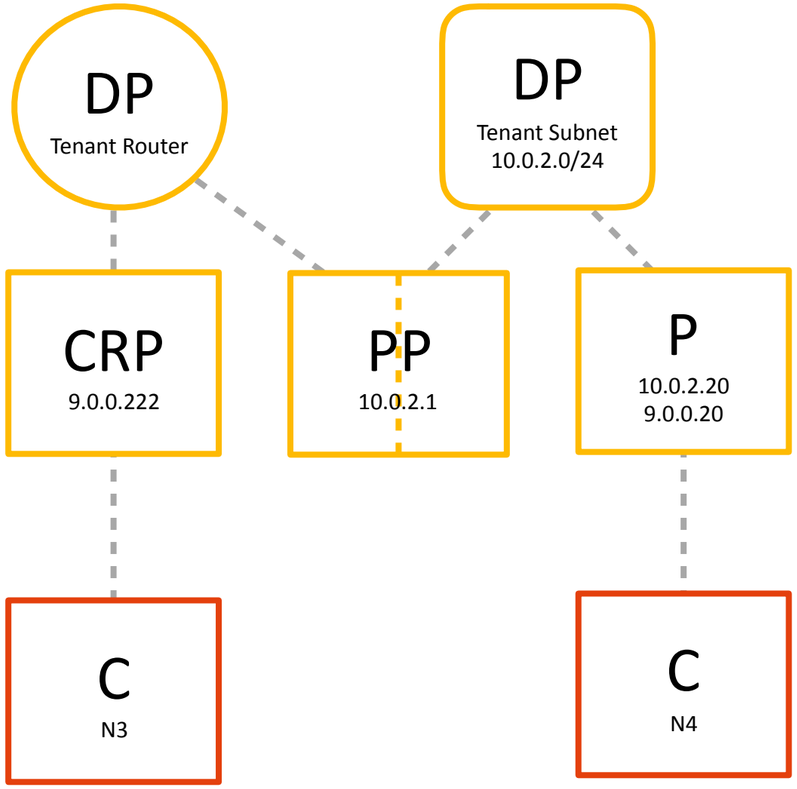
Example

10.0.1.10 → 9.0.0.20

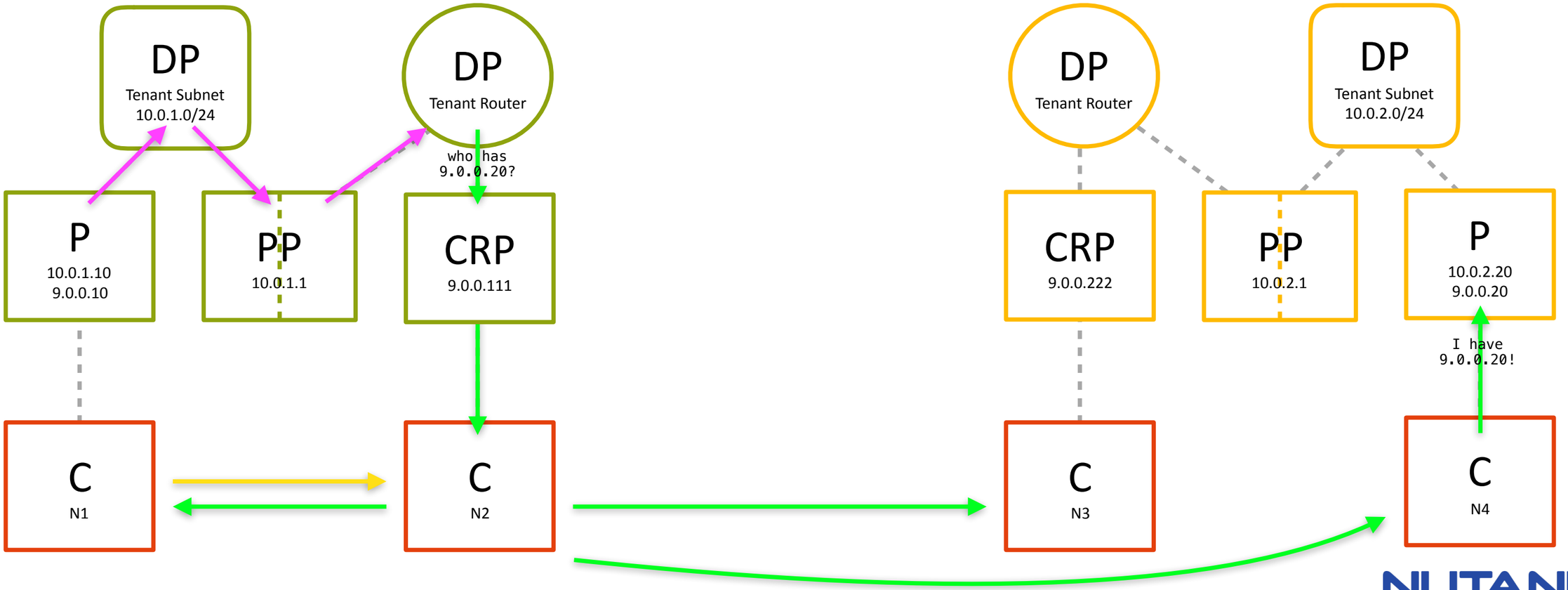


Reference Architecture Example





Reference Architecture Example



(Ab)using l3gateway

Use "l3gateway" instead of "patch" ports.

- Bind the l3gateway 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.

