

# Open Source MANO

Enabling high performance VNFs with  
EPA & SDN Assist

Mark Beierl (Canonical)

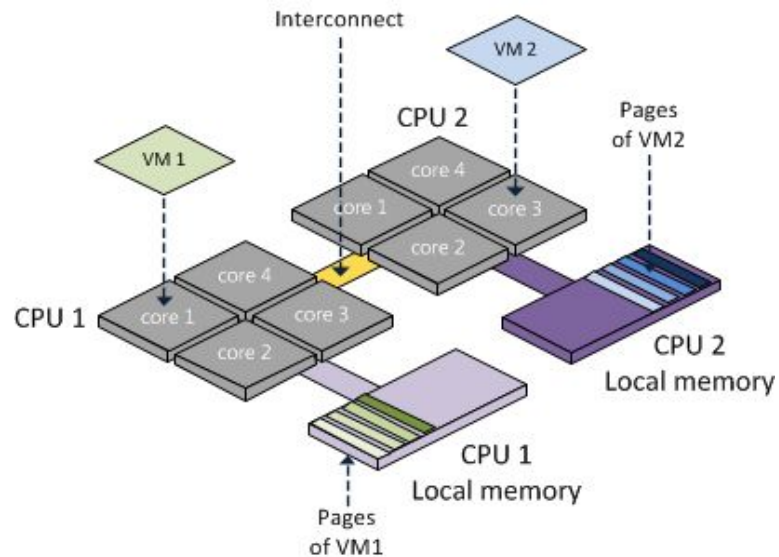
# EPA (Enhanced Platform Awareness)

- Covers a set of techniques for getting more performance
- EPA features include:
  - NUMA node placement
  - CPU Pinning
  - Huge Pages
  - SR-IOV
- OSM supports these since release 0
  - Enabled via the NF descriptor

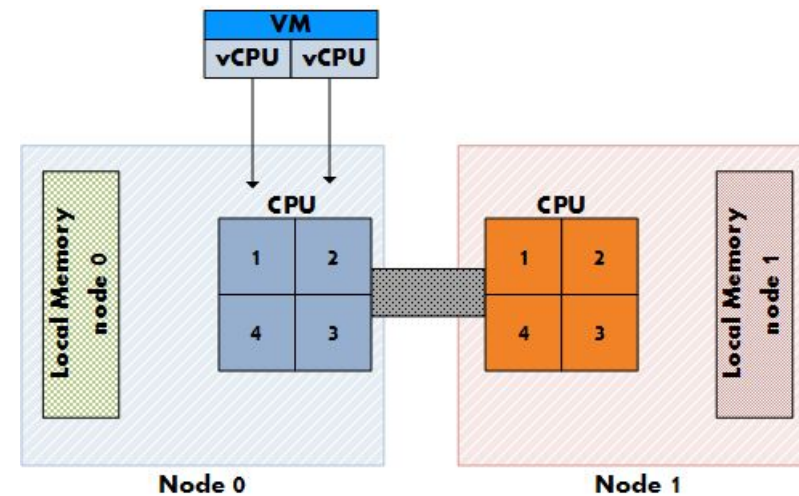


# CPU Pinning & NUMA Awareness

Most NFVI/VIMs support **CPU Pinning** and **NUMA Topology Awareness** capabilities without any need for configuration.



