



KubeVirt

SUMMIT

EUROPE

CLOUD NATIVE THEATER

Breaking the Performance Barrier: High-Performance AI Storage Virtualization with KubeVirt

Jian Li, *SK Telecom*

Yves Weisser, *NetApp*



KubeVirt

SUMMIT
EUROPE

CLOUD NATIVE THEATER



Jian Li

Principal Engineer
SK Telecom



Yves Weisser

Solution Architect
NetApp

Introduction

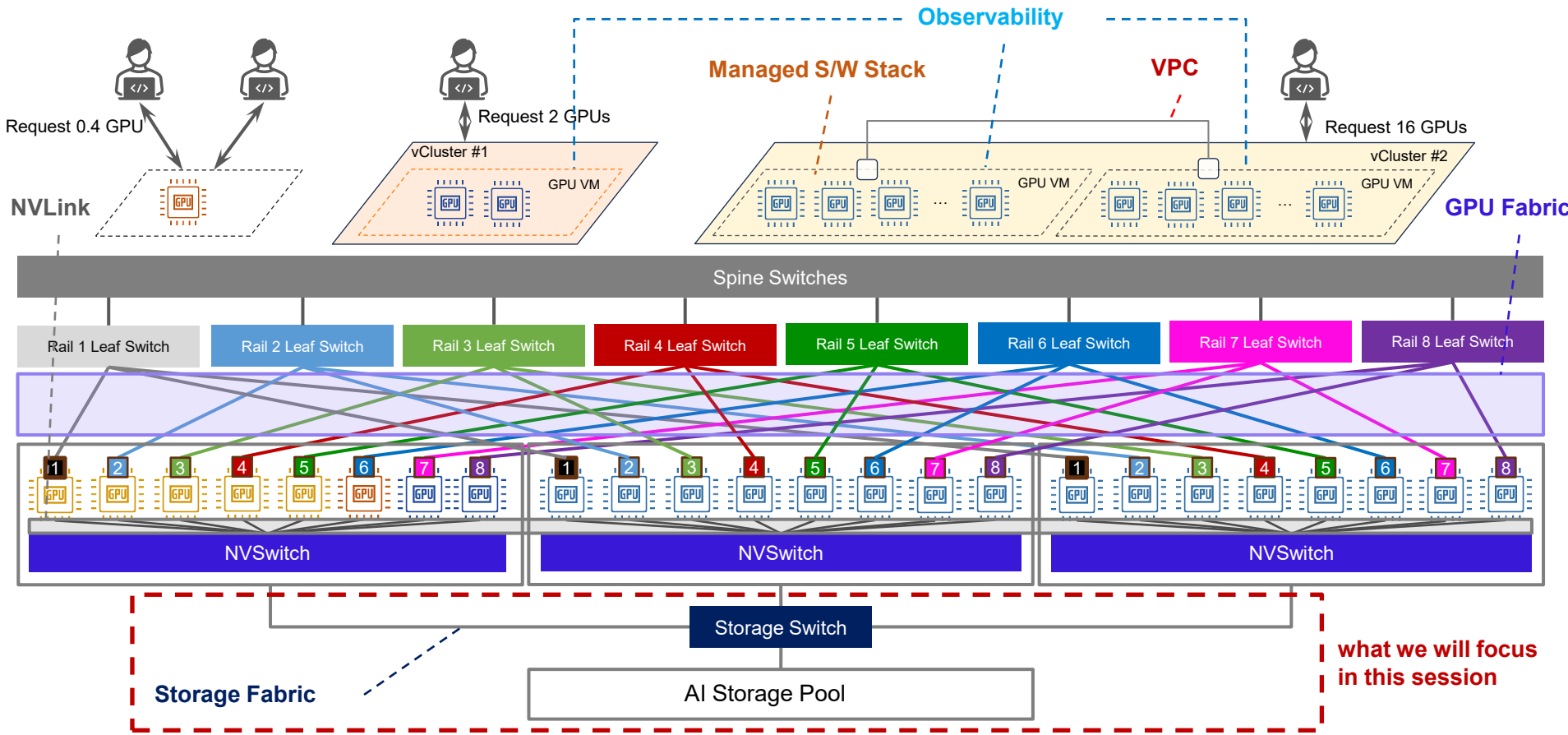
AI Storage Virtualization

AI Storage Integration with KubeVirt

Performance Evaluation Results

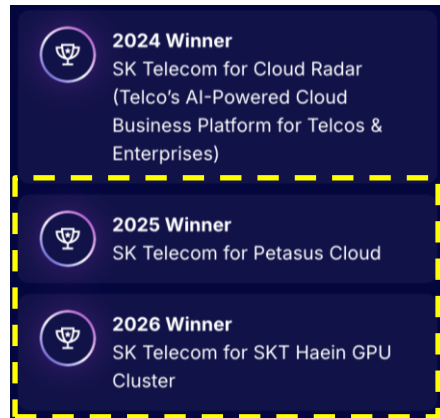
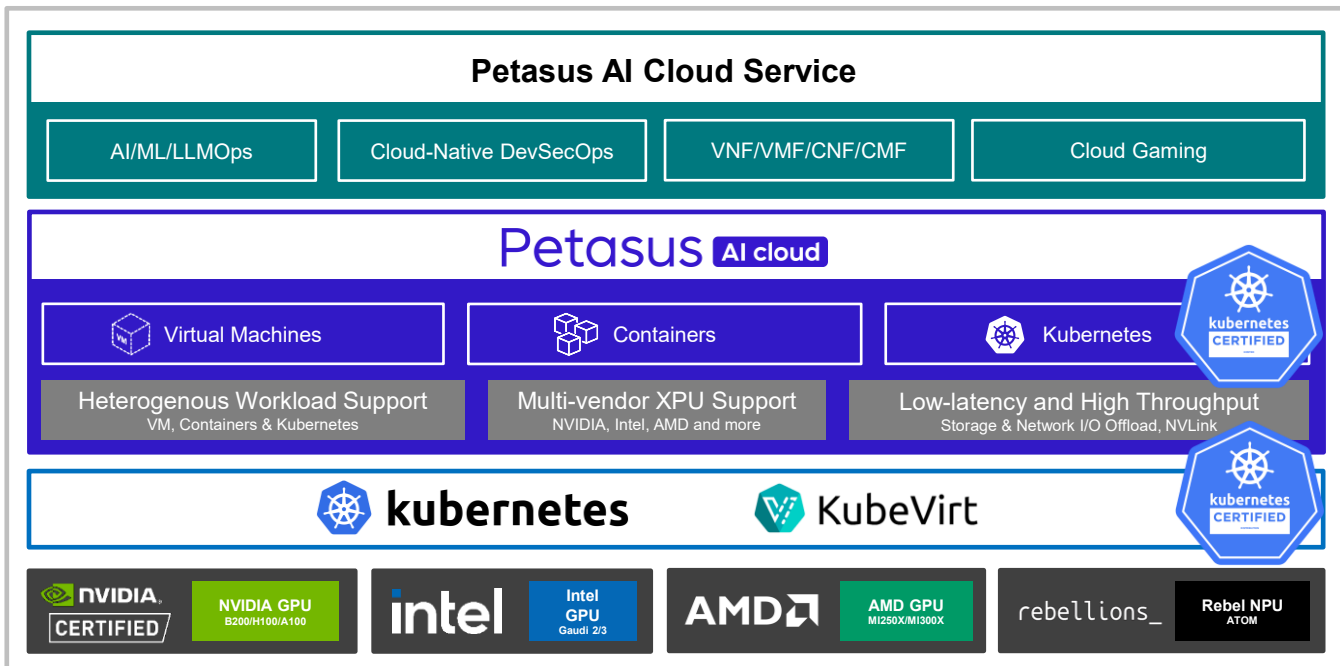
Conclusion

