NSX-T Deep Dive: Layer 3 Routing

Part 1

Agenda

NSX-T Data Center Vision & Architecture

(See Previous Sessions in Introduction Modules)

NSX-T Logical Routing: Deep Dive Part 1

- Terminology
- Feature Overview
- Packet flows and N/S connectivity options
- Multi-tier routing
- Routing features
- High Availability/Resiliency

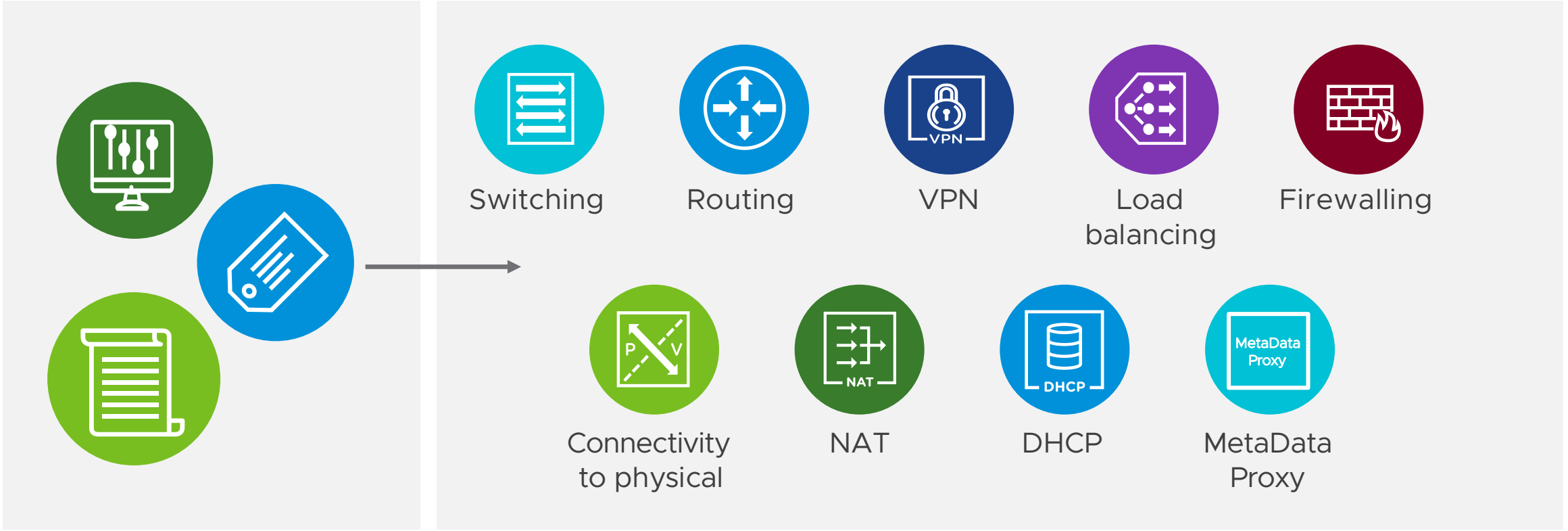
NSX-T Logical Routing: Deep Dive Part 2

(View Part 2 in the Learning Path)

Summary

This Part 1 session on **NSX-T Logical Routing: Deep Dive** will cover all of the listed items. Part 2 will go into an advanced view of the packet flow within the NSX-T routing architecture.

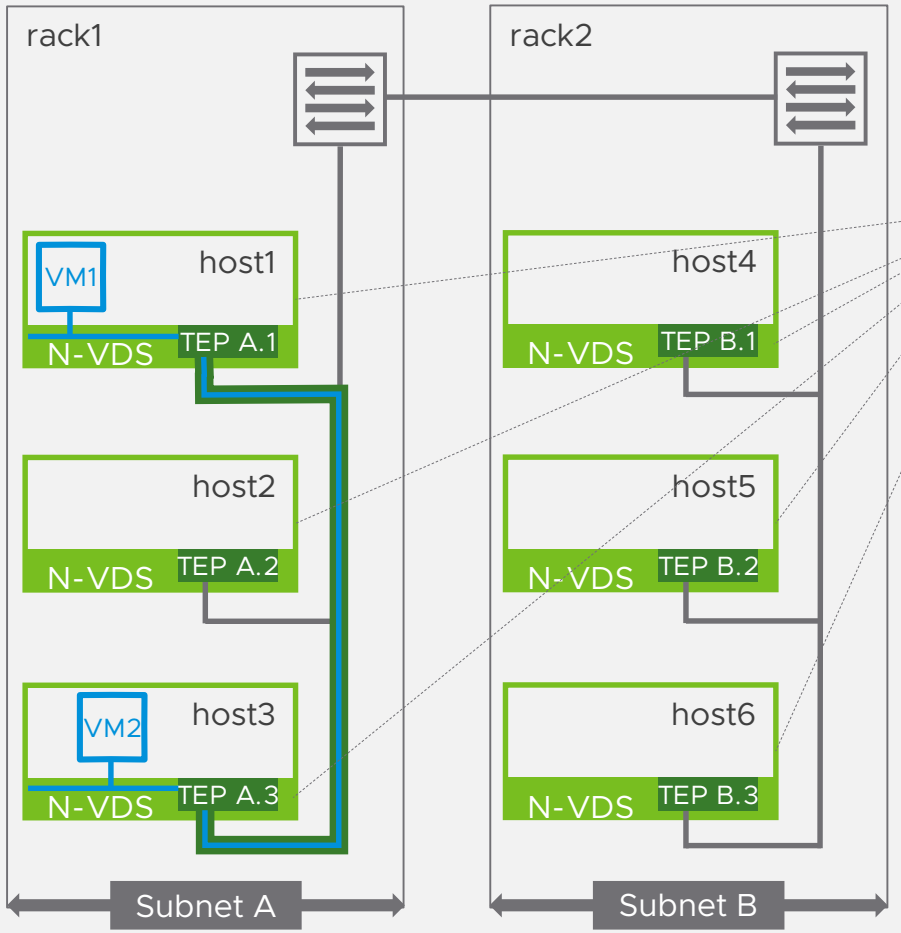
NSX-T Networking and Security Services



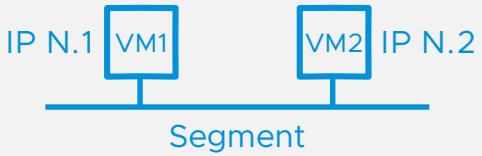
The NSX-T platform is on its 5th release and with every release we are adding more and more features . NSX-t provides all of these networking and security services in software. Every release introduces new features, scale increase and enhancements to the existing features. In this session we will focus on routing.