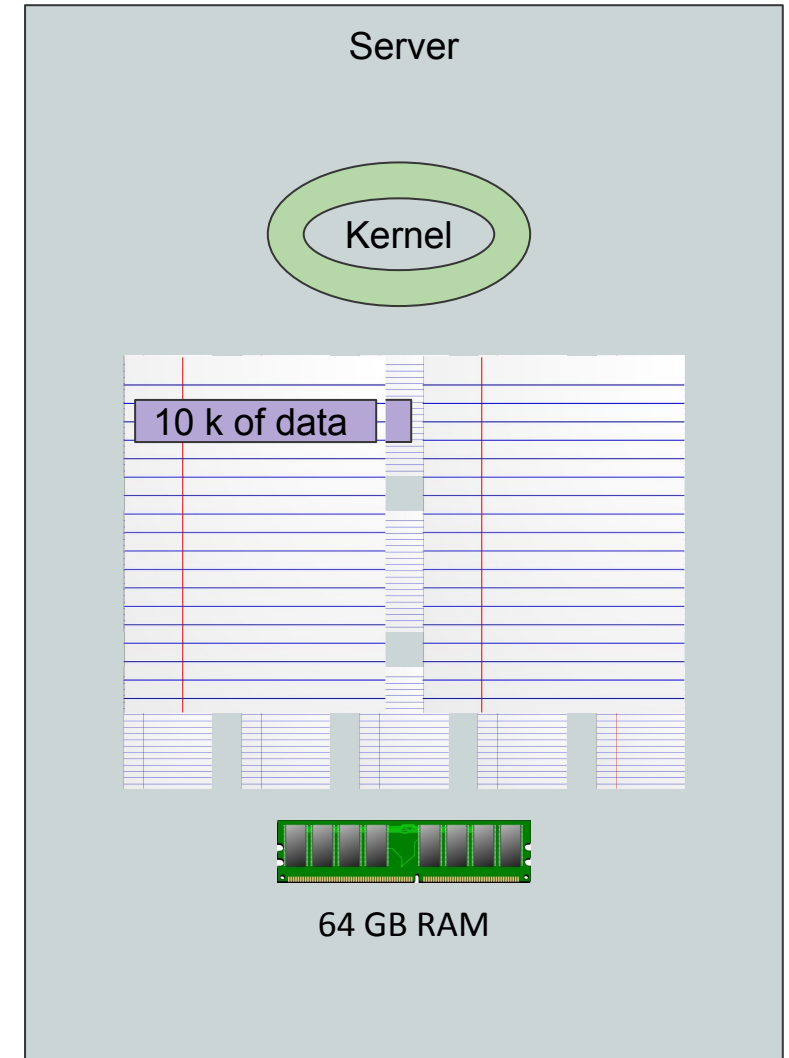
**CPU Pinning:** being able to pin a VM to specific CPUs



**NUMA Topology Awareness:** making the VM aware of the physical CPU topology

# Huge Pages

- Linux kernel maps memory in pages (4k)
- 64 GB RAM = 16,777,216 4k pages
  - Mapping of pages to physical RAM addresses happens in the Translation Lookaside Buffer (TLB)
- TLB is subset of all virtual pages
- Finding memory that is not in TLB is slow
- Recommendation: Huge Pages
  - Changes page size from 4k to something larger
  - Can result in memory waste



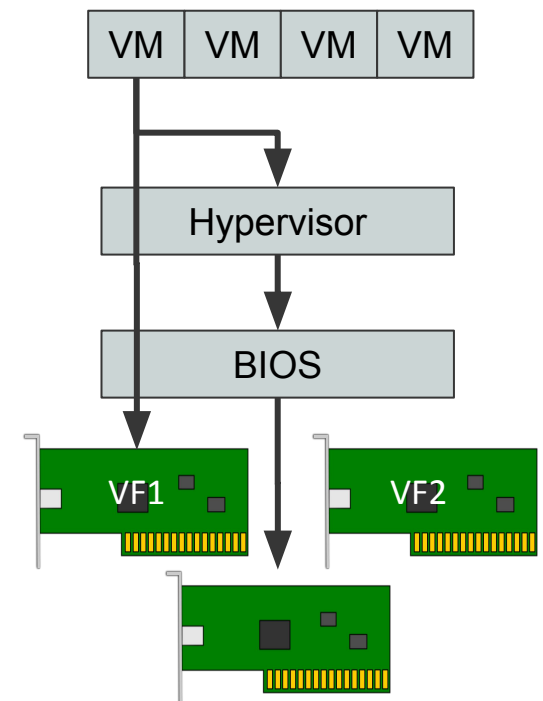
# Memory Huge Pages

**Memory Huge Pages** allows the VNFs to request RAM memory from a special pool where page sizes are bigger, enabling better performance.



**Enabling/changing Huge Pages require a node reload**, and the NFVI servers to allocate a new memory pool with bigger pages, this will not allow VMs set with normal pages to use this new pool, so it should be limited.