Virtualizing AI Infrastructure

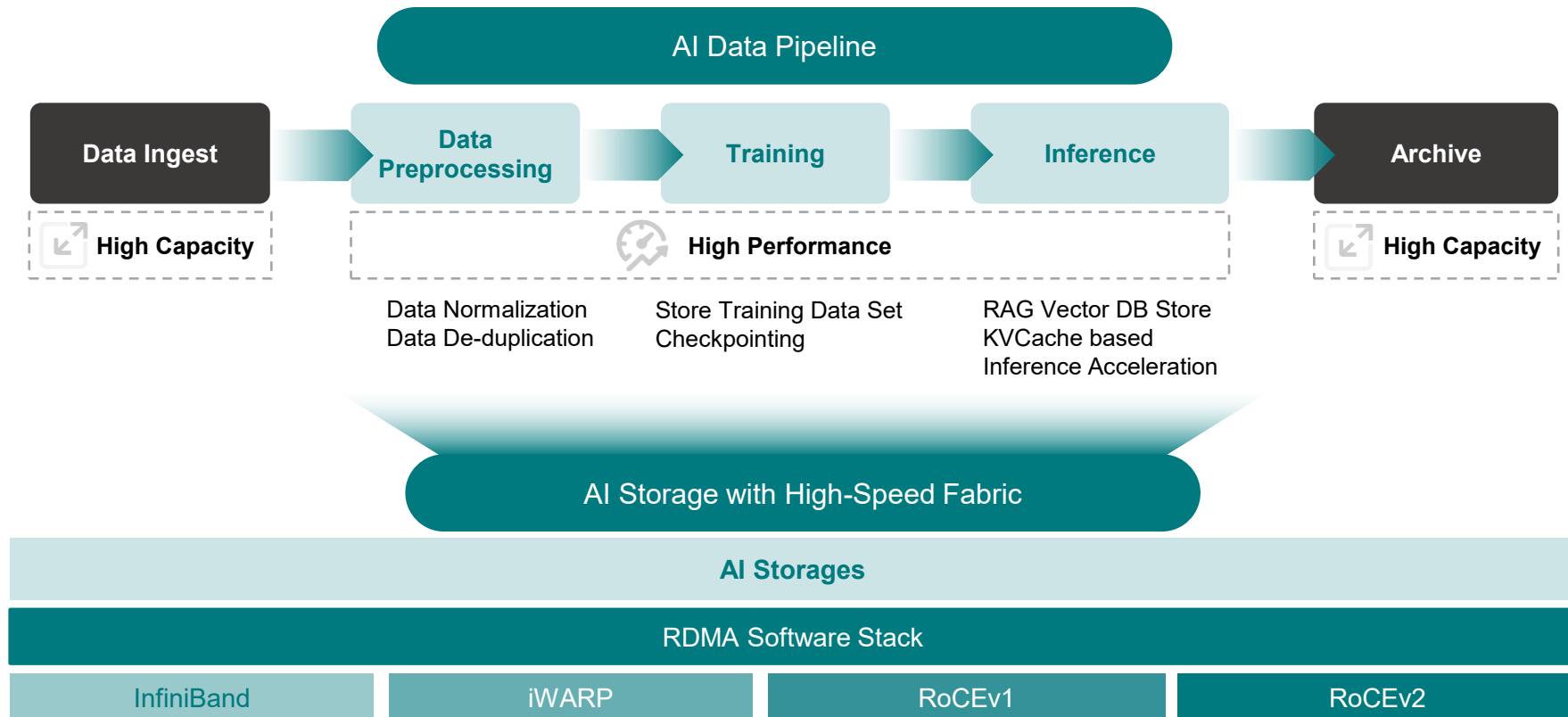


AI Infra. Virtualization with Petasus AI Cloud

Focuses on providing GPUaaS with an emphasis on high performance and cost efficiency



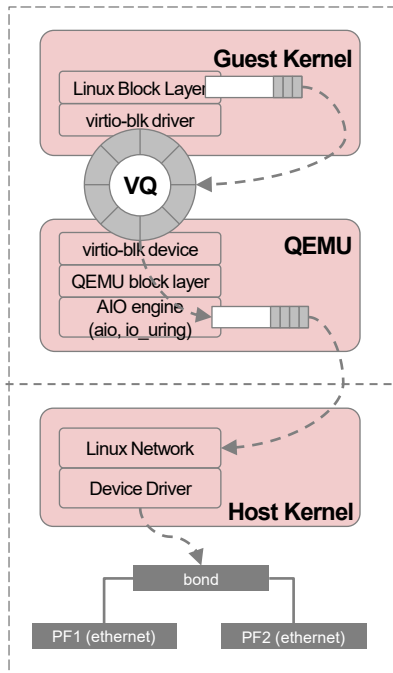
AI Storage Virtualization Overview



AI Storage Virtualization (1/3)

