TECHNICAL WHITE PAPER
Written in 2021
Updated September 2024

Secure Networking for Multi-Tenant High Performance Computing and Machine Learning

Reference architecture and performance study

vmware®
by Broadcom

Contents

Introduction

High performance computing (HPC) environments are crucial to innovation today. They drive drug discovery, electronic design automation, digital movie rendering, and deep learning—to name just a few of many applications. At the same time, an ever-growing need for security is driving HPC environments from the physical world to the virtual.

Traditional bare-metal HPC systems are not able to meet the requirements of dynamic sharing and isolation of resources and thus are incapable of supporting secure multi-tenancy. Aging infrastructures escalate security concerns. Networking security is one of the key benefits that virtualization offers to the enterprise. Virtualized HPC environments also offer significant value. Here, IT can leverage multi-tenancy to get more out of hardware resources and completely separate research projects between users to keep files and data private.

Though public clouds make an array of security policies available, there are still challenges related to security and management flexibility. This is apparent in situations such as clinical genomic sequencing, chip design, or other sensitive areas of research that undergo regulatory compliance and require the highest security. To address these challenges, modern HPC environments need a software-defined networking solution that brings strong security and streamlines security operations.

In this paper, we leverage VMware Cloud Foundation (VCF) and one of its core components, NSX-T Data Center, for HPC workloads. We present a multi-tenant networking architecture and evaluate the performance of HPC applications paired with different NSX-T features, including micro-segmentation with distributed firewall (DFW), encapsulation with GENEVE overlay, and the NSX enhanced data path (ENS)/network stack. Finally, we offer a list of best practices.

HPC workload classification

In broad terms, HPC workloads fall into three categories:

- **High-throughput workloads:** These require no communication or synchronization between tasks that run in parallel. Typical high-throughput applications include Monte Carlo simulations in finance, genome sequence searching in bioinformatics, video rendering in movies, and other parameter-variation simulations. Here, a single program can have hundreds, thousands, or even millions of executions with varying inputs.
- **Parallel-distributed workloads:** These often perform sustained and intense communication within a single job, making their performance sensitive to interconnect bandwidth and latency. For such applications, you can use remote direct memory access (RDMA) to transfer data directly to or from application memory without involving the operating system, thus enabling high bandwidth and low latency. The RDMA design makes it a popular option for HPC systems running parallel distributed workloads. RDMA interconnects can be configured in three ways in VMware virtualized environments: DirectPath I/O, single-root I/O virtualization (SR-IOV), and paravirtualized RDMA (PVRDMA).
- **Machine learning/deep learning (ML/DL):** ML/DL is a type of HPC workload that shares many characteristics with high-throughput workloads. Only large-scale training requires running in a distributed way with multi-node CPUs or accelerators like GPUs. In VMware virtualized environments, admins can configure GPU compute accelerators in three ways: DirectPath I/O, NVIDIA vGPU, and vSphere Bitfusion. vSphere Bitfusion

is a feature that supports remote access to GPUs via networking from VMs anywhere in the data center (read Announcing vSphere Bitfusion – Elastic Infrastructure for AI/ML Workloads for more details).

In this paper, our performance studies focus on two of the above scenarios, both of which depend on Ethernet networks:

1. High-throughput workloads that have I/O to a network file system (NFS)
2. Machine learning workloads running with vSphere Bitfusion

We chose these scenarios because data centers often rely on Ethernet-based virtual networking for administrative traffic, login sessions, NFS I/O traffic, Bitfusion remote networking data traffic, and much more. That's why we set out to demonstrate that virtual networking with advanced security and VMware NSX-T Data Center can support multi-tenancy in an Ethernet environment.

Multi-tenant reference architecture

VCF is a unified, software-defined data center (SDDC) platform that brings together VMware ESXi, vCenter Server, vSAN, NSX-T Data Center, and vRealize Lifecycle Manager. VMware NSX-T Data Center is the network virtualization solution for SDDC and delivers networking and security entirely in software, abstracted from the underlying physical infrastructure. With optimized networking and security policies, NSX-T Data Center can provide secure, multi-tenant virtualized HPC environments that allow organizations to increase overall hardware utilization.

Figure 1 shows an overview of NSX-T Data Center cohesively managing on-premises and off-premises HPC/ML cluster networking, regardless of whether a cluster consists of VMs, containers, or even bare-metal hosts. NSX-T Data Center abstracts network operations from the underlying physical networks onto a virtualized layer, such that the switches, routing, firewalls, and load balancers are distributed across the entire environment.