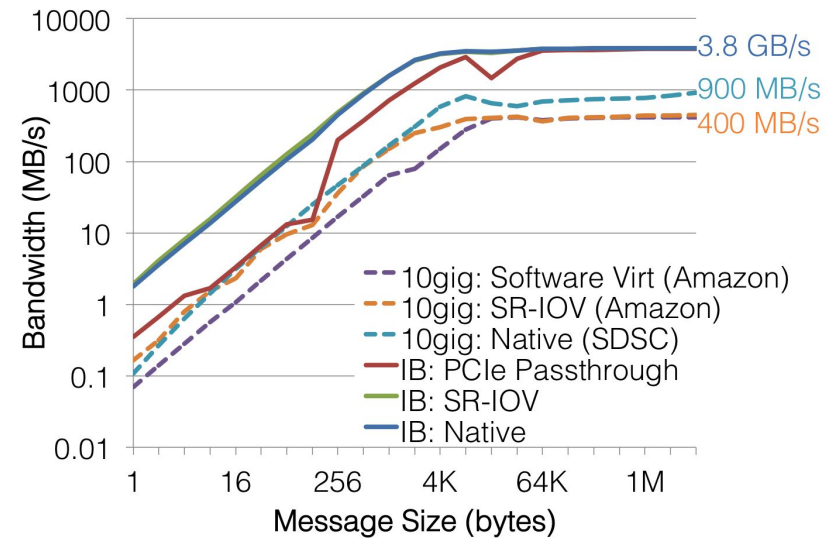
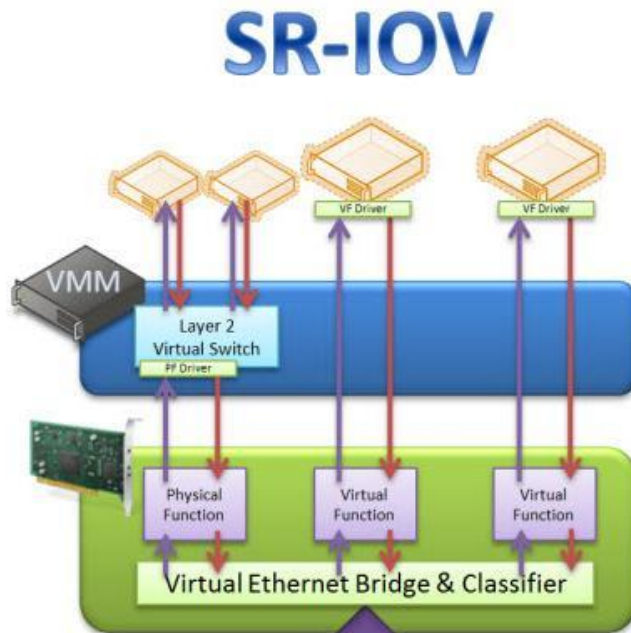
- Hypervisor must maintain map of which VM sent which packet so response goes to correct VM
- Single Root I/O Virtualization
  - Allows device to appear to be multiple separate physical PCIe devices
    - Physical Function (PF) - the primary function of the device
    - Virtual Function (VF) - associated with PF, shares physical resources of device
  - Bypasses map so lookup is not necessary