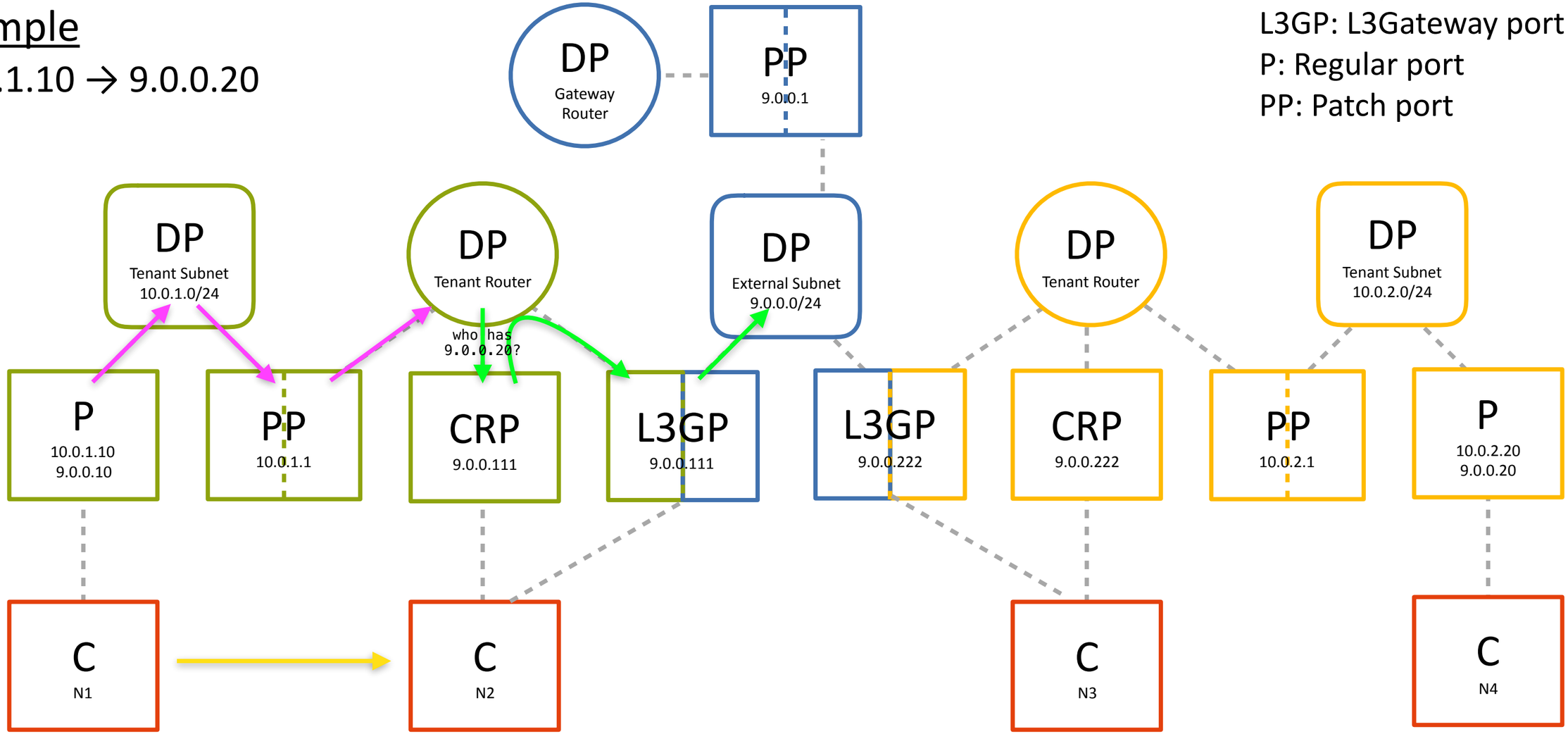
Datapath Example

Example

10.0.1.10 → 9.0.0.20

Legend

C: Chassis
CRP: Chassis redirect port
DP: Datapath
L3GP: L3Gateway port
P: Regular port