Figure 1. Unified networking infrastructure for multi-cloud HPC/ML

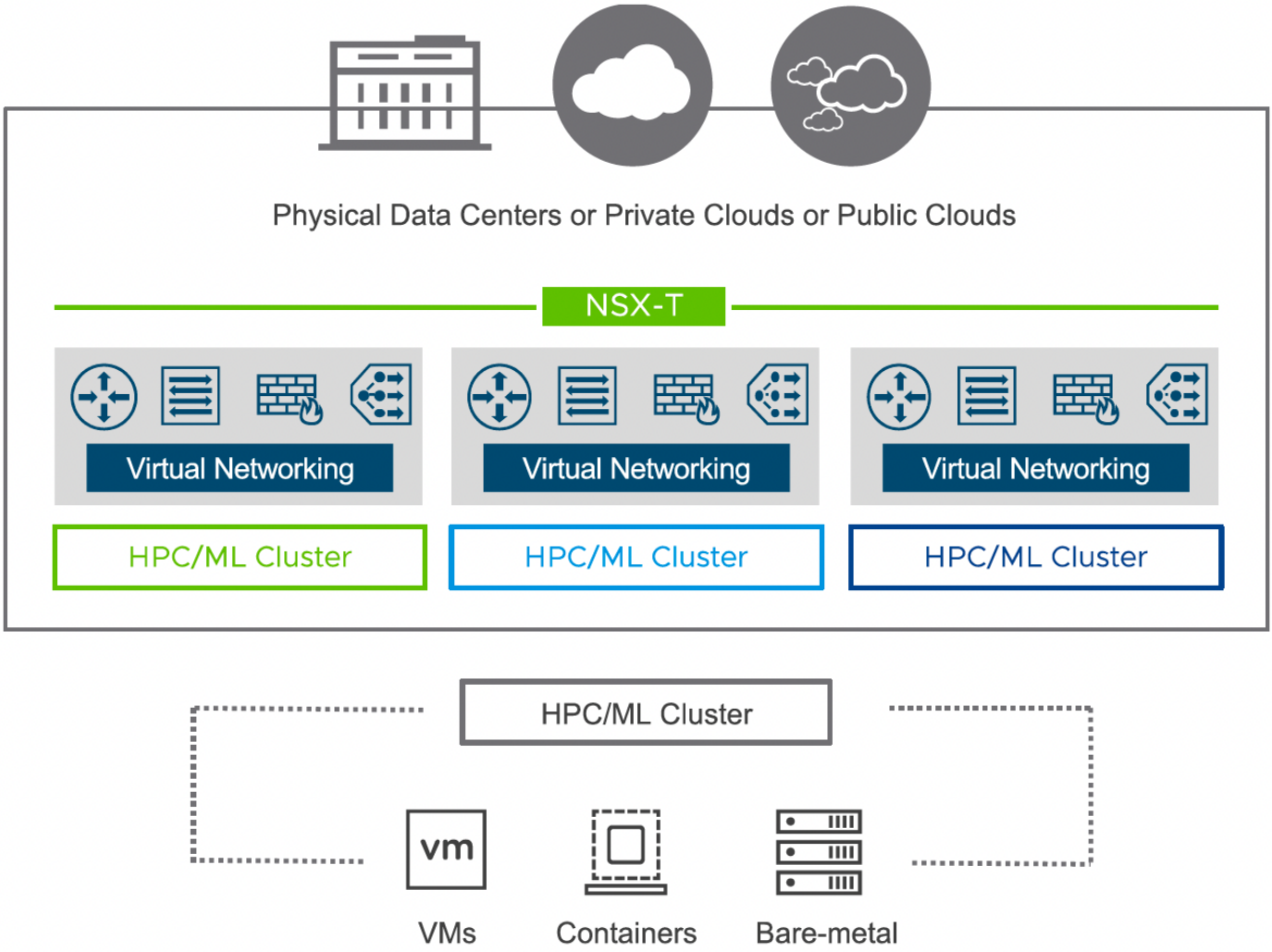
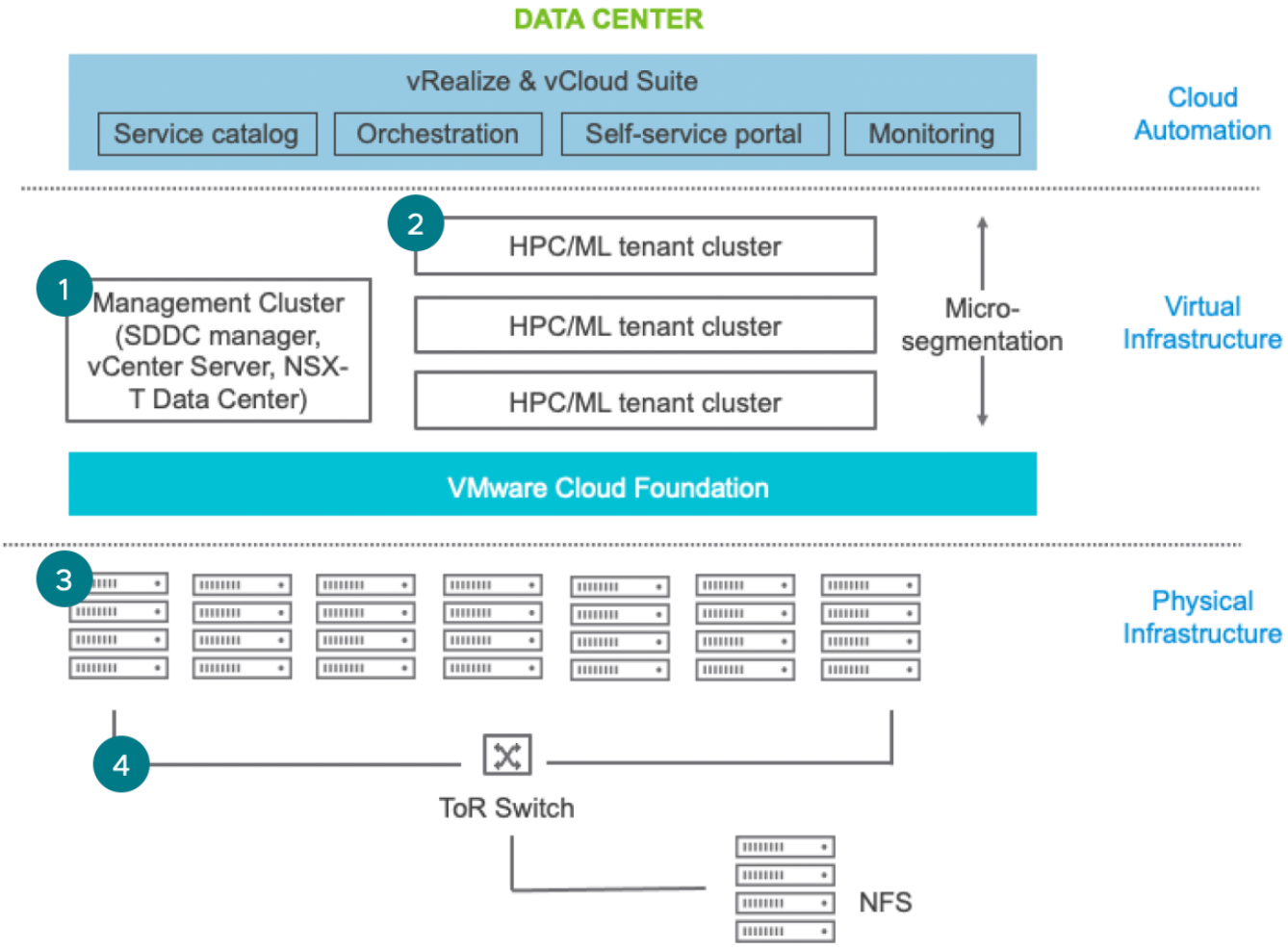


Figure 2 shows a reference architecture example of a multi-tenant private cloud for HPC/ML with four main components: management cluster, compute cluster, storage, and networking.

Figure 2. A private cloud reference architecture for multi-tenant HPC/ML



1 Management cluster

The management cluster runs VMs that manage the virtualized HPC (vHPC) environment. It includes the VCF management components and some service VMs. The VCF management components include:

- **SDDC Manager:** Provides a management interface to VCF and enables automatic updates.
- **vCenter Server:** Provides centralized management of the ESXi hosts and VMs; coordinates resources for the entire cluster.
- **NSX-T Data Center Manager:** Provides virtual switching, routing, load balancing, distributed firewalls, and overlays to HPC tenant clusters only for Ethernet networks. Admins must separately configure an RDMA interconnect.
- **vRealize and vCloud Suite (optional):** Include cloud automation functions, such as a service catalog and self-service portal, which allow individual departments or research labs to instantiate HPC/ML resources they

need without waiting for IT to do it. It also allows administrators to manage IT resources while adhering to business policies.

- Other services such as DNS (not shown in the figure).