# SR-IOV

**SR-IOV** allows VNFs to have direct access to a virtualized PCI of a NIC, thus giving it better throughput.



*SDSC Early SR-IOV results*

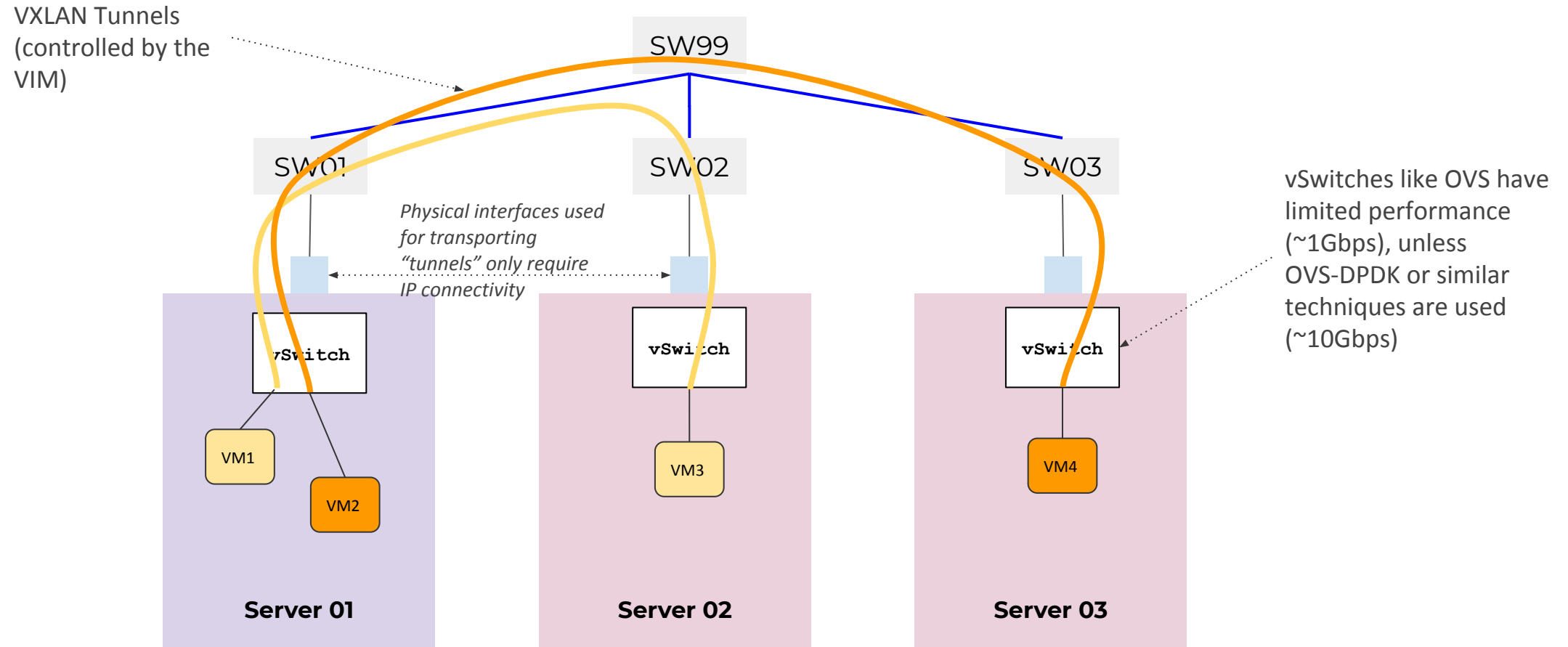
Enabling SR-IOV requires a node reload for reconfiguration of the IOMMU virtualization mode. It also requires physical interfaces to be dedicated to this feature.