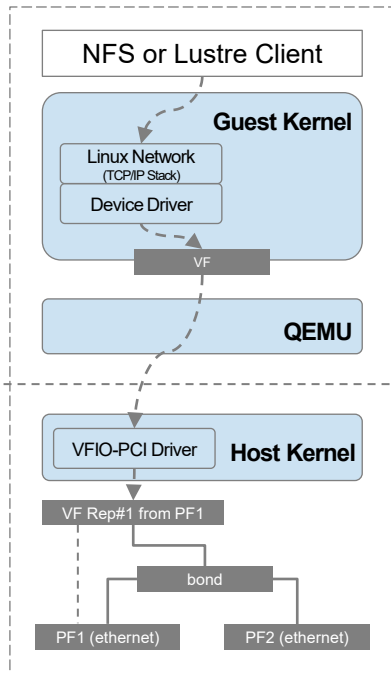
CSI based NVMe-over-TCP

Pros: Easy to integrate with K8S
Cons: perf. degradation

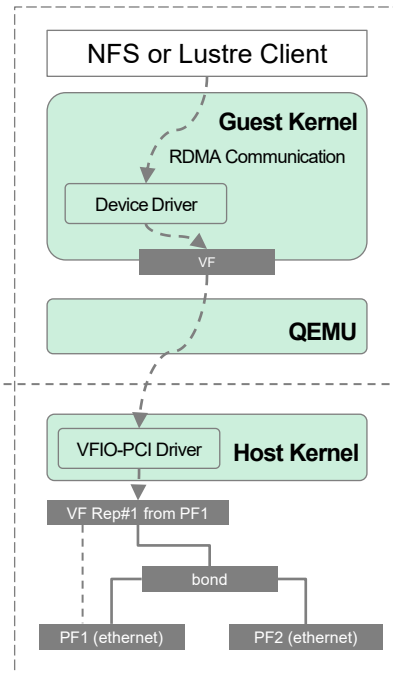


VF LAG based Native NFS/Lustre over TCP VF LAG based Native NFS/Lustre over RDMA

Pros: High I/O throughput
Network optimization is not required
Cons: Need to manage additional S/W* in VM
Additional config** for realizing HA



Pros: BareMetal grade performance
Cons: Need to manage additional S/W* in VM
Additional config** for realizing HA
Additional optimizations*** on RDMA



Can be resolved using Petasus

Host User Space
Host Kernel Space

*) Lustre and multipath NFS Kernel Module

**) SR-IOV VF LAG based HA

***)) Need to enforce ECN, PFC for NIC and switch

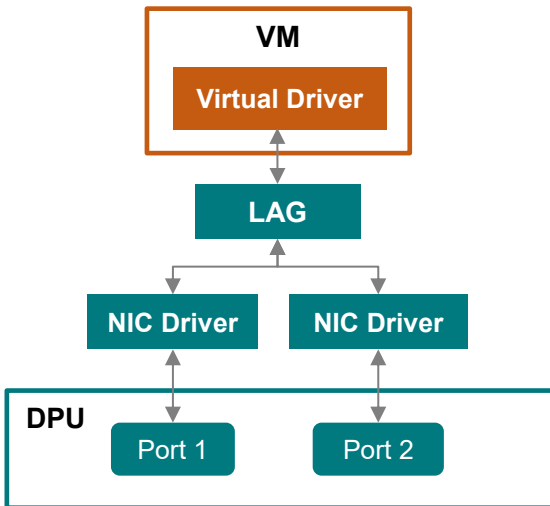
AI Storage Virtualization (2/3)