2 Compute cluster

Different scientific and engineering groups dedicate the tenant virtual clusters, deployed on physical compute nodes, to running HPC/ML workloads. Each tenant cluster comprises a set of VMs that run on a shared underlying infrastructure.

Typically, an admin installs HPC job schedulers (for example, Slurm, Univa Grid Engine, or IBM Spectrum LSF) onto each tenant virtual cluster. The admin can deploy the management component of the job scheduler, such as the Slurm manager, within the management cluster to further boost resource utilization.

vSphere Bitfusion, meanwhile, allows applications running inside a VM to access one or more GPUs on remote nodes. It also supports multiple VMs sharing a single GPU.

3 Storage

Often, departments use NFS for home directories and project space mounted across all nodes, although they can leverage a parallel file system (not shown in the figure) for large-scale application data. Admins can configure high-speed interconnects, such as RDMA, to achieve low latency and high bandwidth for HPC application message exchanges or normal access to the parallel file system.

4 Networking

Management traffic typically communicates with 10/25G Ethernet among management nodes and compute nodes. A top-of-rack (ToR) switch connects all management and compute nodes.

Network design and considerations for HPC/ML workloads

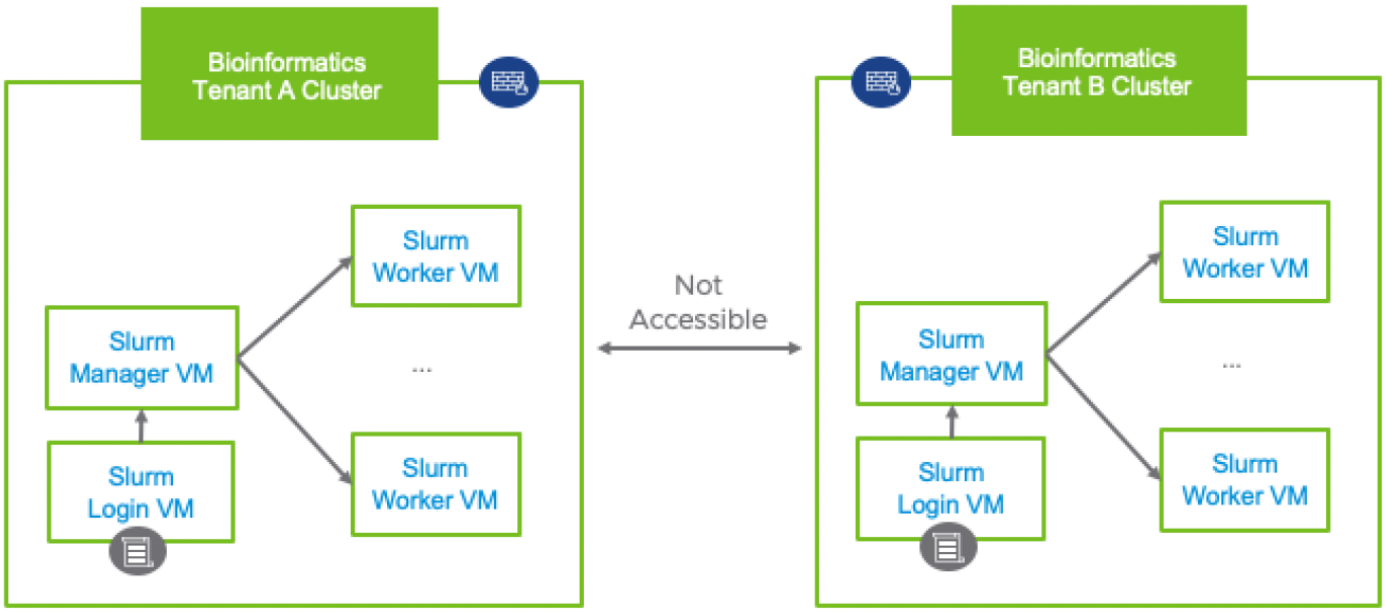
In an HPC/ML context, there are three important NSX-T features to consider:

- Distributed firewall
- VLAN vs GENEVE overlay
- Enhanced data path

Distributed firewall

Distributed firewall is a key feature for enforcing micro-segmentation. We distribute the enforcement at a fine-grain level, targeting specific VMs, resource pools, clusters, ESXi hosts, and data centers. Figure 3 illustrates how IT can apply distributed firewall rules to the HPC/ML tenant clusters to guarantee complete networking isolation.

Figure 3. Strict networking policies are set such that VMs within one bioinformatics tenant cluster can only communicate with each other. In addition, we can further enforce the rule that users within a tenant can only SSH to the Slurm Login VM of their bioinformatics cluster to protect the Slurm Manager or Slurm Worker VMs.



VLAN vs. GENEVE overlay

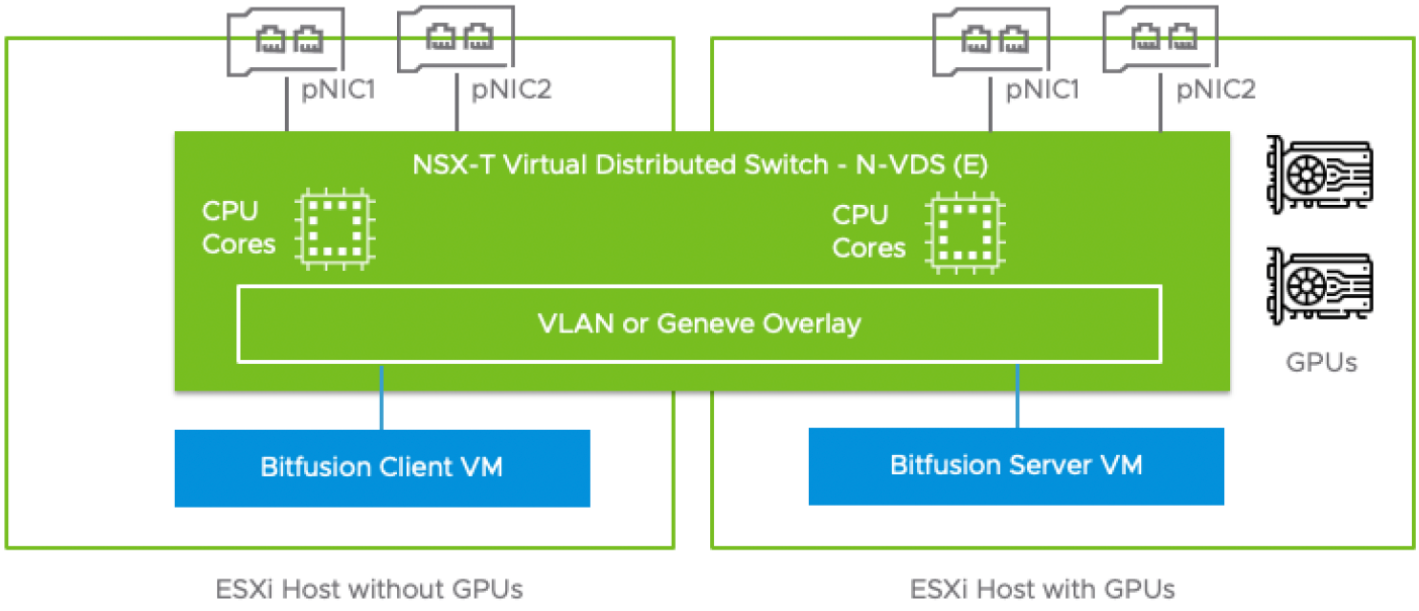
NSX-T supports virtual local area networks (VLANs) and GENEVE overlay. VLANs segment single physical networking into multiple, isolated virtual domains. Admins can tag each virtual logical network with a VLAN ID. GENEVE is a network encapsulation mechanism that allows IT to create logical networks that span physical network boundaries. NSX-T adopts the GENEVE tunneling mechanism to provide an overlay capability. It encapsulates logical networks in the user datagram protocol (UDP) and identifies every logical network by segment ID without a VLAN tag. As a result, many isolated Layer 2 networks can coexist underlying a common Layer 3 using the same VLAN ID.