Network Virtualization – Overlay Model

Logical switch – segment (L2 Broadcast Domain)



Logical View



VM (vnic)	Location
Mac VM1	TEP A.1
Mac VM2	TEP A.3

TEP: Tunnel End Point

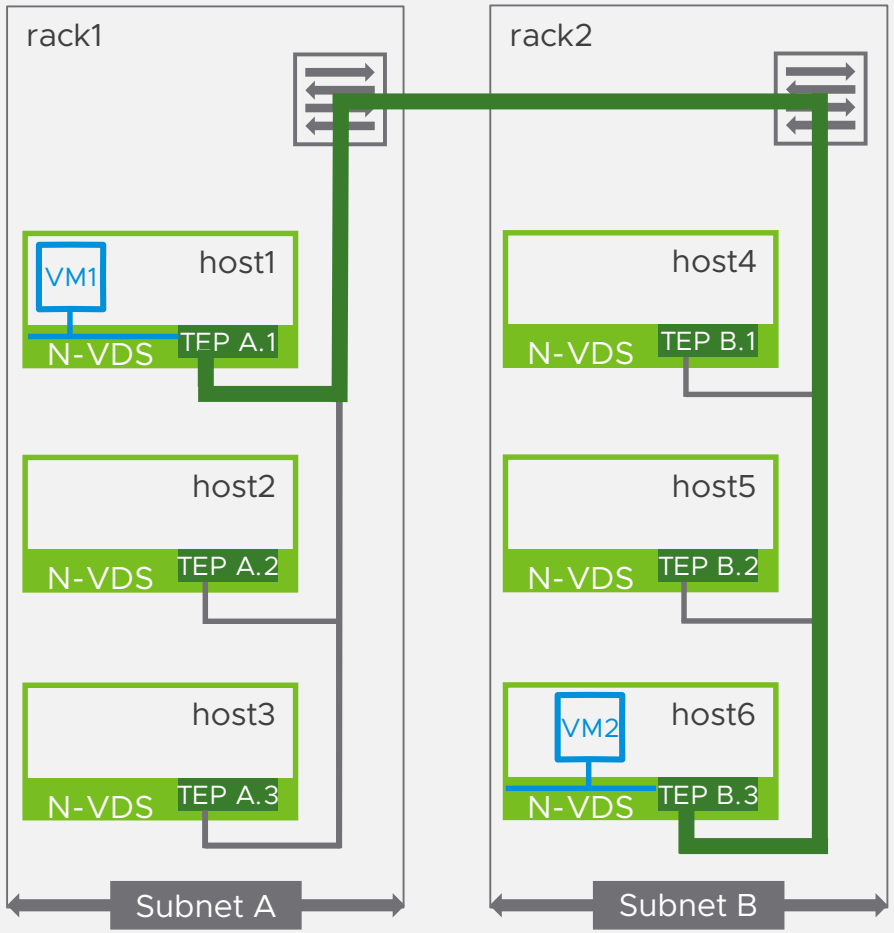
The N-VDS or, NSX Virtual Distributed Switch, is the NSX data plane component. Logical switches now called Segments are instantiated on the hypervisors.

The Segments are extended between the hypervisors by IP tunnels utilizing the IETF Geneve overlay.

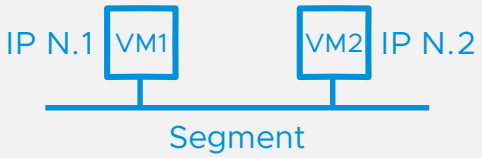
NSX maintains a table locating the position of the virtual elements in the physical network communicated through the Central Control Plane of the Management Cluster.

Network Virtualization – Overlay Model

Logically switch over L2/L3 physical fabric



Logical View



VM (vnic)	Location
Mac VM1	TEP A.1
Mac VM2	TEP B.3

TEP: Tunnel End Point

The NSX Overlay is agnostic to the physical fabric connectivity. The Tunneling end points (TEPs) of the hypervisors provide the tunnels connecting the segments. This allows the logical segments connectivity any type of L2 or L3 physical fabric.

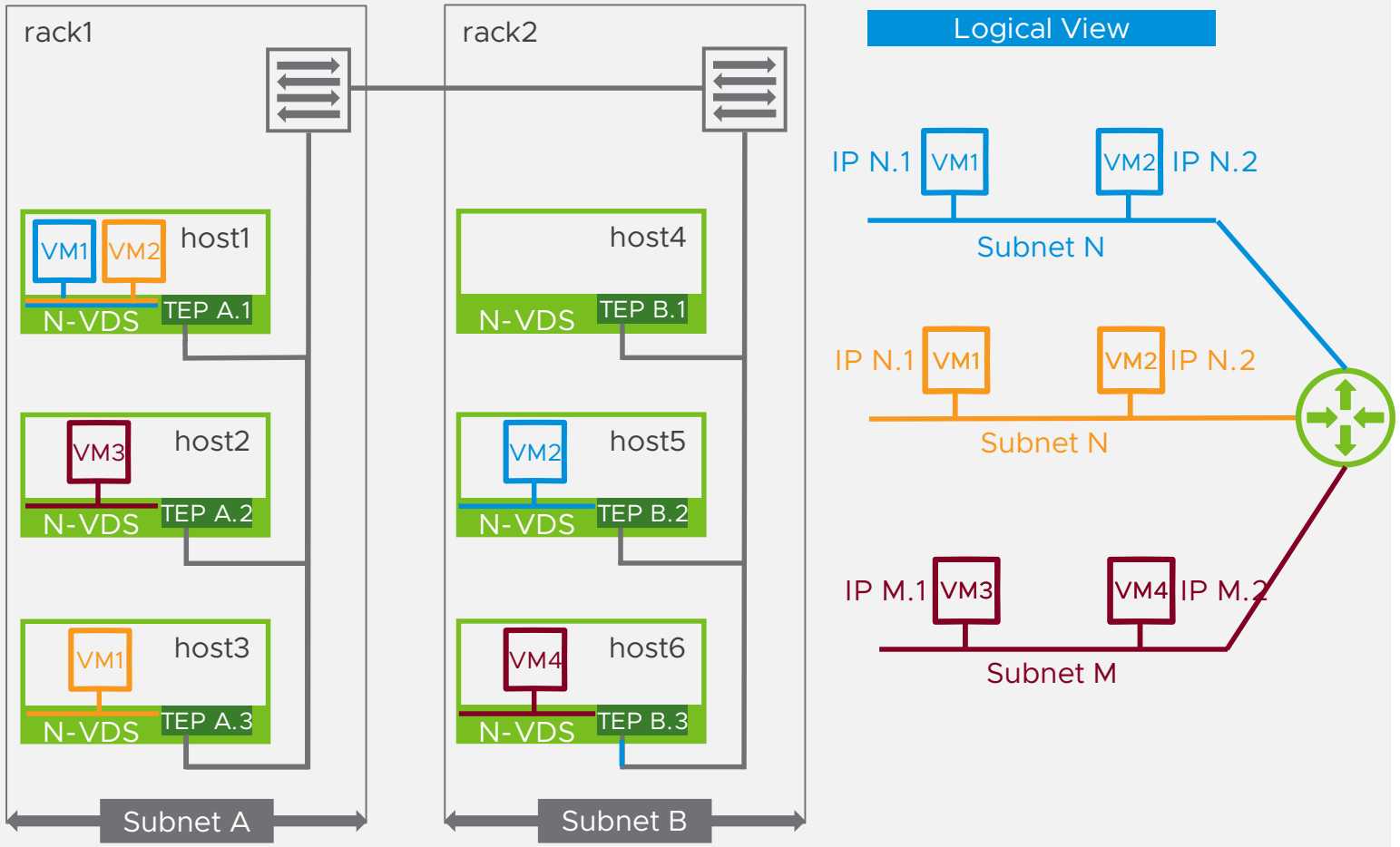
The physical fabric requirements are the following:

- IP connectivity
- 1600 byte MTU minimum (jumbo frame recommended)

The physical switch fabric should be a simple IP factory of connectivity for hosting the NSX application and security platform.

Complete Virtual Networks

Virtualize networks similar to virtualizing compute

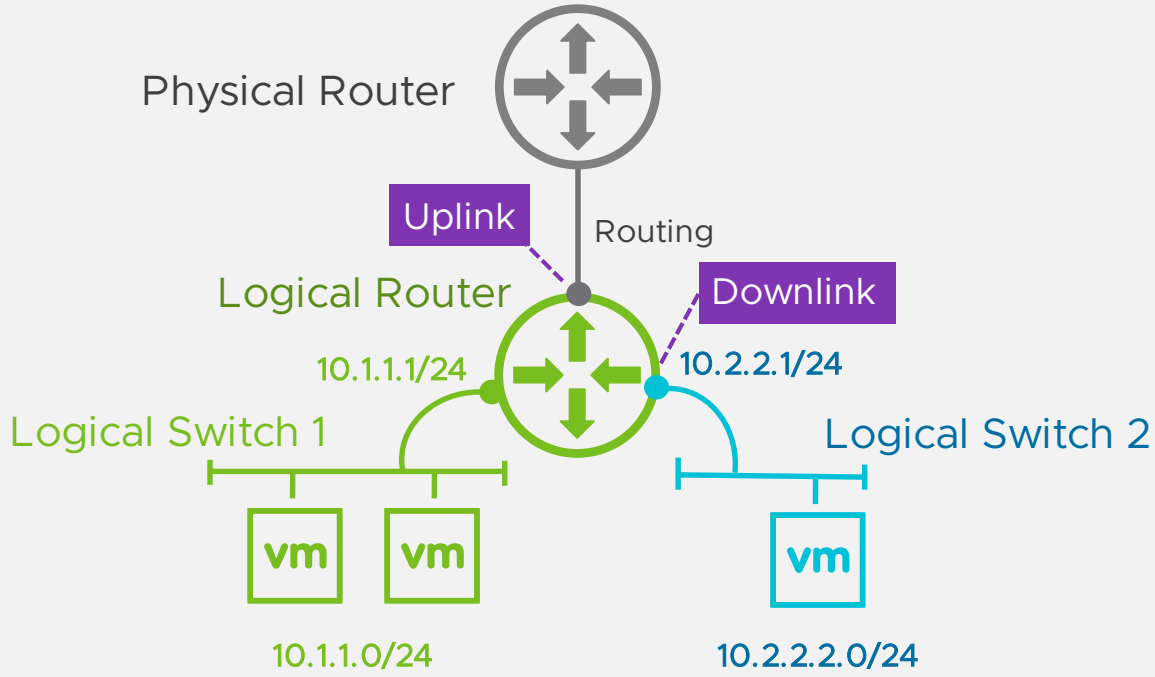


NSX injects distributed routing functionality into every hypervisor. Distributed routing connects logical segments within the tunneled overlay network.

Therefore, elaborately switched and routed connectivity of the virtual workloads are seamlessly connected via the virtual transport network.

Complex application needs are simplified when leveraging the NSX overlay. This lends agility to workload deployment, operations and workload lifecycle management.

Terminology: Introducing the Logical Router(LR)



The Logical Router provides two distinct types of routed connectivity.