PP: Patch port



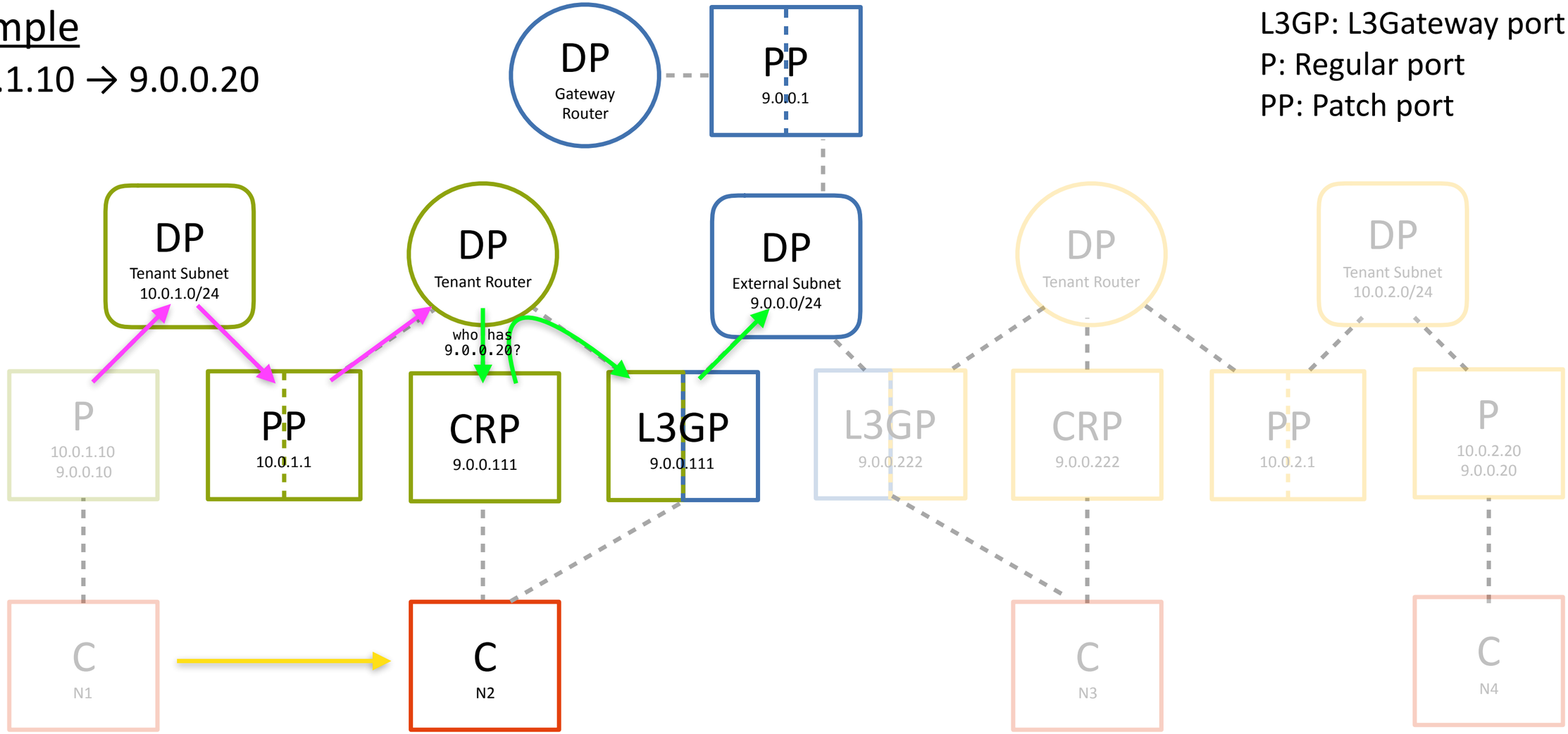
Datapath Example

Legend

C: Chassis
CRP: Chassis redirect port
DP: Datapath
L3GP: L3Gateway port
P: Regular port
PP: Patch port

