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

Reference architecture and performance study
Table of contents

Introduction ..................................................................................................................................... 3

HPC workload classification .......................................................................................................... 3

Multi-tenant reference architecture ............................................................................................. 4
  Management cluster 5
  Compute cluster 5
  Storage 5
  Networking 5

Network design and considerations for HPC/ML workloads ..................................................... 6
  Distributed firewall 6
  VLAN vs. GENEVE overlay 7
  Enhanced data path 7
  Performance implications 8

HPC and ML benchmarking .......................................................................................................... 8
  Test methodology 9
  Testbed 9
  Selection of workloads 10

Performance best practices ........................................................................................................... 15

Summary ........................................................................................................................................ 16

Authors .......................................................................................................................................... 16
**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. Similar to other virtualization benefits, it also offers significant value for virtualized and cloud HPC environments. When multi-tenancy is exploited to increase hardware resource utilization, full isolation of research projects is needed to ensure that project files and data are not accessed inappropriately by other users. Though public clouds make an array of security policies available, there are still challenges related to security and management flexibility, particularly for situations where the highest security is required — for example, clinical genomic sequencing, chip design, or other sensitive areas of research that undergo regulatory compliance. 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 Enhanced Data Path/Network Stack (ENS). Finally, we offer a list of best practices.

**HPC workload classification**

In broad terms, HPC workloads fall into three categories:

- **High-throughput workloads.** These are usually “embarrassingly parallel” because they 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 or thousands — or even millions — of executions with varying inputs.

- **Parallel-distributed workloads.** In contrast to high-throughput workloads, parallel-distributed workloads often perform sustained and intense communication within a single job, making their performance sensitive to interconnect bandwidth and latency. For such applications, Remote Direct Memory Access (RDMA) can be used 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, and Paravirtualized RDMA.

- **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. GPU compute accelerators, in particular, can be configured in three ways in VMware virtualized environments: 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 this [blog](https://example.com) for more details).

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

1. High-throughput workloads that have I/O to a network file system (NFS).
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-based 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 switching, routing, firewalls and load balancing are distributed across the entire environment.

![Figure 1: Unified Networking Infrastructure for Multi-Cloud HPC/ML.](image)
Figure 2 shows an example reference architecture of a multi-tenant private cloud for HPC/ML with four main components: management cluster, computer cluster, storage, and networking.

Management cluster

The management cluster runs VMs that manage the virtualized HPC (vHPC) environment. Primarily, 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 users to perform automatic updates.
- vCenter Server: provides centralized management of the hosts and VMs, coordinating resources for the entire cluster.
- NSX-T Data Center manager: provides virtual switching, routing, load balancing, distributed firewalls, and overlays to HPC tenant clusters. It applies only to Ethernet networking, and RDMA interconnect should be configured separately.
- vRealize and vCloud Suite (optional): adds 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 create resources for them. It also enables administrators’ management of IT resources while ensuring compliance with business policies.
- Other services such as DNS (not shown in the figure).

Compute cluster

The tenant virtual clusters, which are deployed on physical compute nodes, are dedicated to running HPC/ML workloads for different scientific and engineering groups. Each tenant cluster comprises a set of VMs that run on a shared underlying infrastructure.

Typically, HPC job schedulers (for example, Slurm, Univa Grid Engine, or IBM Spectrum LSF) are installed onto each tenant virtual cluster. The management component of the job scheduler, such as the Slurm manager, can be deployed 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.

Storage

Often, NFS is used for home directories and project space mounted across all nodes, although a parallel file system (not shown in the figure) can instead be leveraged for large-scale application data. Exploiting high-speed interconnects such as RDMA can achieve low latency and high bandwidth for HPC application message exchanges or normal accessing of the parallel file system.