Options for Link Aggregation (LAG) to achieve HA and high-performance

SW LAG by Host CPU

Pros: applicable on any NICs
Cons: limited performance

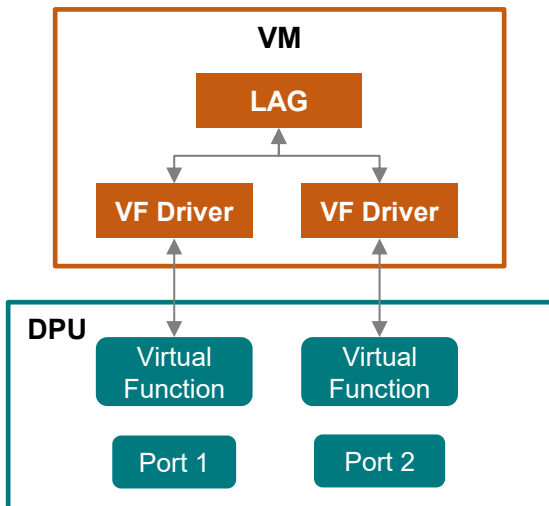
Hypervisor



SW LAG by VM CPU

Pros: improved performance
Cons: only for HA, not for aggregation

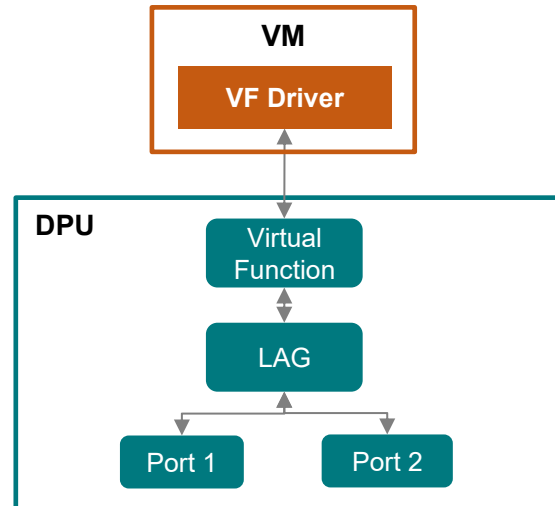
Hypervisor



HW LAG by VM CPU

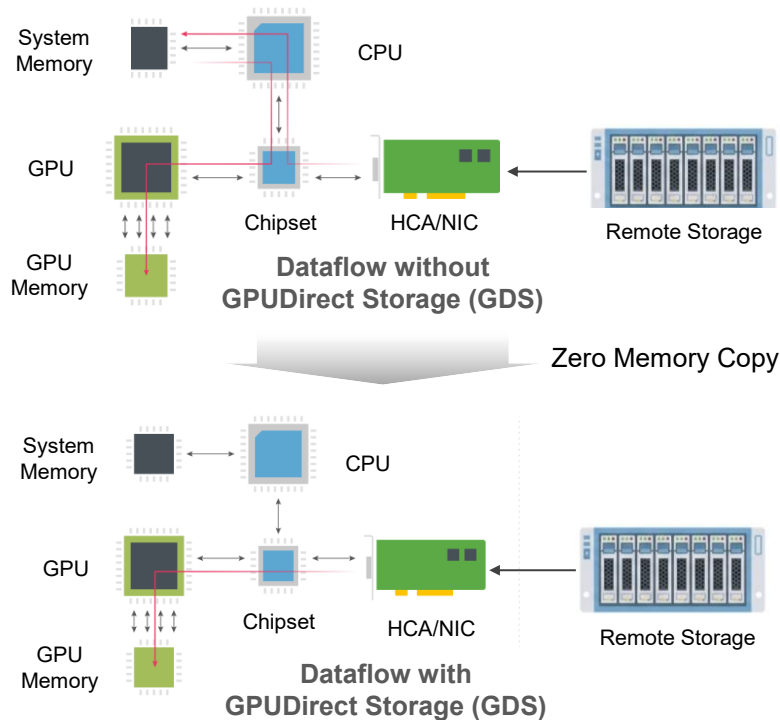
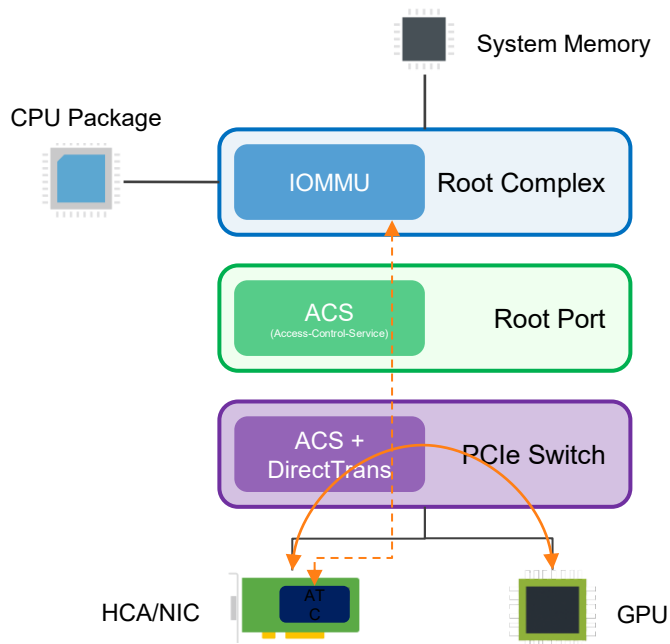
Pros: line-rate performance with aggregation
Cons: only available on ConnectX NICs