Enhanced data path

In addition to the distributed firewall and VLAN vs. GENEVE, admins can leverage the enhanced networking stack (ENS) in NSX-T (also known as enhanced data path mode) to accelerate remote networking performance. NSX-T ENS primarily targets network functions virtualization (NFV) workloads that require a high performance data path, such as Telco and 5G, where improved packet throughput is critical. Here, deep learning training through vSphere Bitfusion GPU remoting also benefits from improved networking throughput. ENS supports both VLAN and GENEVE traffic.

As shown in Figure 4, traffic (either VLAN or GENEVE) between a Bitfusion client VM on a CPU node and a Bitfusion server VM on a GPU node can benefit from ENS when they connect to an ENS-enabled virtual switch (N-VDS (E)).

Figure 4. NSX-T ENS for accelerating networking performance



Performance implications

Table 1 illustrates the design considerations and performance implications of the three NSX-T features for running HPC/ML workloads.

Table 1. Design considerations and performance implications

Design decisions	Design justifications	Performance implications
Distributed firewall (DFW)	DFW provides protection of workload at the virtual NIC (vNIC) level.	• DFW adds some CPU overhead as it runs in the hypervisor kernel. • DFW adds network processing latency because it interposes a firewall.
GENEVE overlay	Overlay networks create isolated, multi-tenant broadcast domains across data center fabrics to deploy elastic, logical networks that span physical network boundaries.	• Overlay encapsulation adds additional bytes to each packet. • Overlay encapsulation and decapsulation use additional CPU resources. • Communicating with non-overlay endpoints (for example, an NFS server) must go through NSX Edge.
Enhanced data path	Enhanced data path (ENS) provides superior network performance. It is beneficial for running low latency and high throughput HPC/ML workloads.	Admins must assign dedicated CPU cores to manage the traffic to and from vNICs. This reduces the cores available for computational workloads on a fully loaded system.

HPC/ML performance

Given the potential benefits of these three NSX features, we quantitatively assessed the performance implications for their use in an HPC/ML environment with performance benchmarking.

This section describes the test methodology, workload selection, test setup, hardware/software details, and performance results.

Test cases

There are four test cases each for VLAN and GENEVE, as shown in table 2, when we consider testing with and without a distributed firewall (DFW) for NSX-T virtual distributed switch (N-VDS) and N-VDS with enhanced data path (ENS, or E) enabled.

Table 2. We performed eight test cases

	N-VDS		N-VDS (E)	
VLAN	w/o DFW (baseline)	w/ DFW	w/o DFW	w/ DFW
GENEVE	w/o DFW	w/ DFW	w/o DFW	w/ DFW

Testbed

The testbed in this study was a 21-node VCF cluster, of which 4 nodes acted as the management cluster. We used the remaining 17 nodes for computing: 16 as CPU-only servers and 1 GPU-enabled server. Each node had dual-port 25GbE network connectivity. For ENS, we dedicated 2 CPU cores on each host to support 1 core per NIC. The NFS server was a Dell PowerScale F200 array of 4 nodes, and it had 2x 25GbE connectivity to the main cluster. Tables 3 and 4 illustrate the testbed hardware and software details.

Table 3. Hardware details for testbed setup

Hardware component	Details
CPU server	Dell PowerEdge R740 Intel Xeon Gold 6248R Processor @ 3.0GHz 2 sockets, 24 cores per socket 384GB memory
GPU server	Dell PowerEdge C4140 Intel Xeon Gold 6248 Processor @ 2.5GHz 2 sockets, 20 cores per socket 192GB memory 4x NVIDIA V100 connected with NVLINK

Hardware component	Details
Switch	Dell PowerSwitch S5232F-ON
NIC	NVIDIA (Mellanox) ConnectX-4 Lx dual port 10/25GbE
NFS	Dell PowerScale F200, 4 nodes, 2x 25GbE

Table 4. Software details for testbed setup

Software component	Details
VMware Cloud Foundation (VCF)	4.2
NSX-T/NSX Manager	3.1
ESXi	7.0.2
vSphere Bitfusion	3.5.0
VM guest operating system	RHEL 8.1 for qperf and BioPerf Ubuntu 20.04.2 LTS for BERT
NVIDIA container for BERT	nvcr.io/nvidia/tensorflow:21.07-tf1-py3

Selection of workloads

We chose three benchmarks:

- **A microbenchmark:** qperf, a Linux utility, to measure TCP latency and bandwidth performance.
- **An HPC high-throughput application:** BioPerf, a bioinformatics benchmark suite that contains 10 highly popular bioinformatics programs commonly used to evaluate HPC cluster throughput performance.
- **A deep-learning application:** BERT. This deep-learning model has been widely applied to solve various natural language tasks. The benchmark source code is based on NVIDIA's NGC BERT Tensorflow container and the scripts within it. There are three major phases in language modeling: pre-training, fine-tuning, and inference. We adopted fine-tuning for benchmarking for a moderate computational load.

Performance test methodology and results

Latency and bandwidth measured using Linux qperf

We created two socket size VMs on two separate CPU-only nodes in the compute cluster, so that the number of vCPUs matched the number of physical cores on a single NUMA socket. We sized its memory to fit into one NUMA socket as well. For example, in our setup, each VM had 24 vCPUs, 160GB memory on a dual-socket physical node with 48 physical CPU cores, and 384GB memory. We fully reserved the VM's CPU and memory for maximum performance.