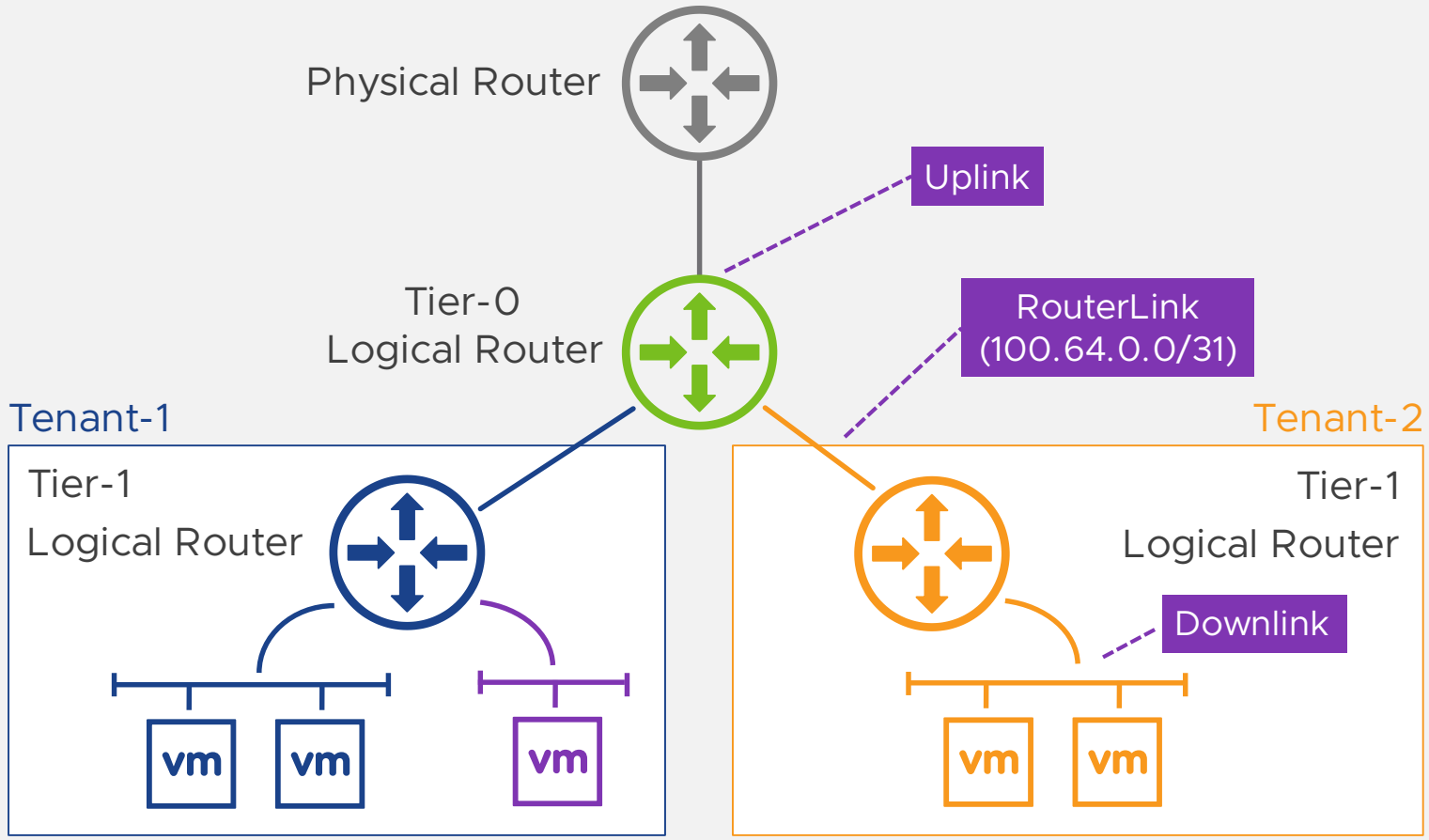
East to West routing when connecting two or more logical segments within the virtualized network overlay.

Additionally, North to South routing is performed where the logical router peers with the physical infrastructure. This is also referred to as the edge routers.

The logical routers for edge connectivity also provide centralized network services such as Network Address Translation(NAT), load-balancing, perimeter firewall, VPN etc

Logical Routing: Multi-tier Topology

NSX Tier-0 and Tier-1 routers for tenant and edge connectivity



Tier-0 and Tier-1 Logical Routers

There are two distinct roles that enable an elegant NSX tiered routing model.

The Tier-0 logical router connects to physical infrastructure. The Tier-0 external routing is managed manually.

The Tier-1 logical router is used as a per tenant first hop router with auto plumbed connectivity to its Tier-0 router.

The multi-tier model has several benefits:

- Tenant isolation
- Separate control for Infrastructure and tenant admin
- Eliminates dependency on physical infrastructure when a new tenant is provisioned

Components of the Logical Router

Distributed Router (DR) and Services Router (SR)

Distributed Router (DR)

- Runs locally in the transport nodes participating in the NSX fabric
- Typically runs as kernel module in the hypervisor
- Provides distributed E-W routing
- Traffic between different subnets on same hypervisor doesn't leave the hypervisor

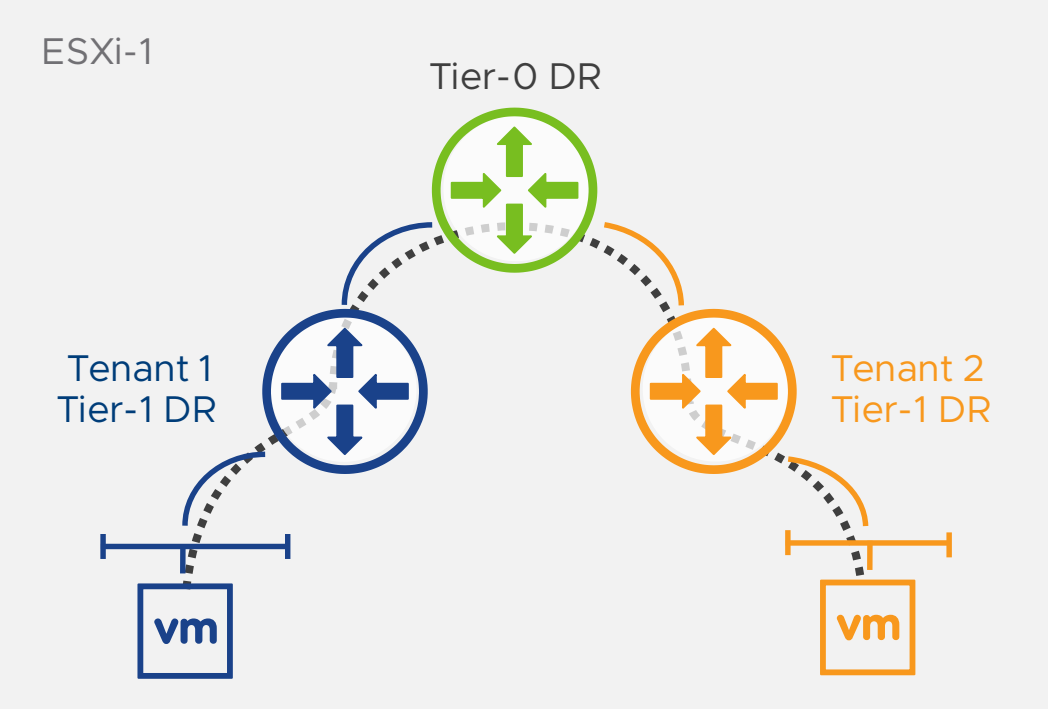
Services Router (SR)