Networking

Management traffic typically communicates with 10/25 Gb/s 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

Distributed firewall is a key feature for enforcing micro-segmentation. The enforcement is distributed at a fine-grain level, such as specific VMs, resource pools, clusters, ESXi hosts, and data centers. As shown in Figure 3, distributed firewall rules can be applied to the HPC/ML tenant clusters to ensure full networking isolation.
Secure Networking for Multi-Tenant High-Performance Computing and Machine Learning Environments

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. Each virtual logical network can be tagged with a VLAN ID. GENEVE is a network encapsulation mechanism that allows users 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 User Datagram Protocol 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 distributed firewall and VLAN vs. GENEVE, the Enhanced Networking Stack in NSX-T (also known as Enhanced Data Path mode) can be leveraged 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 traffics.

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>
<thead>
<tr>
<th>DESIGN DECISIONS</th>
<th>DESIGN JUSTIFICATIONS</th>
<th>PERFORMANCE IMPLICATIONS</th>
</tr>
</thead>
</table>
| Distributed firewall (DFW) | Distributed Firewall 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 as it basically interposes a firewall. |
| GENEVE | 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 (e.g., an NFS server) needs to go through NSX Edge |
| Enhanced Data Path | Enhanced Data Path provides superior network performance. It is beneficial for running low latency and high throughput HPC/ML workloads. | Dedicated CPU cores need to be assigned to manage the traffic to and from vNICs, reducing the cores available for computational workloads on a fully loaded system. |

Table 1: Design considerations and performance implications.

HPC and ML benchmarking
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 test methodology, workload selection, experimental setup, hardware/software details, and performance results.

Test methodology

As shown in Table 2, if we consider VLAN or GENEVE Overlay with or without DFW, and then with or without ENS, we have eight possible test cases. VLAN traffic without DFW establishes a performance baseline. N-VDS denotes NSX-T Virtual Distributed Switch, and N-VDS (E) denotes NSX-T Virtual Distributed Switch with ENS enabled.

<table>
<thead>
<tr>
<th>Test case</th>
<th>N-VDS</th>
<th>N-VDS (E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VLAN w/o DFW (baseline)</td>
<td>w/o DFW</td>
<td>w/o DFW</td>
</tr>
<tr>
<td>VLAN w/ DFW</td>
<td>w/ DFW</td>
<td>w/ DFW</td>
</tr>
<tr>
<td>GENEVE w/o DFW</td>
<td>w/o DFW</td>
<td>w/o DFW</td>
</tr>
<tr>
<td>GENEVE w/ DFW</td>
<td>w/ DFW</td>
<td>w/ DFW</td>
</tr>
</tbody>
</table>

Table 2: Test cases.

Testbed

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

**Table 3: Hardware details for testing.**

<table>
<thead>
<tr>
<th>HARDWARE</th>
<th>SPECIFICATIONS</th>
</tr>
</thead>
</table>
| CPU Server | DellEMC R740  
Intel Xeon Gold 6248R @ 3.0GHz  
2 sockets, 24 cores per socket  
384 GB memory |
| GPU Server | DellEMC PowerEdge C4140  
Intel Xeon Gold 6248 @ 2.5GHz  
2 sockets, 20 cores per socket  
192 GB memory  
4 x NVIDIA V100 connected with NVLINK |
| Switch | Dell EMC S5232F-ON |
| NICs | Mellanox ConnectX-4 LX Dual Port 10/25GbE |
| NFS | Dell PowerScale F200, 4 nodes, 2 * 25 GbE |
Secure Networking for Multi-Tenant High-Performance Computing and Machine Learning Environments

Table 4: Software details for testing.

<table>
<thead>
<tr>
<th>Software</th>
<th>Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>VMware Cloud Foundation</td>
<td>4.2</td>
</tr>
<tr>
<td>NSX-T Manager</td>
<td>NSX-T 3.1</td>
</tr>
<tr>
<td>ESXi hypervisor</td>
<td>7.0.2</td>
</tr>
<tr>
<td>vSphere Bitfusion</td>
<td>3.5.0 release</td>
</tr>
<tr>
<td>VM Guest OS</td>
<td>RHEL 8.1 for qperf and BioPerf</td>
</tr>
<tr>
<td></td>
<td>Ubuntu 20.04.2 LTS for BERT</td>
</tr>
<tr>
<td>NVIDIA Container for BERT</td>
<td>nvcr.io/nvidia/tensorflow:21.07-tf1-py3</td>
</tr>
</tbody>
</table>

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 that are commonly used to evaluate HPC cluster throughput performance.
- A deep-learning application: BERT (Bidirectional Encoder Representation from Transformers). 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 as the computational load is moderate for our cluster setting.