Figures 5-8 show, in microseconds, the round-trip latency comparisons among the different NSX-T network settings from table 2. We measured latency over a range of message sizes. The graphs show performance both with and without DFW and the ratios of the two as dashed lines. Here, a lower ratio is better, and a ratio of 1.0 indicates that no overhead incurred when we applied DFW.

Figure 5. Latency of NSX-T VLAN w/ & w/o DFW

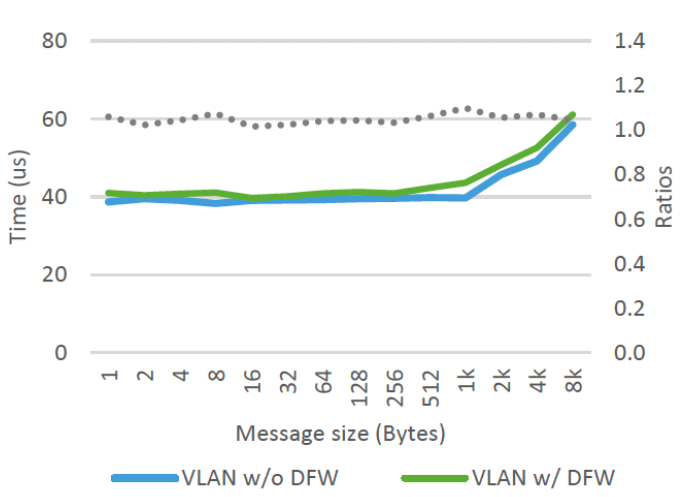


Figure 6. Latency of GENEVE overlay w/ & w/o DFW

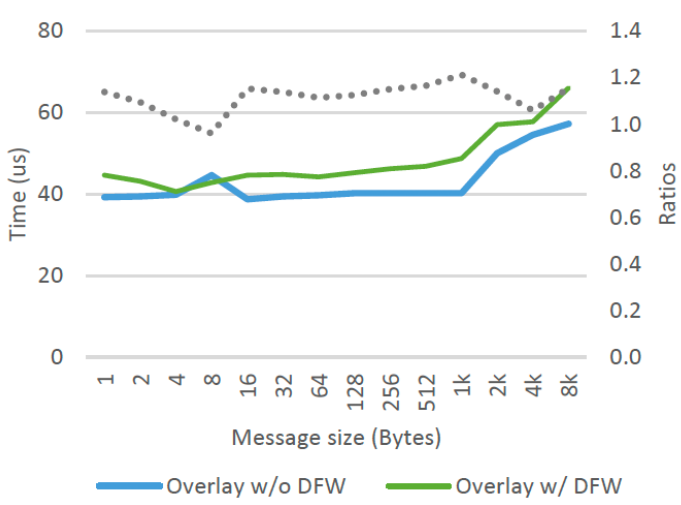


Figure 7. Latency of NSX-T ENS VLAN w/ & w/o DFW

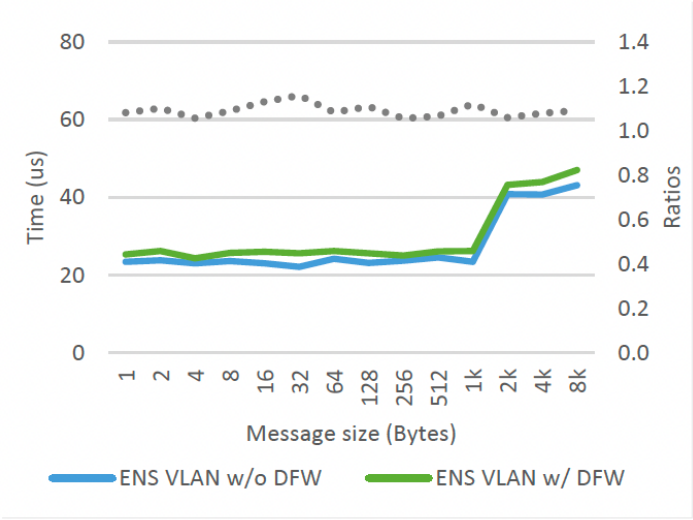
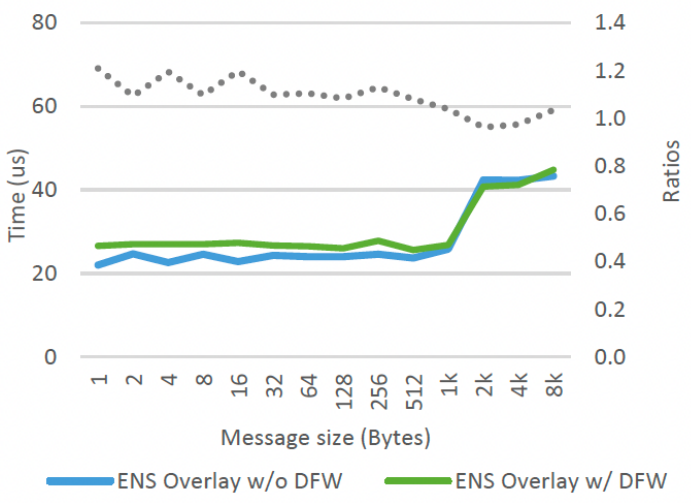