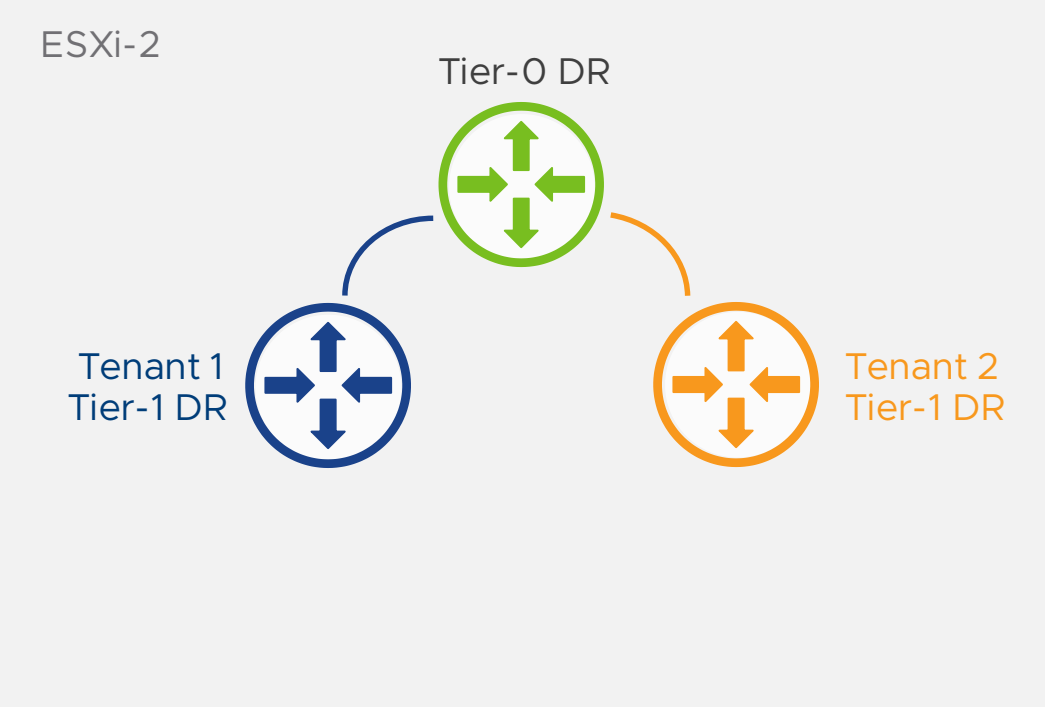
- Responsible for providing on/off ramp gateway services including N/S routing
- Provides centralized services like
  - NAT, BGP, LB, Edge Firewall, connectivity to the physical
- The SR is instantiated as a service on an appliance called the Edge Node

Logical Routing: Multi-tier Topology

Multi-tier distributed routing

The distributed routing model for NSX instantiates Tier-0 and Tier-1 routers on every hypervisor to prevent hair-pinning.