Example

10.0.1.10 → 9.0.0.20



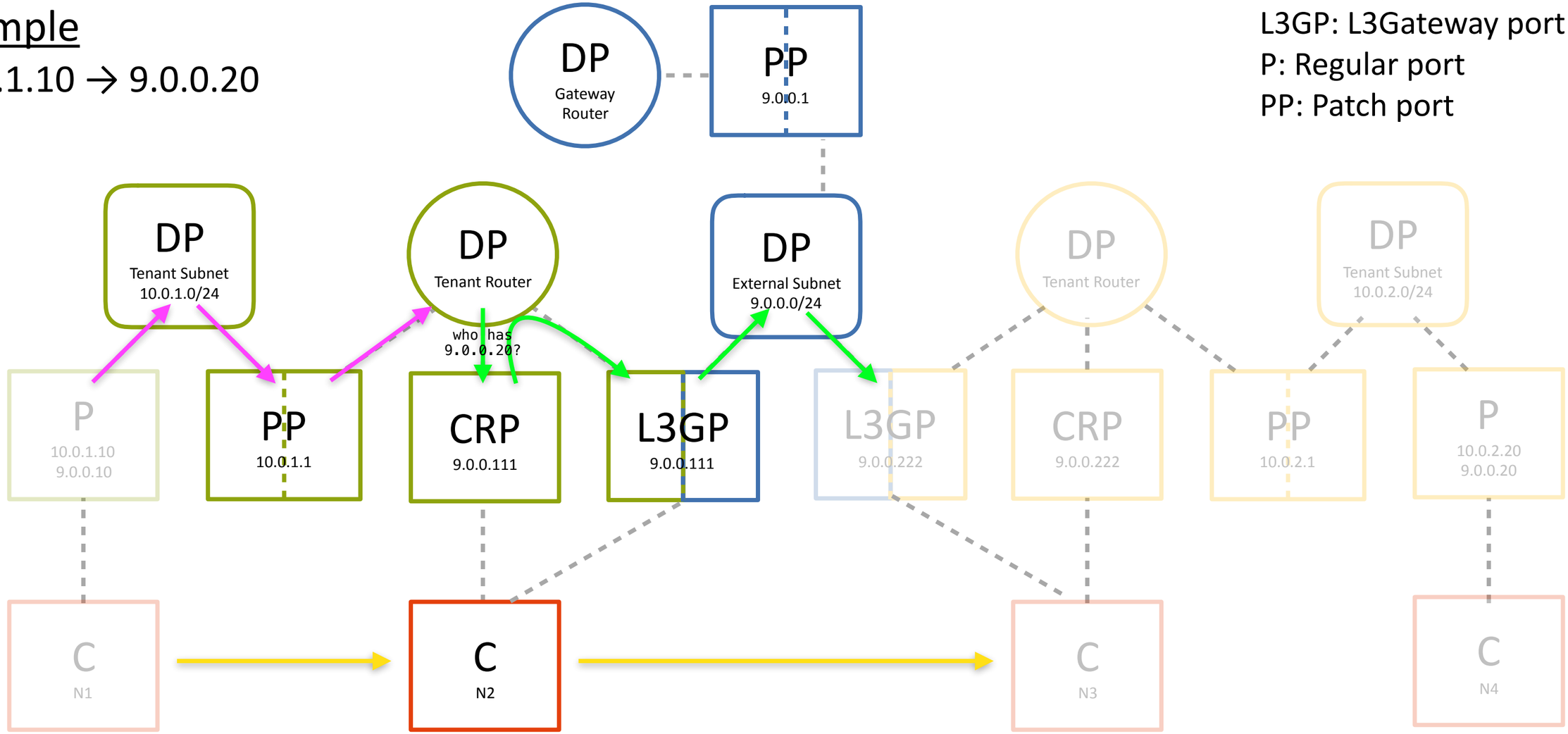
Datapath Example

Legend

C: Chassis
CRP: Chassis redirect port
DP: Datapath
L3GP: L3Gateway port
P: Regular port
PP: Patch port