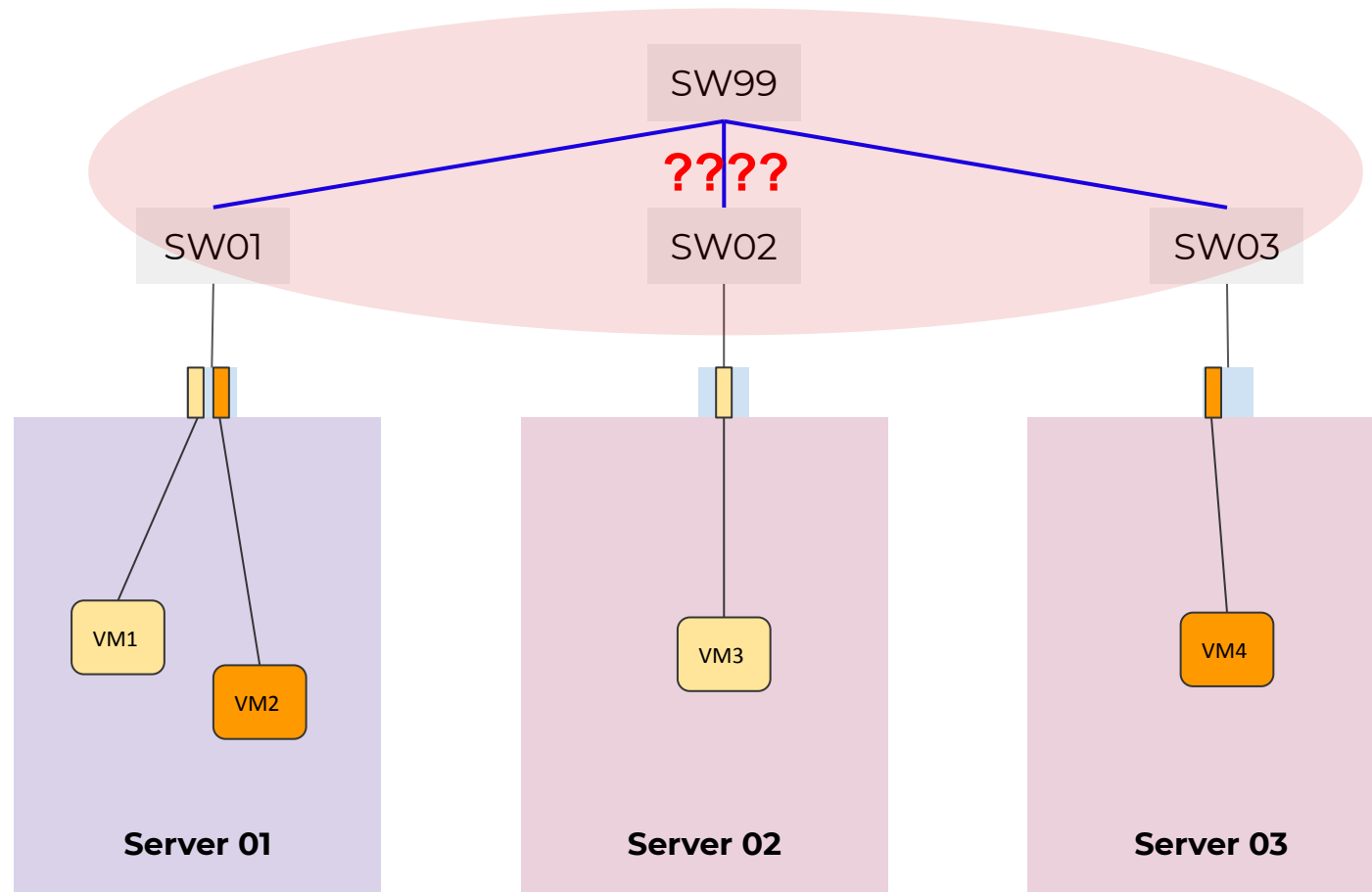
- OSM supports EPA enablement in Descriptors
- However
  - OSM does not change OpenStack server configuration
  - SR-IOV must already be enabled in the compute node
  - Huge pages must be enabled in the compute node
  - Descriptor must be made aware of NUMA topology
  - OSM does not know:
    - Number of Numa Nodes
    - Number of CPU cores
    - Number of CPUs/threads per core
- All this must be known before launching a service

- SDN Controller
  - Separates the network control functions from forwarding functions
  - Creates overlays that exist on top of physical network
  - Manages flow control of switches “under” the overlay
- OSM currently supports:
  - Arista Cloudvision
  - Floodlight OpenFlow
  - Juniper Contrail
  - OpenDaylight (ODL) OpenFlow
  - ONOS (OpenFlow or VPLS)