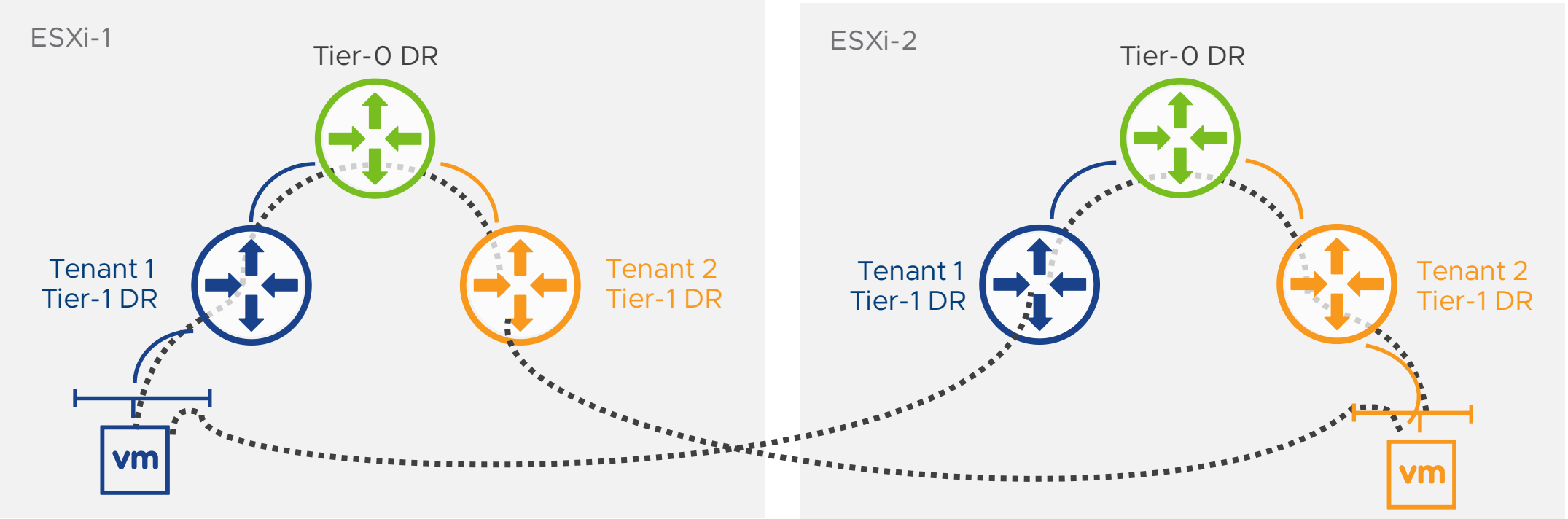


Logical Routing: Multi-tier Topology

Multi-tier distributed routing

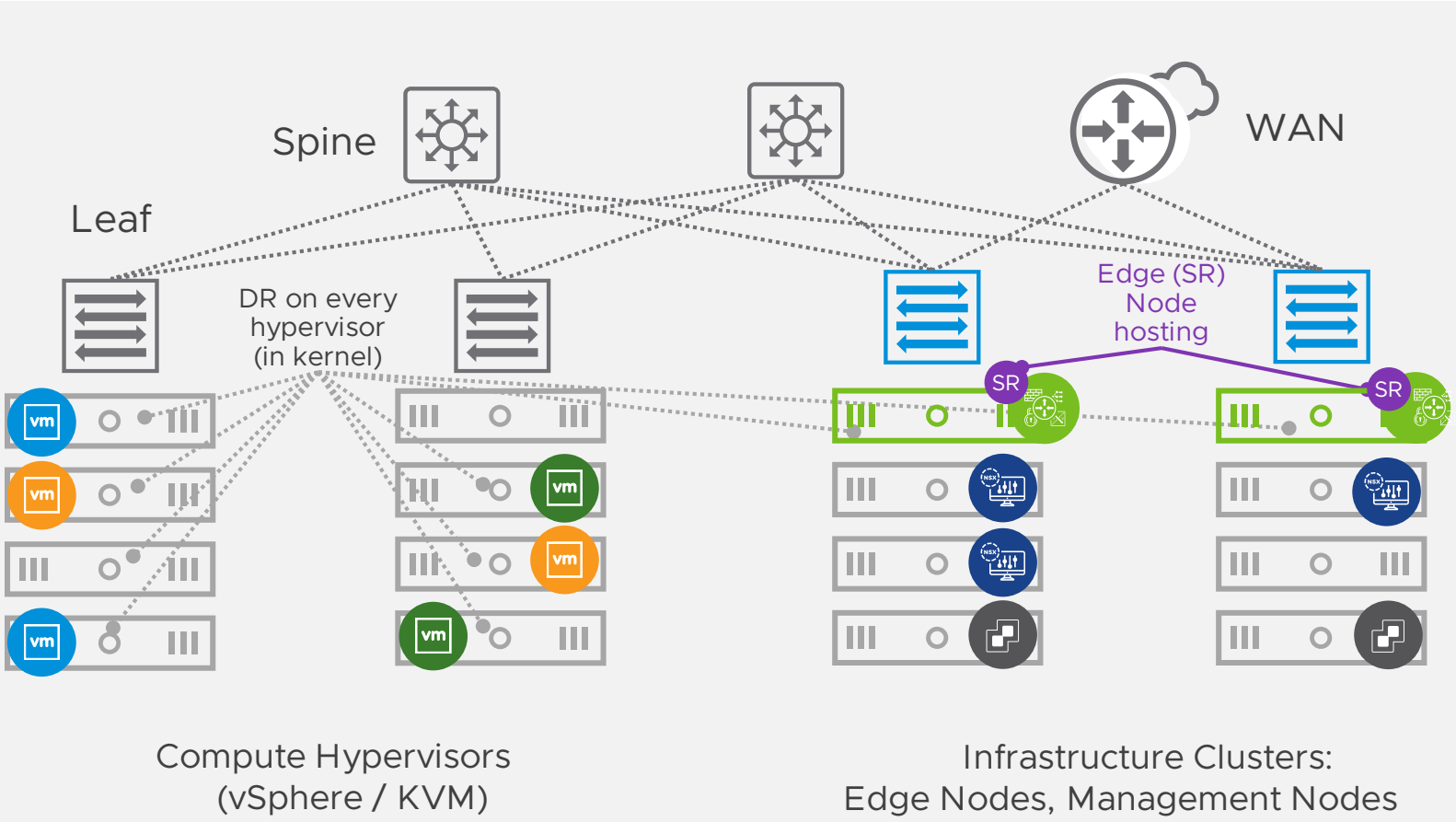
The distributed routing model for NSX instantiates Tier-0 and Tier-1 routers on every hypervisor to prevent hair-pinning.

The fully distributed routing model of NSX-T performs all of the logical routing on the source host of the workload



Logical Routing Topology

High-level view of logical routing



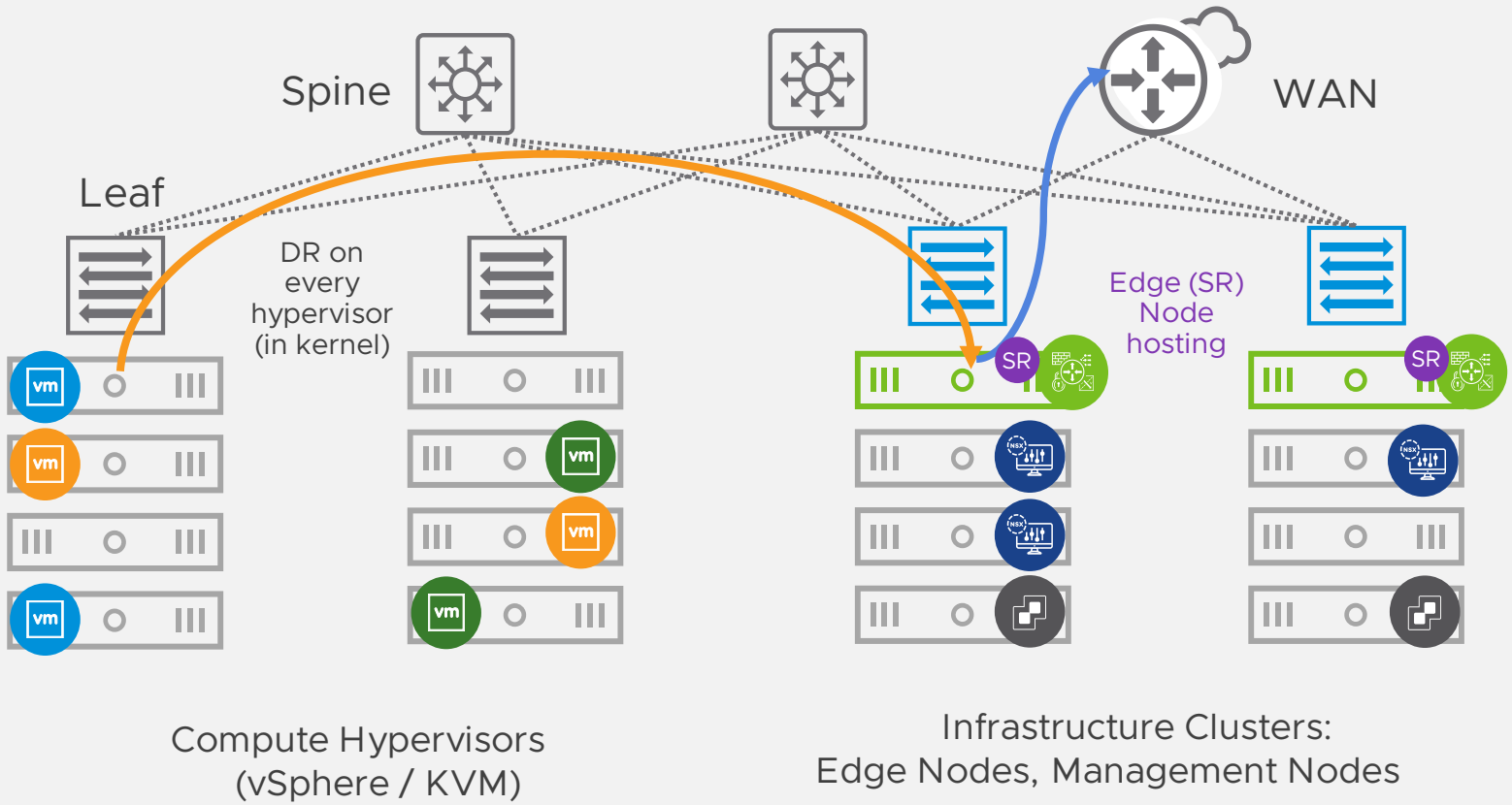
Packets enter and leave the NSX overlay via the SR component of the Edge Nodes.

The Edge Nodes are typically clustered for operational value and provide a specific demark for physical connectivity.

The workloads are hosted on the compute hypervisors and are the endpoints for the east to west communication flow. Distributed routing is performed in the kernel of these hypervisors.