Example

10.0.1.10 → 9.0.0.20



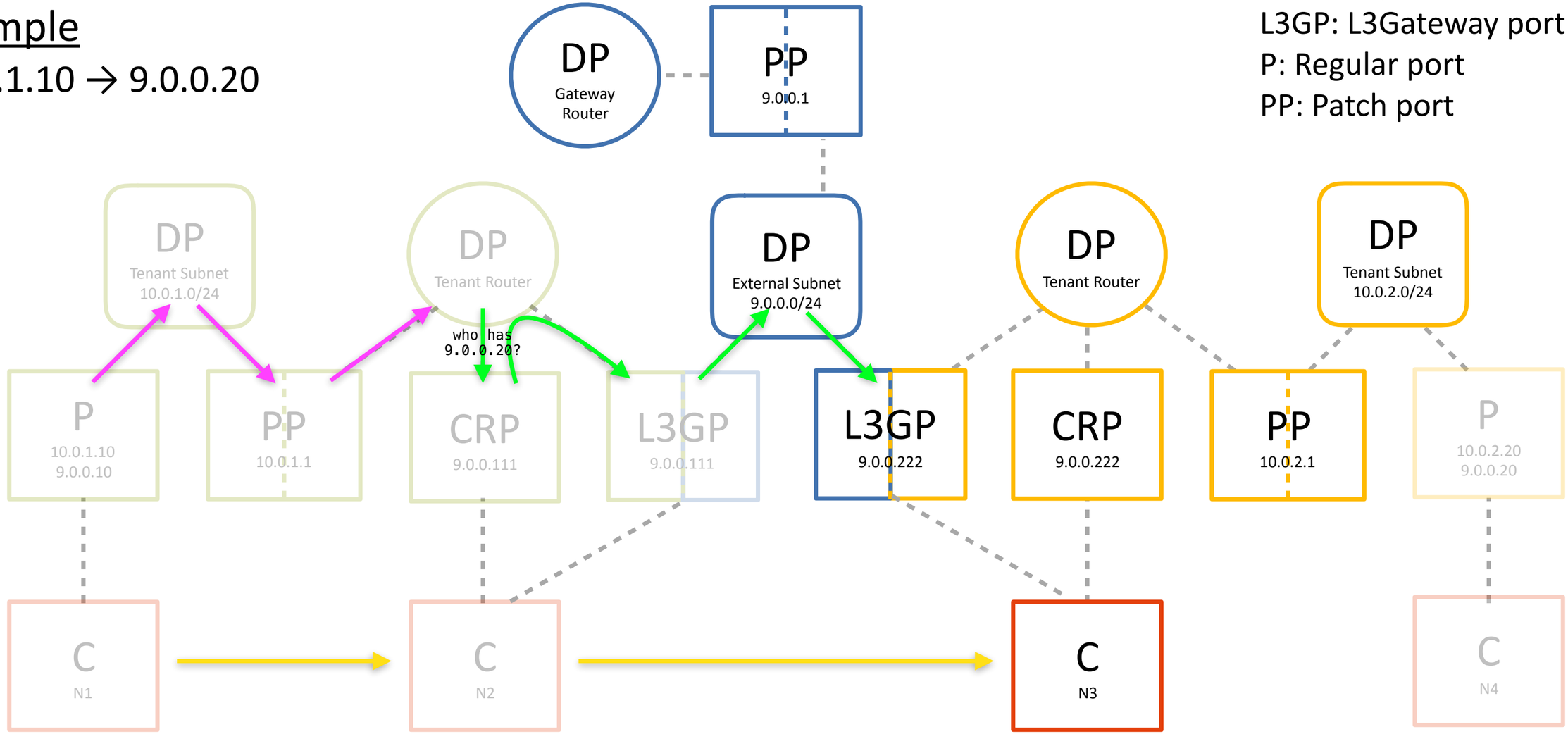
Datapath Example

Legend

C: Chassis
CRP: Chassis redirect port
DP: Datapath
L3GP: L3Gateway port
P: Regular port
PP: Patch port