Figure 8. Latency of NSX-T ENS GENEVE overlay w/ & w/o DFW



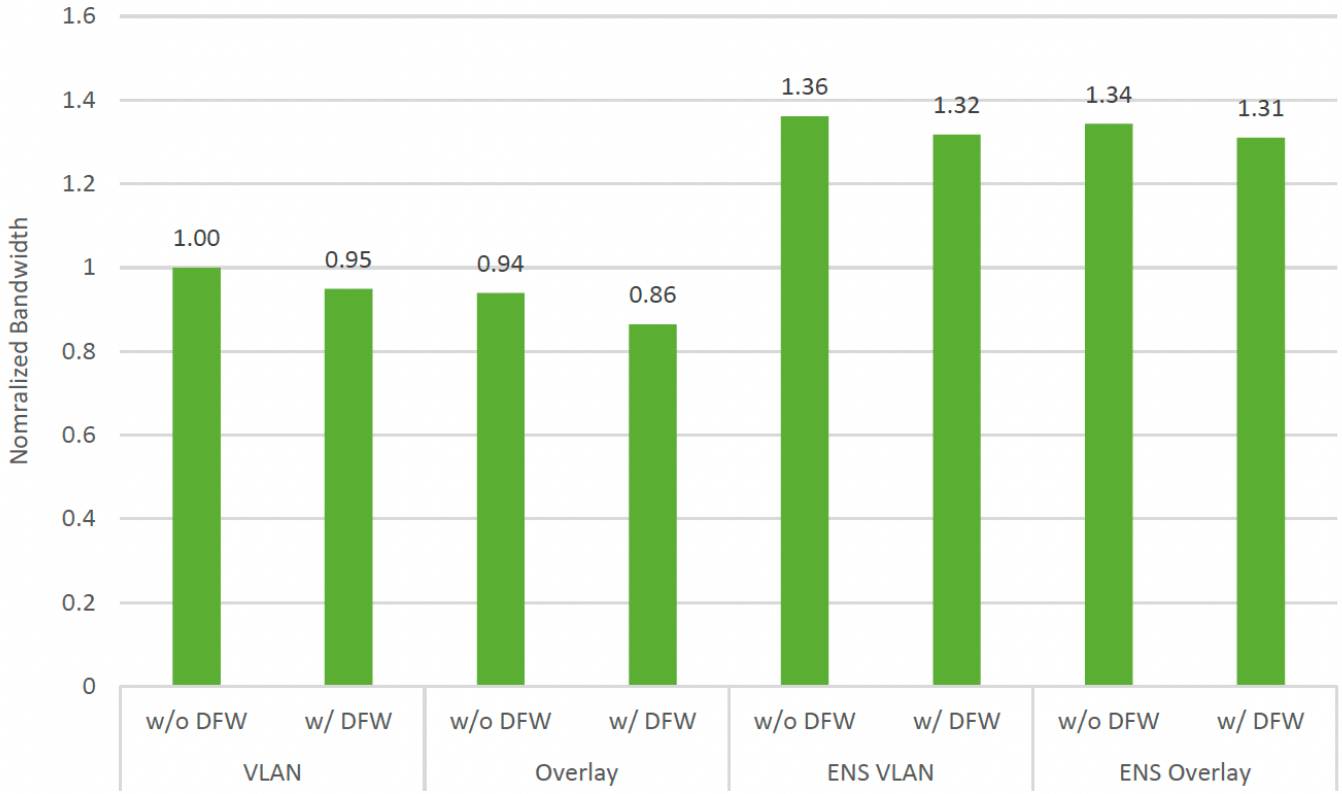
The latency figures show:

- DFW introduces small latency overheads in all cases, including VLAN, GENEVE overlay, ENS VLAN and ENS overlay.
- GENEVE overlay performance is close to that of VLAN. Similarly, ENS overlay performance is close to ENS VLAN.
- ENS demonstrates significantly better performance. For example, latency for 2-byte messages with VLAN without DFW is about 40µs, while ENS VLAN without DFW is nearly half—only about 23µs.

Figure 9 shows normalized bandwidth comparisons, where NSX-T VLAN without DFW establishes a baseline bandwidth. Here, a ratio higher than 1.0 means the bandwidth is better than the baseline case. This figure shows:

- With DFW, there is bandwidth degradation in all cases. For example, VLAN with DFW bandwidth is 5% lower than VLAN without DFW, and GENEVE overlay with DFW is 8% lower than GENEVE overlay without DFW.
- GENEVE overlay also leads to bandwidth degradations relative to VLAN, but with ENS, the difference is much smaller.
- ENS also demonstrates significantly better bandwidth performance than the default data path in both VLAN and overlay cases.

Figure 9. Normalized bandwidth (Baseline: NSX-T VLAN w/o DFW, higher is better)



Throughput measured using BioPerf with its high-throughput workload

As shown in figure 10, we created 2 tenant HPC clusters. Each tenant cluster had 32 worker VMs across all 16 CPU-only nodes in the cluster, where each VM was socket size. Each tenant cluster ran the same bioinformatics benchmark suite. This configuration followed the best practices for creating a multi-tenant HPC throughput computing environment to ensure hardware resources were fully utilized.

To achieve security isolation among tenants, we set up DFW rules both for each individual tenant and for the whole cluster. For each tenant, we only allowed external users to have SSH access to the login node. Internally, all of a tenant's nodes could communicate with each other, and they could access public services such as NFS and DNS. At the cluster level, we enabled ICMP and DHCP public services and rejected all other traffic.

Figure 10. Test setup for multi-tenant HPC testing. We set DFW rules to enforce security isolation between tenants. We used NSX T0 and T1 gateways only for overlay segments.