Logical Routing Topology

Routing to the Edge Cluster

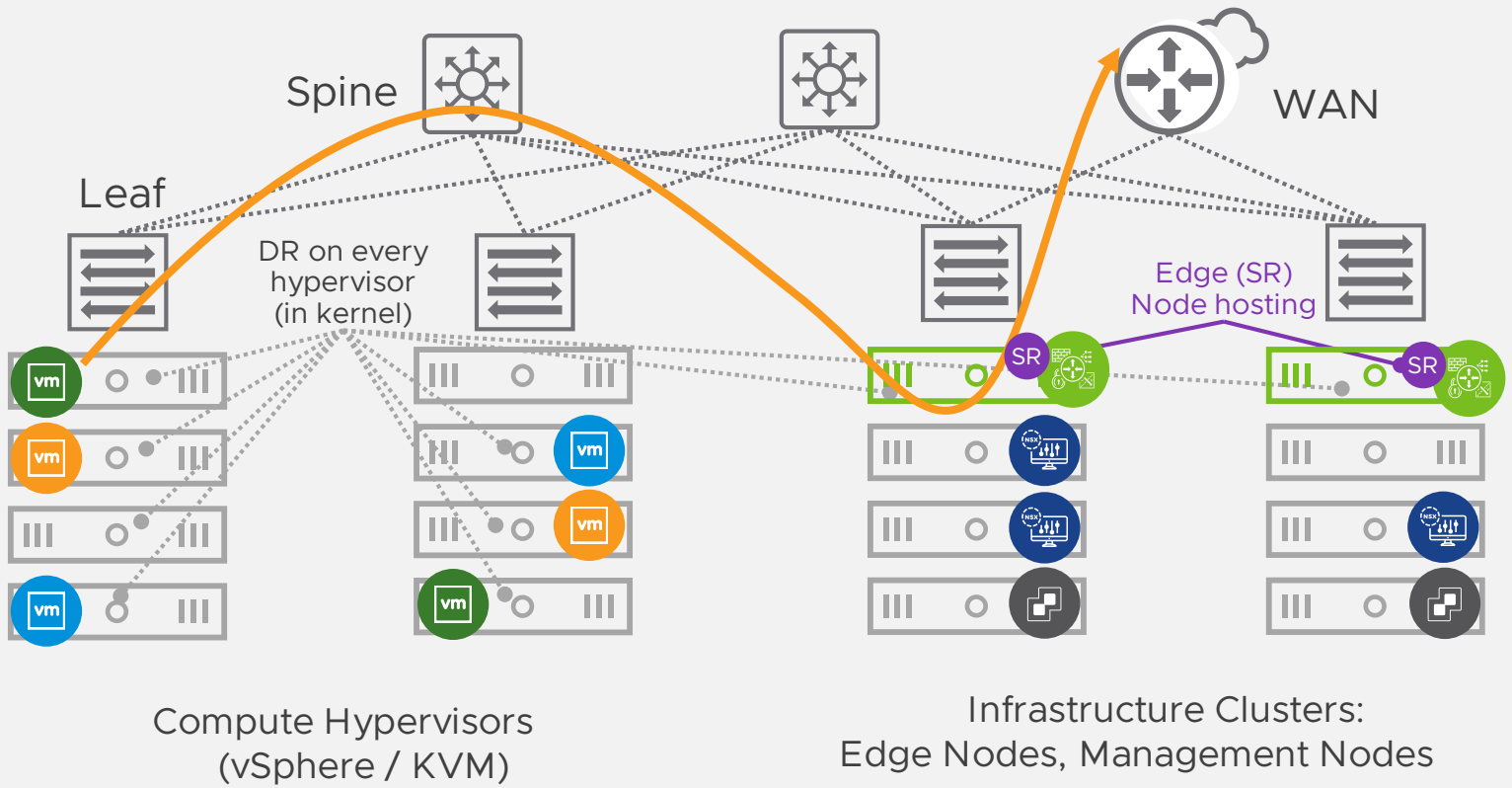


East to West switching and routing is performed in the kernel.

When a virtual workload is sending a packet is destined to an endpoint outside the overlay, routing begins on the source hypervisor and the TEPs communicate the packet to the edge nodes.

Logical Routing Topology

Routing to physical fabric



When the packet is delivered to the edge node by the TEPs of the source hypervisor to the Edge node, the packet has already been ‘routed’ to the T0 on the source host.

The packet is then routed through the T0 uplink to its SR on the Edge node. The SR has a configured connection to the outside physical router. The packet is then communicated out the SR to its adjacent physical router.