Hypervisor



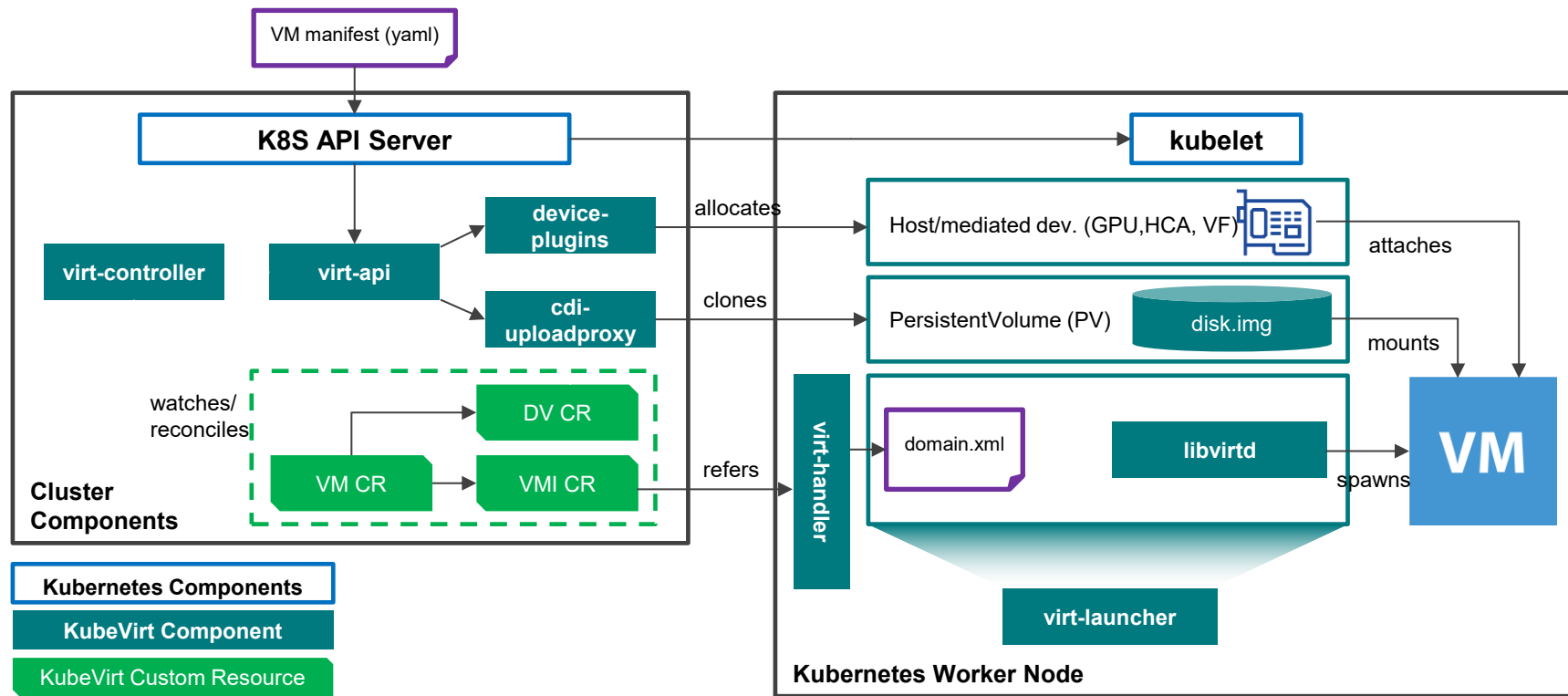
AI Storage Virtualization (3/3)