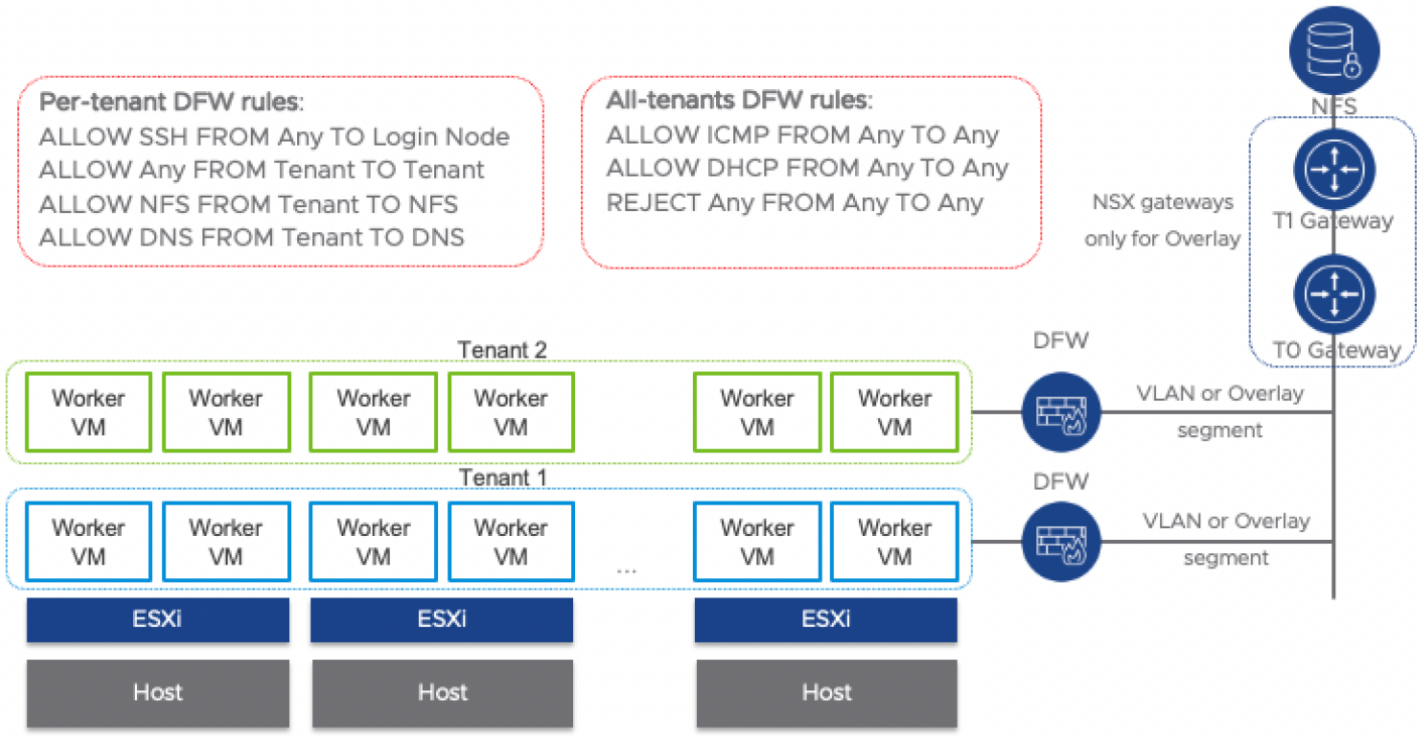


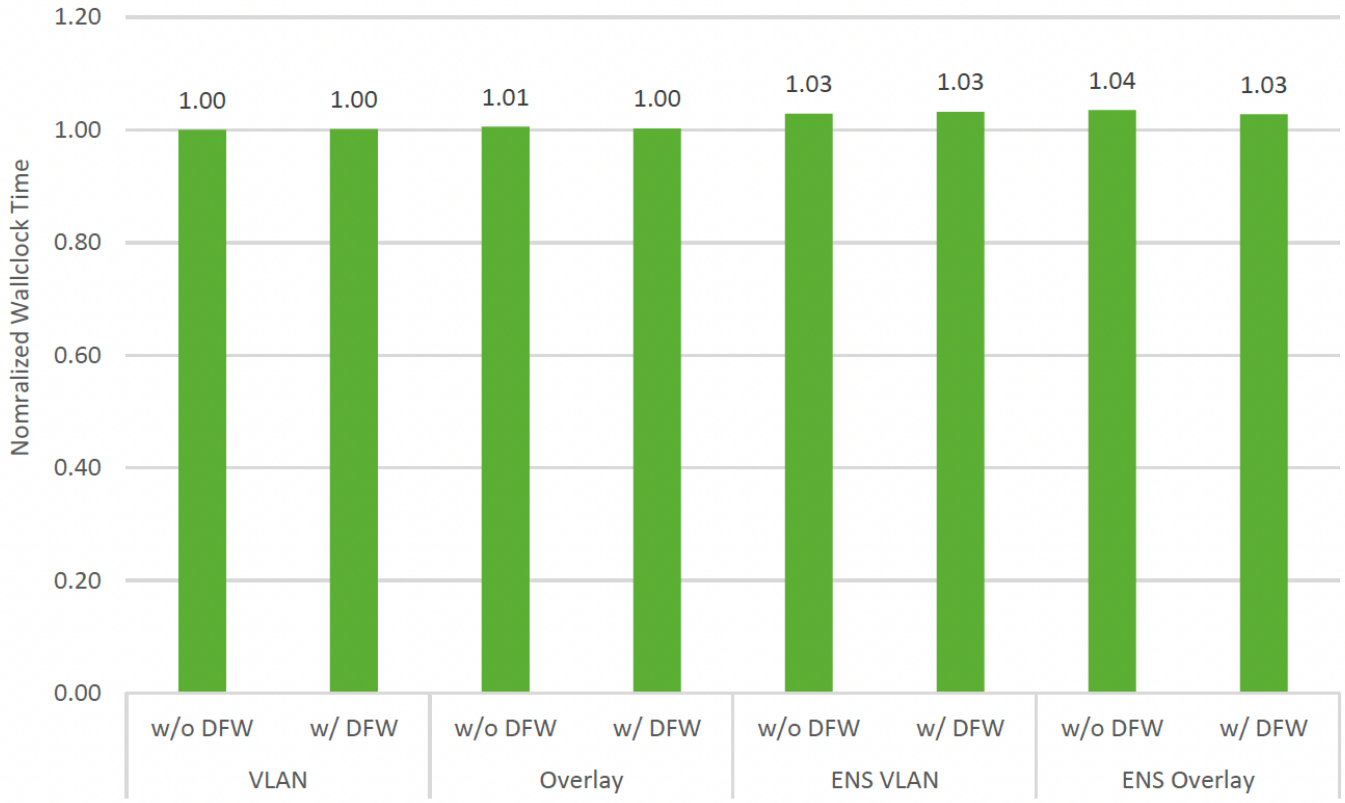
Figure 11 shows the normalized wall-clock time average of two HPC tenant clusters running BioPerf jobs. NSX-T VLAN without DFW establishes the baseline. A ratio higher than 1.0 means a case required more time to finish all jobs, thus indicating overhead relative to the baseline case.

The ENS cases (ENS VLAN and ENS Overlay) show slight degradation relative to non-ENS cases (VLAN and overlay). As shown from microbenchmarks, ENS, with its optimized networking stack, provides superior latency and bandwidth.

However, we needed to assign dedicated CPU cores to manage incoming and outgoing traffic. In this testing, the 2 HPC tenant clusters simultaneously ran computationally intensive jobs that fully loaded the system. While the increased networking performance was beneficial for accelerating application NFS I/O, the required core reservation for ENS offset the networking performance gains and lead to a performance degradation for high-throughput workloads.

DFW or GENEVE overlay barely added any overhead to the high-throughput workloads. This indicates that users who run high-throughput workloads can take advantage of NSX-T Data Center's micro-segmentation capabilities and enjoy full network security without experiencing a performance impact.

Figure 11. Normalized wall-clock time for two HPC tenant clusters to finish all jobs (Baseline: NSX-T VLAN without DFW, lower is better)



ML/DL training speed measured using BERT (testbed set up with Bitfusion VMs)

Machine learning and deep learning are considered part of the HPC workload because they require a large number of computations and data movement. For this study, we created 2 socket size VMs with 1 on a GPU host and 1 on a CPU host. The VM on the CPU host acted as the Bitfusion client, and the VM on the GPU host acted as the Bitfusion server. We installed the GPU host with 4 NVIDIA V100 GPUs, which we configured as DirectPath I/O (passthrough) info the Bitfusion server VM.

Figure 12 compares the BERT fine-tuning training speed in sentences per second, so a higher score is better. We normalized all the results to that of VLAN without DFW to facilitate the comparison. The results show that when using the default networking stack, there is no visible impact from DFW or overlay. Even with ENS, the differences are within 1% of each other. This result aligns well with our HPC throughput application results above.

Applying ENS improved the BERT fine-tuning training performance by 1% or 2%. This is different from our above HPC throughput application results because we didn't fully load the system under test in this case. We expect to see higher improvement from ENS for workloads where the communication-to-computation ratio is higher.