# Using Virtual Interfaces (VIRTIO)

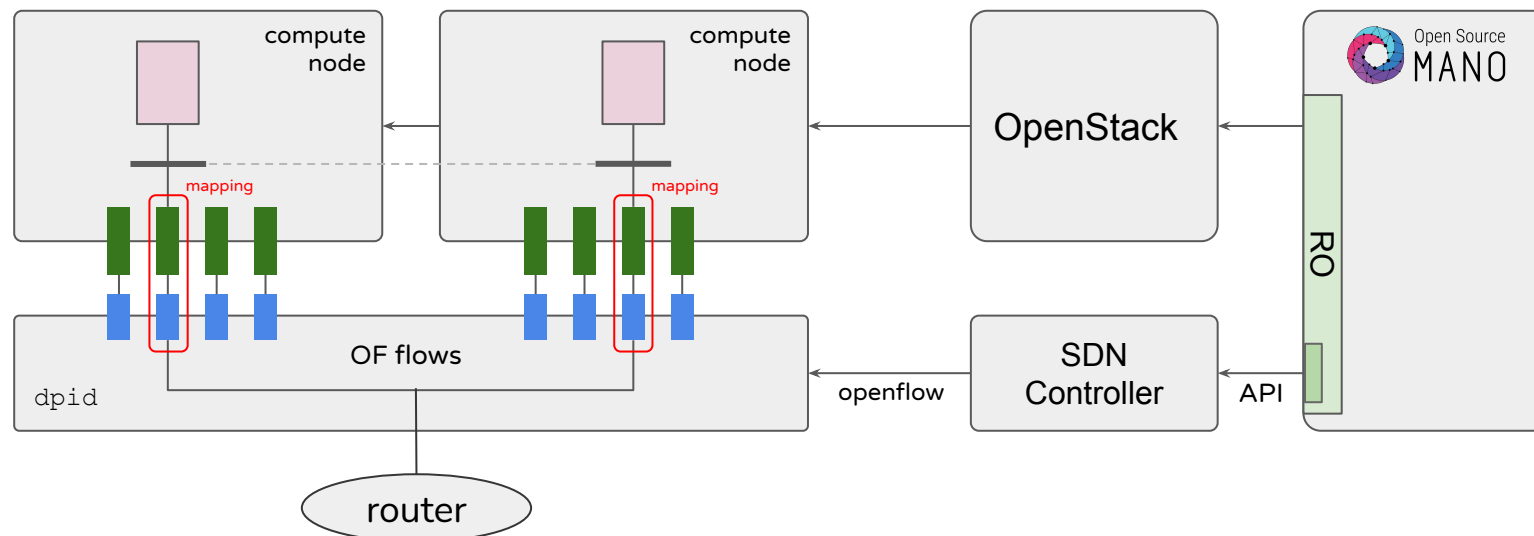


# Using Physical Interfaces (SR-IOV/PASSTHROUGH)



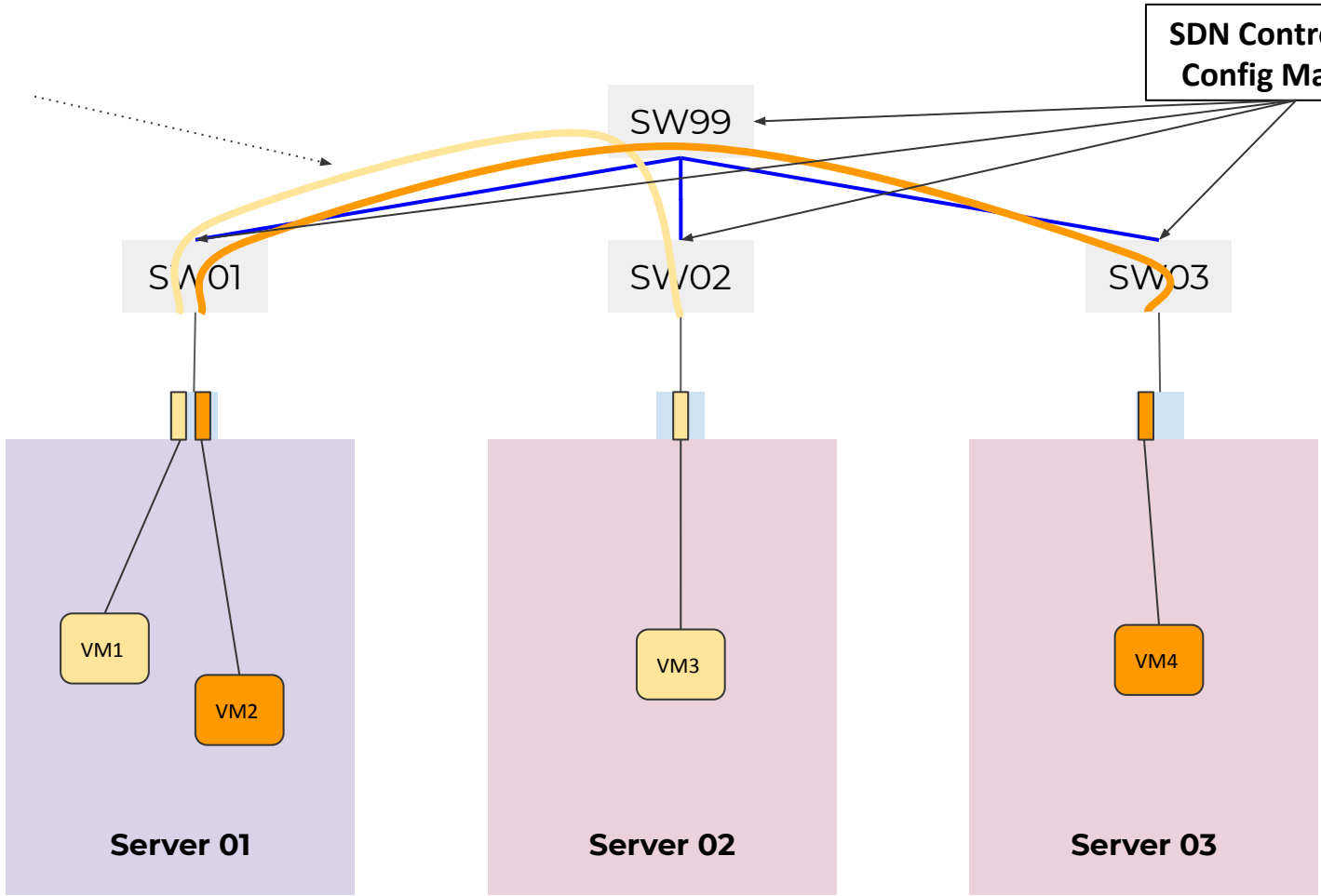
SR-IOV and Passthrough features expose the instance directly to the physical NIC, so **who takes care of the end-to-end connectivity?**