ATS (@HCA/NIC) with Direct Translation (@PCIe Switch) enables GDS for VM

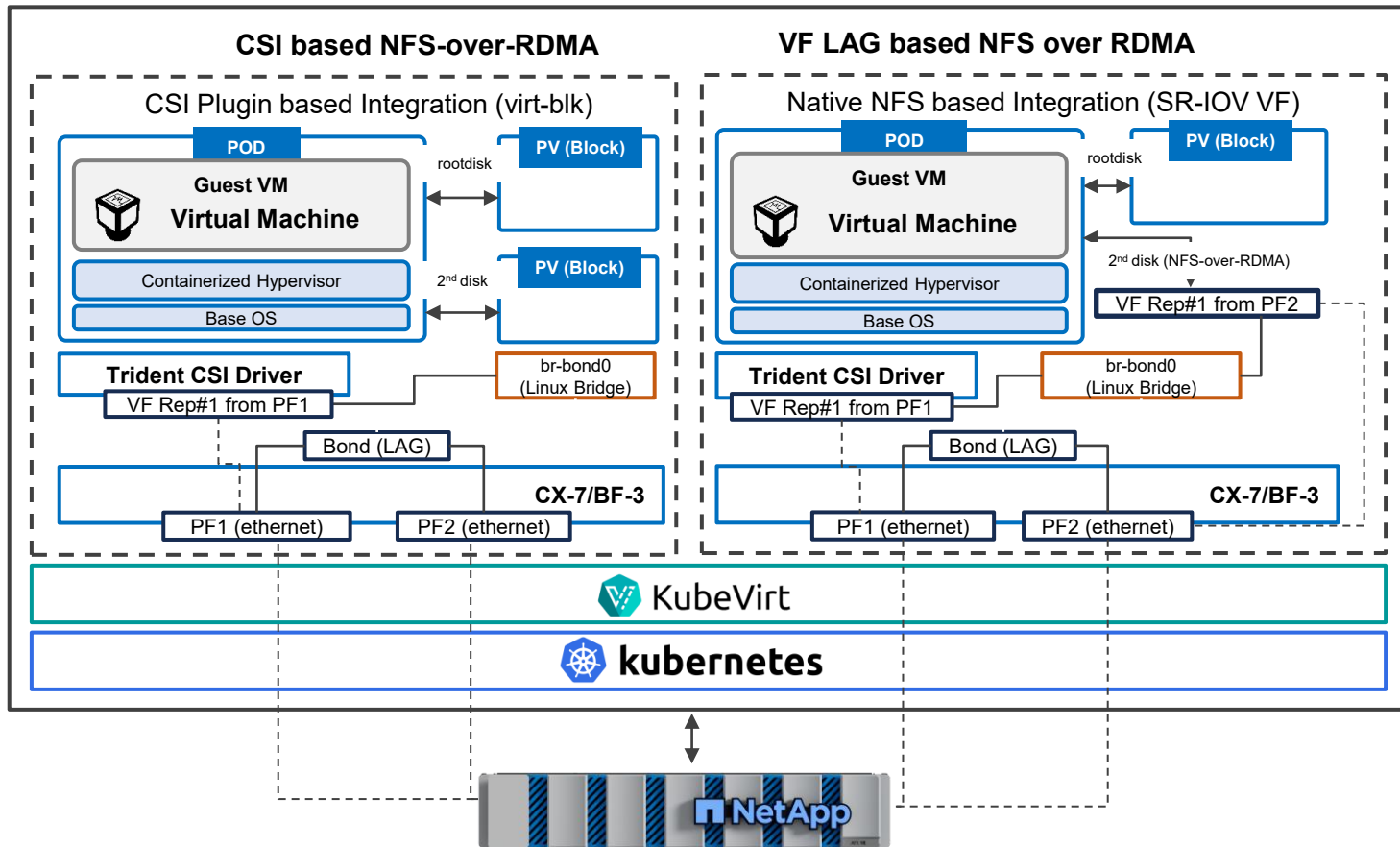


Device Attachment at KubeVirt

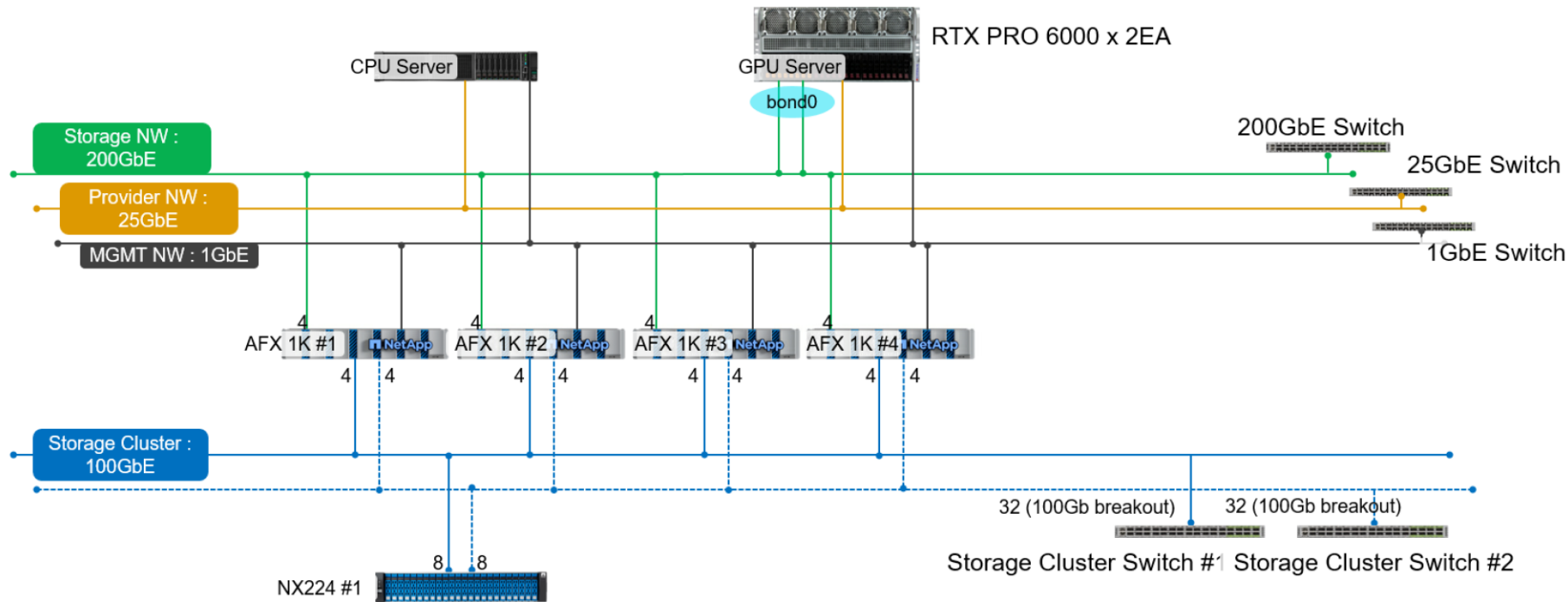
Uses **Kubernetes-native** way to **manage** VM resource (CR) and **provision** containerized VM



KubeVirt Integration Architecture



Experiment Network Topology



Thank you