Figure 12. Normalized BERT fine-tuning training speed in sentences per second (Baseline: NSX-T VLAN without DFW, higher is better)

Performance best practices

The performance test results reveal that:

- DFW adds little latency and has moderate impact on network bandwidth. We recommend applying DFW rules to achieve micro-segmentation in a multi-tenant environment for both HPC throughput workloads and ML/DL workloads using Bitfusion.
- GENEVE overlay has nearly the same network latency as basic VLAN, but it decreases network bandwidth. That's because encapsulation adds additional bytes to each packet, and the bandwidth can be further reduced when coupling with DFW. Thus, we only recommend GENEVE overlay when an overlay network is desired, such as when logical networks that span physical network boundaries are required.
- For HPC high-throughput workloads, which can benefit from using more cores, we recommend using the default data path in NSX-T Data Center instead of ENS and leaving all the cores for the workloads.
- Following our approach, admins can support multiple tenants simultaneously by creating multiple virtual clusters on the shared physical infrastructure. Admins can use DFW from NSX-T Data Center to achieve security isolation among the tenants and implement CPU over-provisioning to improve overall resource utilization.
- Admins don't need to utilize all the CPU cores for ML/DL workloads that use Bitfusion for GPU remoting. We recommend configuring ENS on both the client and server sides. The superior networking performance from ENS will reduce the communication overhead during remote GPU calls.

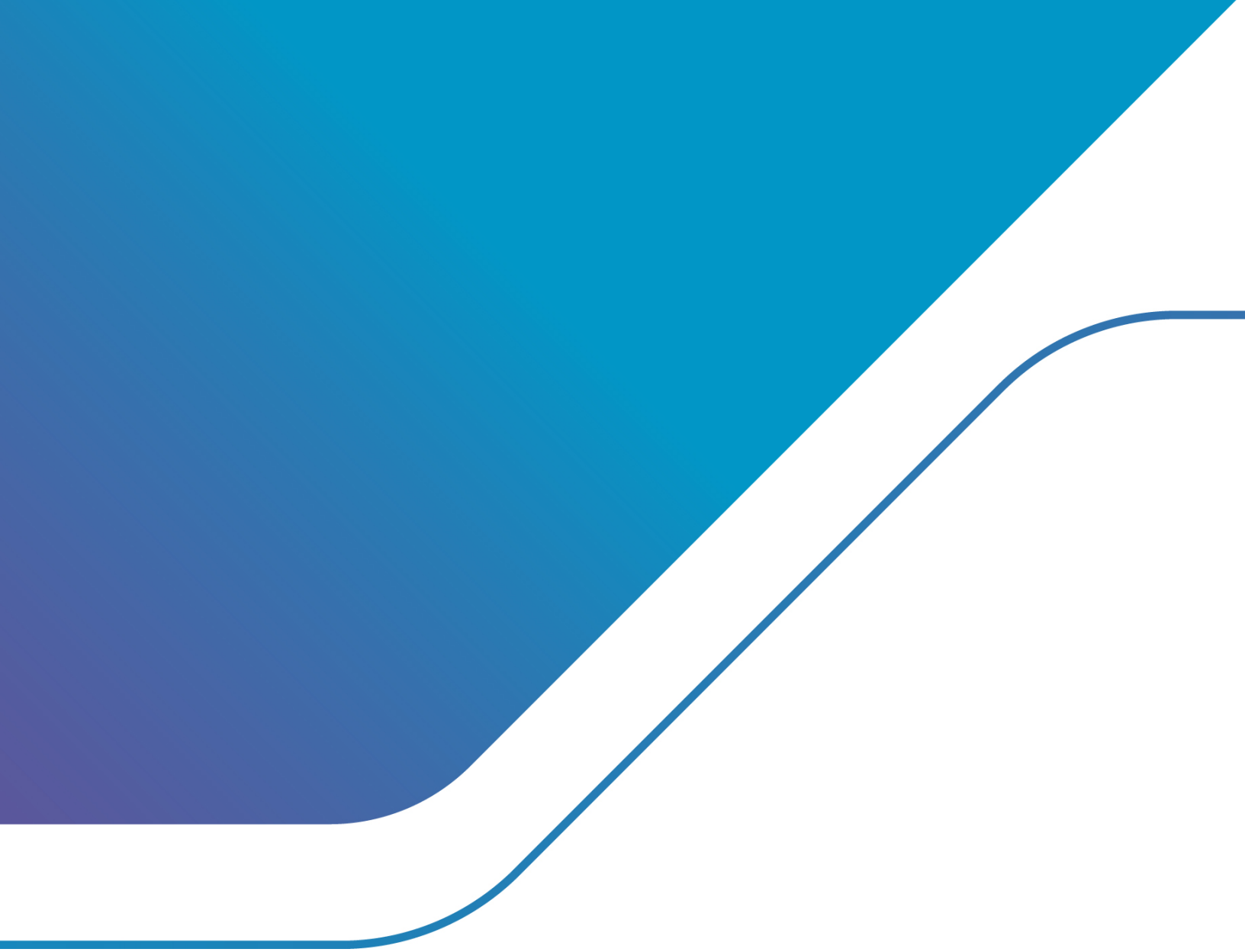
Conclusion

HPC enables computational scalability for breakthrough innovations, and virtualizing HPC infrastructure adds value. Among the various benefits of virtualization, networking security is a critical component to enable secure multi-tenancy. VMware NSX-T Data Center can provide a multi-tenant secure networking solution for HPC and ML workloads with minimal performance overhead.

About the authors

Michael Cui (VMware alum) was a senior member of technical staff in the VMware Office of the CTO, focusing on virtualizing high performance computing. His expertise spans broadly across distributed systems and parallel computing. His daily work ranges from integrating various software and hardware solutions, to conducting proof-of-concept studies, to performance testing and tuning, and to technical paper publishing. In addition, Michael serves on Hyperion's HPC Advisory Panel and participates in paper reviewing in several international conferences and journals, such as IPCCC, TC, and TSC. Previously, he was a research assistant and part-time instructor at the University of Pittsburgh. He holds both doctoral and master's degrees in computer science from the University of Pittsburgh.

Na Zhang (VMware alum) was a staff engineer working on high performance computing and machine learning within VMware's Office of the CTO. She has broad experience in the areas of design and implementation of vHPC tools, ML-as-a-service platform architecture, performance tuning and best practices for HPC and ML workloads, accelerator solutions, and integration of HPC middleware with VMware products. Based in Boston, Na is the author of multiple research and technical papers, many designed to support VMware customers to optimize their HPC and ML deployments. She has also served on the technical program committee for several international HPC events, including SC, vHPC, and HPC&S. She received her PhD in Applied Mathematics from Stony Brook University in 2015. Her research primarily focused on the design and analysis of parallel algorithms for large and multi-scale simulations running on world-class supercomputers.



vmware® by Broadcom