1. OSM orchestrates SR-IOV or Passthrough
  - Proper assignment of I/O physical interfaces to the VM (PFs or VFs = Physical or Virtual Functions)
1. OSM SDN Assist gives the ability to create L2 connections between VFs
  - Interconnecting VMs
  - Attaching external traffic sources



# SDN Assist

VXLAN, OpenFlow,  
VLAN, etc.  
(depends on the SDN  
Controller or Manager)



OSM's SDN Assist feature takes care of the "underlay" connectivity **whenever it sees VLDs with SR-IOV or PASSTHROUGH ports that need to connect between each other.**

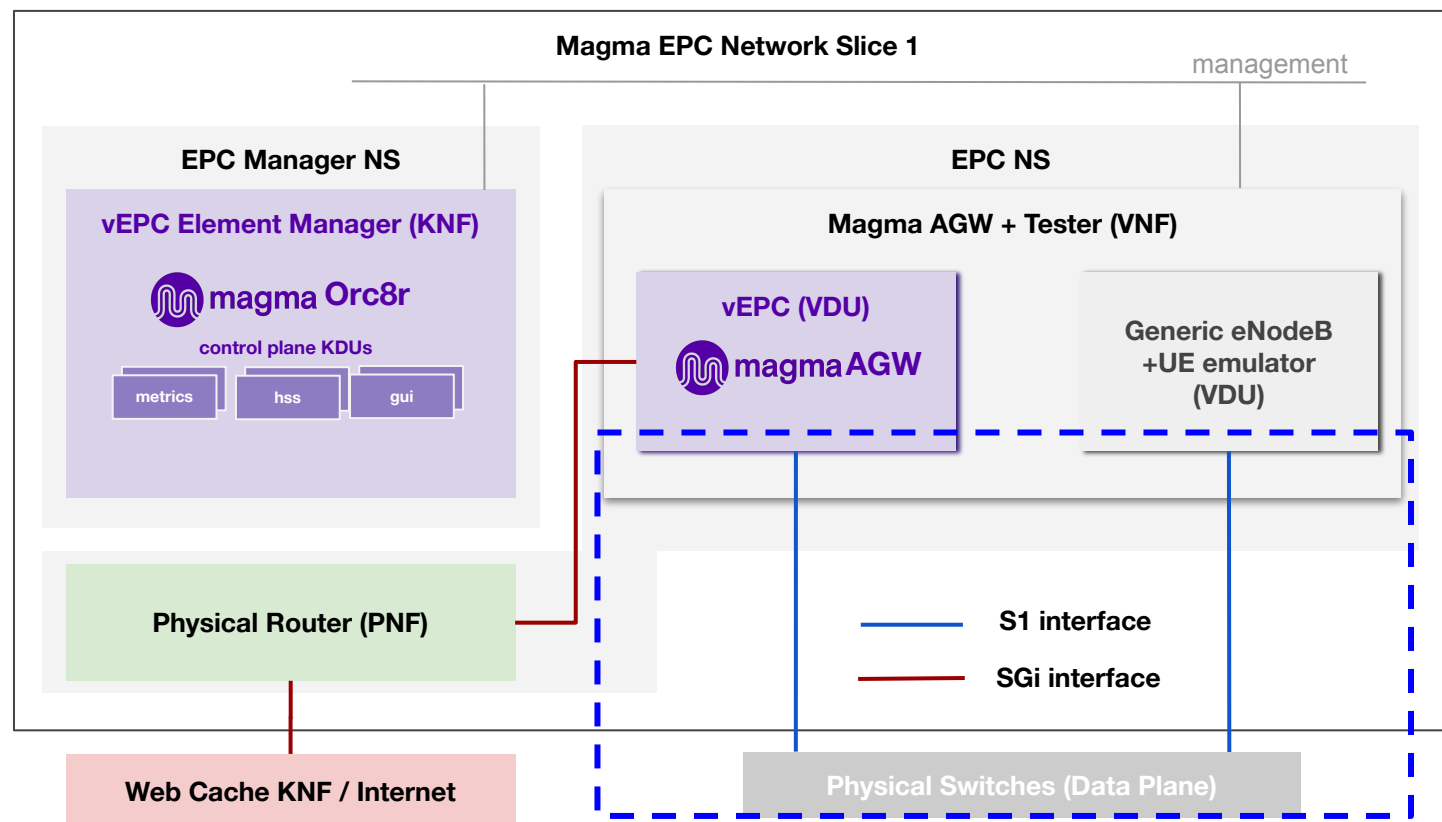
\* Supported as of REL7.1.0 → ONOS, Arista, Open Daylight and Floodlight

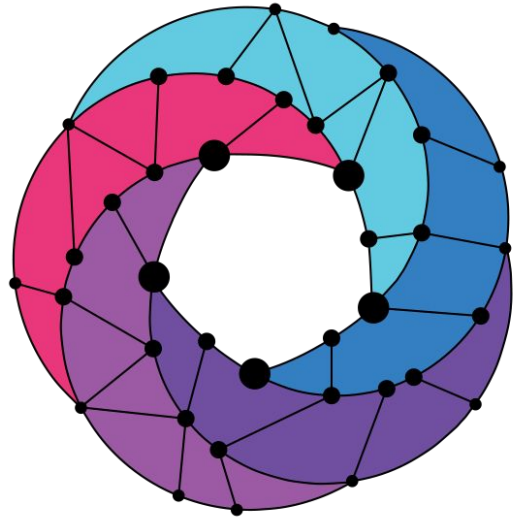


- Like EPA, OSM does not manage SDN Controller or OpenStack
  - Compatible SDNC must be installed
  - Must be reachable from OSM
- Some plugins need additional information
  - Port mapping files for PCI ports
- VIM account must have admin privileges
  - Needs get PCI information

# EPA in our Network Service

In our example, we can configure the S1 data interface, currently using VIRTIO drivers (OVS/VxLAN) to use SR-IOV instead. We can also set the descriptor to request CPU Pinning, memory Huge Pages, and stick the VDUs to a single NUMA node. Today in OSM, all these optimizations are applied automatically when selecting SR-IOV in one of the interfaces, in order to match the packet processing capabilities that the direct connection to the NIC will allow for.





# Open Source MANO

Find us at:

[osm.etsi.org](https://osm.etsi.org)  
[osm.etsi.org/wikipub](https://osm.etsi.org/wikipub)