Gateway/Logical Router: Centralized Services

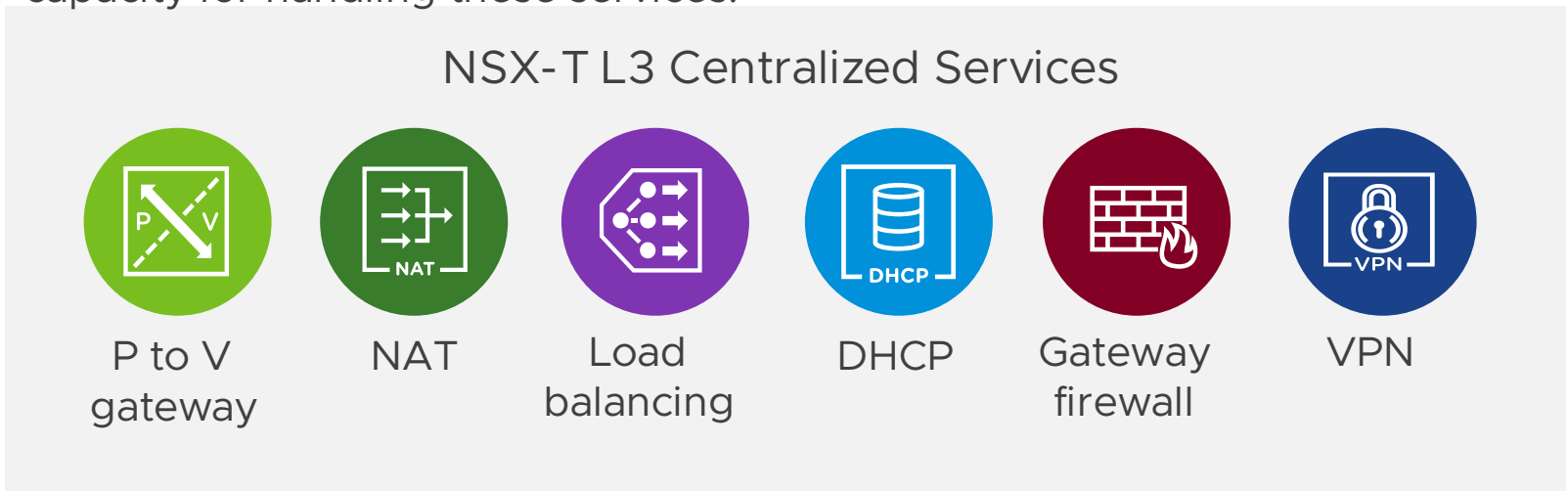
Implemented on Edge Node

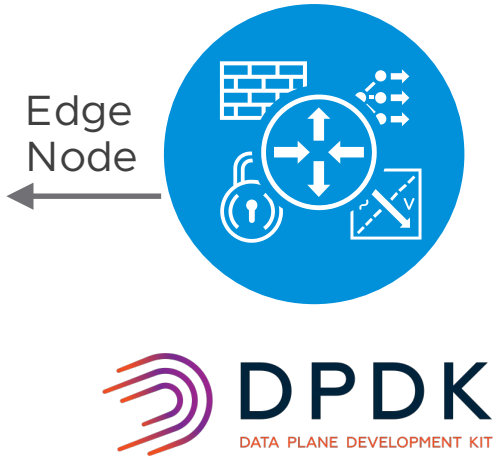
Some services are centralized as they are required to be found on a specific device. This may be due to the services stateful nature or the service is providing connectivity to a physical device.

The Edge Nodes are appliances with a pool of capacity for handling these services.

For high performance connectivity, the Edge Nodes leverage offloads such as the Data Plane Development Kit (DPDK).

In addition to maximum performance, the Edge is built for resiliency with various Active/Active and Active/Standby deployment models





DPDK
DATA PLANE DEVELOPMENT KIT

NSX-T Terminology: Edge Node

What is the Edge Node

The Edge nodes can utilize two form factors: Bare Metal and Virtual machine. The Edge cluster may only utilize a single choice. Here is short summary of the various features, sizes, deployment choices and availability models

Edge Nodes

Edge-nodes are appliances with pools of capacity for hosting any services which are not distributed

Complete placement flexibility – User can assign a service (SR) for a particular logical router to an Edge Node

Form factor choice – Virtual Machine or Bare Metal | Both OVA and ISO flavors available

Built for resiliency – A/A and A/S models available

Sizing choice – 3 sizes available (small, medium, large)

Leverages DPDK technology for fast packet processing

DPDK
DATA PLANE DEVELOPMENT KIT

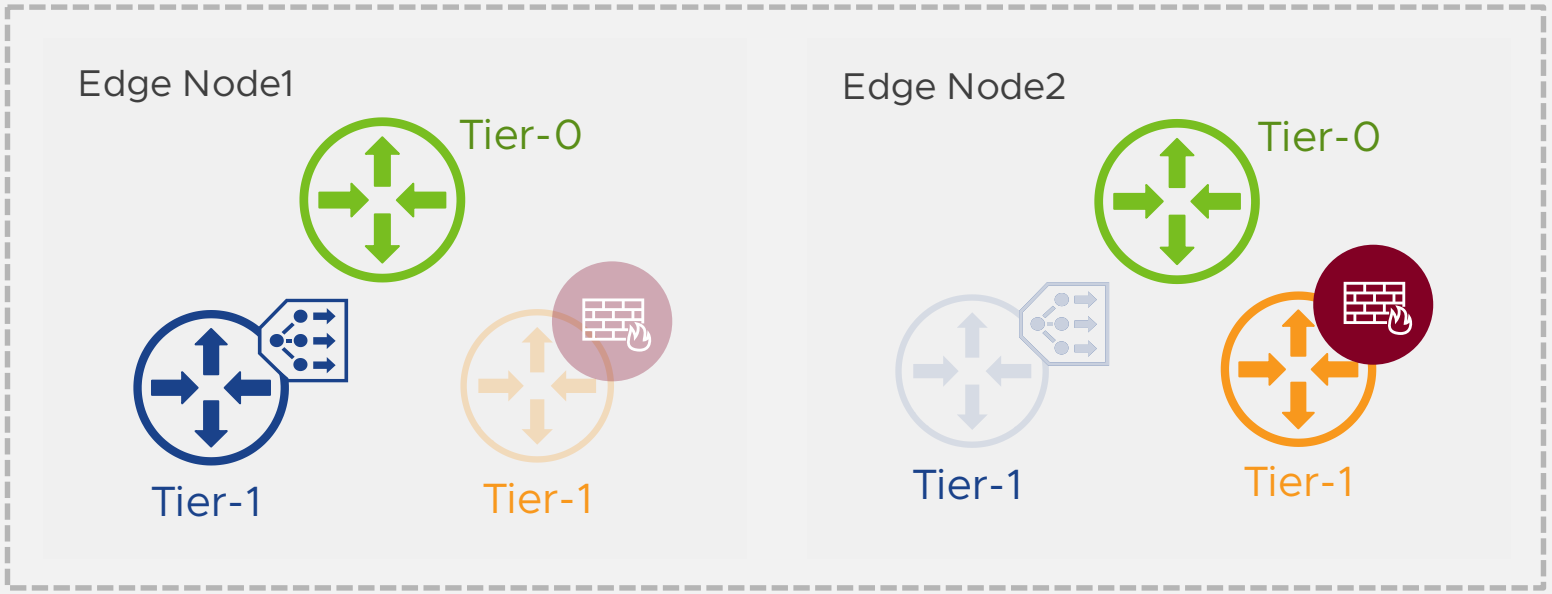
High Availability

Logical router services

NSX-T Edge nodes run in an edge-cluster to provide high-availability for Routing and Services.

Active – Active Model

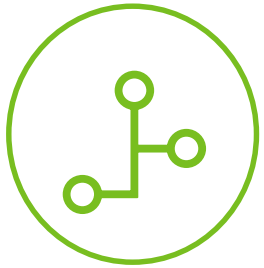
Active – Standby Model



The Edge Nodes are clustered within an Edge Cluster. This provides for highly available routing and the services hosted by the services router (SR) component.

This section introduces the value of the NSX-T Edge Cluster and edge nodes. Part 2 of the NSX-T Logical Routing: Deep Dive will discuss the specific functionality of the connectivity of NSX-T Logical routing for the DR, the Edge Nodes and the SR components

Key Takeaways



Logical Routing

NSX-T Logical Routing

- **Distributed Routing (DR)** optimizes traffic flows for East-West traffic.
- **Centralized Routing for** North-South traffic on **High performance Edge nodes**
- **DPDK Enabled Edge nodes** provide capacity to host North-South connectivity to physical and centralized services **(SRs).**
- **High availability per Logical Router –** A/A and A/S models available.