Example

10.0.1.10 → 9.0.0.20

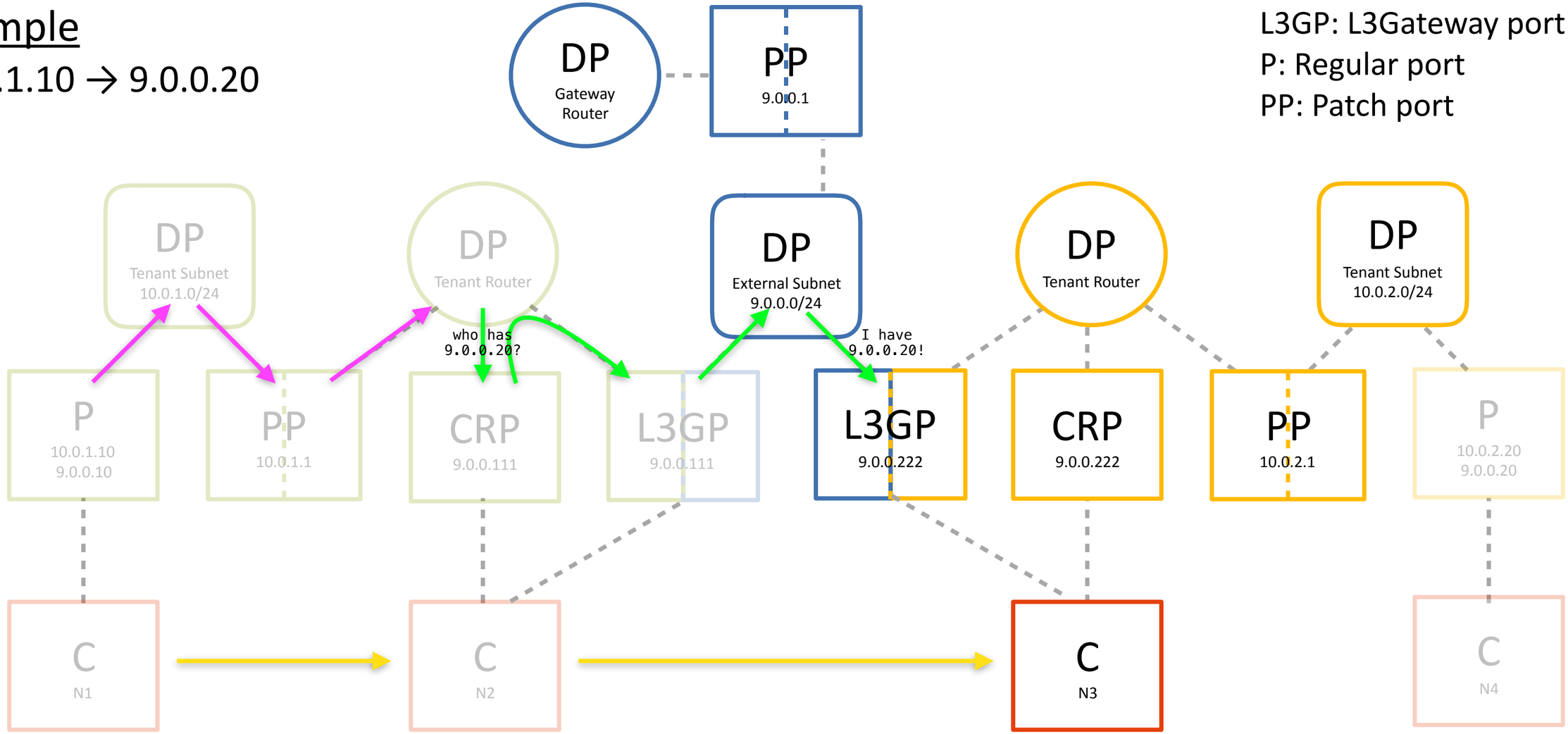


NUTANIX

Datapath Example

Example

10.0.1.10 → 9.0.0.20



Legend

C: Chassis
CRP: Chassis redirect port
DP: Datapath
L3GP: L3Gateway port
P: Regular port
PP: Patch 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 l3gateway thing is obviously a hack. How can we do better?