The experiment setup and performance results for the above three benchmarks are described below.

Microbenchmark — latency and bandwidth

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

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

From the latency figures we can see:

- 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 40 µs, while ENS VLAN without DFW is nearly half — only about 23 µs.

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. From the figure, we can see:
• With DFW, there is bandwidth degradation in all cases. For example, VLAN with DFW bandwidth is 5 percent lower than VLAN without DFW, and GENEVE Overlay with DFW is 8 percent 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 5:** Qperf latency of NSX-T VLAN w/o DFW.

**Figure 6:** Qperf latency of NSX-T Overlay w/o DFW.

**Figure 7:** Qperf latency of NSX-T ENS VLAN w/o DFW.

**Figure 8:** Qperf latency of NSX-T ENS Overlay w/o DFW.
Secure Networking for Multi-Tenant High-Performance Computing and Machine Learning Environments

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

High-throughput application — BioPerf

As shown in Figure 10, we created two tenant HPC clusters. Each tenant cluster had 32 worker VMs across all 16 CPU-only nodes in the cluster, where each VM was socket-sized. Each tenant cluster ran the same bioinformatics benchmark suite. This configuration follows best practices for creating a multi-tenant HPC throughput computing environment to ensure maximum utilization of hardware resources.

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: Environmental setup for multi-tenant HPC testing. DFW rules are set up to enforce security isolation between tenants. Note that NSX T0 and T1 gateways are only used 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.

From the figure, we can see 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, dedicated CPU cores must be assigned to manage incoming and outgoing traffic. In this testing, the two 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 performance degradation for high-throughput workloads.

At the same time, DFW or GENEVE Overlay barely add any overhead to high-throughput workloads. This indicates that high-throughput workload users can take advantage of NSX-T Data Center’s micro-segmentation capabilities and enjoy full network security without experiencing 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)

Deep learning application with Bitfusion – BERT

Machine learning and deep learning are considered part of the HPC workload, as they require a large number of computations and data movement. For this study, we created two socket-size VMs with one on a GPU host and one 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 four NVIDIA V100 GPUs, which were configured in DirectPath I/O (passthrough) mode into the Bitfusion server VM.

Figure 12 compares the BERT finetuning training speed in sentences per second, so a higher score is better. Note that all results are normalized 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 one percent of each other. This result aligns well with our HPC throughput application results above. At the same time, applying ENS improved the BERT finetuning training performance by one to two percent. This is different from our above HPC throughput application results because the system under test was not fully loaded in this case. We expect to see higher improvement from ENS for workloads where the communication-to-computation ratio is higher.
Performance best practices

Based on the performance results:

- 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 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 to users’ workload.

- Following our approach, multiple tenants can be supported simultaneously by creating multiple virtual clusters on the shared physical infrastructure. DFW from NSX-T Data Center can be used to achieve security isolation among the tenants, while CPU over-provisioning can improve overall resource utilization.

- For ML/DL workloads which use Bitfusion for GPU remoting, giving all the CPU cores to user workloads won’t be necessary. 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.
Summary

HPC enables computational scalability for breakthrough innovations and virtualizing HPC infrastructure adds much 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.

Authors

Michael Cui is 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 is a staff engineer working on High-Performance Computing (HPC) and Machine Learning (ML) within VMware’s Office of the CTO. She has broad experience in the areas of design and implementation of vHPC tools, machine-learning-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 a number of international HPC events, including SC, vHPC and HPC&S. She received her Ph.D. 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.