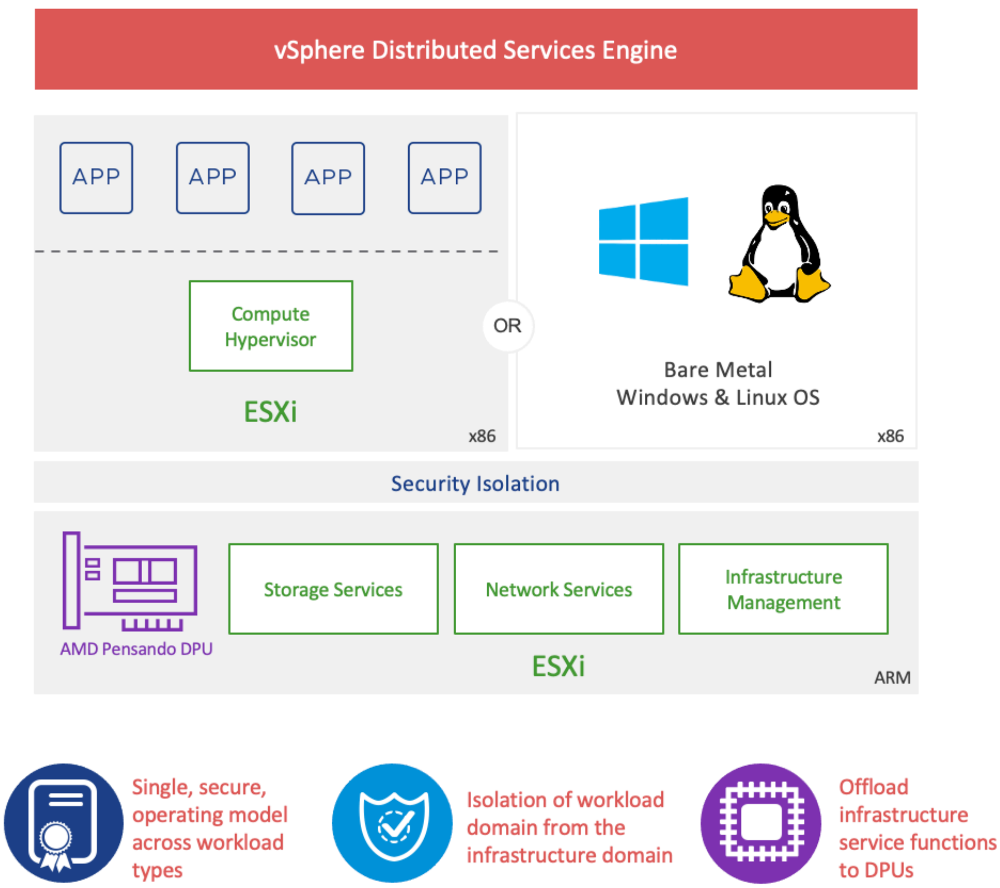
VMware® vSphere® Distributed Services Engine with AMD Pensando DPU

July 2023

Overview

The VMware® vSphere® Distributed Services Engine (vDSE) is the next major step in the evolution of VMware vSphere and VMware Cloud Foundation (VCF), focused on delivering infrastructure for modern VM and container-based applications, as well as supporting bare-metal workloads. vDSE delivers virtual infrastructure as a distributed control fabric through tight integration with data processing units (DPUs) running inside OEM supported x86 platforms. VMware is leading an industry-wide initiative to deliver this solution to its customers by bringing together best-of-breed DPU silicon (including AMD) and best-of-breed server OEMs (Dell Technologies, HPE, Lenovo). vDSE gives VMware and AMD customers the ability to accelerate and secure workloads on next-generation composable hardware systems:

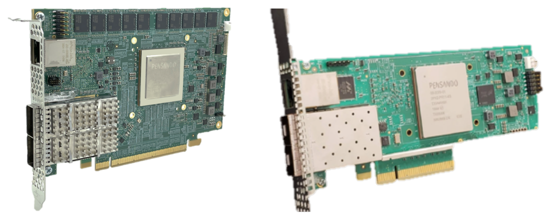


vSphere Architecture (source: VMware, Inc. Used by permission).

VMware vSphere Distributed Services Engine with AMD Pensando DPU

Why AMD?

The core intellectual property coming from AMD is the *data processing unit* or *DPU*. This technology represents a leap forward in bringing simplicity and efficiency to not only today's hyperscale and public cloud providers (where similar 'function accelerator' technology is already widely deployed), but also to large enterprise IT customers, as they continue to build and manage compute infrastructure across hybrid cloud environments.



10/25/100G DPU-based adapters

These environments typically require managing very large on-premises compute environments, an area where VMware and the vSphere platform has for years played a key role in making x86 servers more efficient and optimized. Over the next few months as server vendors begin to adopt the latest technologies from AMD and VMware, the Distributed Services Engine (running vSphere on DPUs) will leverage the AMD EPYC™ CPU and the AMD Pensando™ DPU to maximize and efficiently use enterprise compute resources and data center infrastructure by driving higher workload consolidation (ex. reduced consumption of CPU cores for x86 networking), offloading infrastructure services, and supercharging applications by providing best in class performance.

World's Most Intelligent DPU

Market Leadership

- At least one generation ahead of competitors[1]
- First to market with 100G/200G@<60W

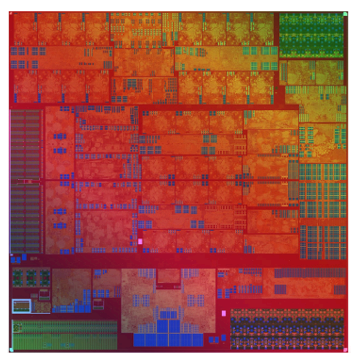
Performance Leadership

- 144x P4 packet processors, native datapath for I/0, device emulation
- Supports tens of millions of network flows
- Concurrent services at line rate performance
- Network | Security | Storage | Telemetry

Architectural Leadership

- Fully programmable control, data and management planes
- Future-ready: 200GE, best investment for new server lifecycle aligned with add-on VMware features (ex. storage, security, bare-metal)

"Elba" | Software-in-Silicon



2nd Generation
2 x 200G
7 nm

Introducing VMware vSphere 8

Many customers in the enterprise already leverage VMware vSphere in their data centers today. As a part of their journey to vSphere 8, VMware has done significant work to bring the benefits of cloud and cloud infrastructure directly to on-premises workloads. This includes the same software-defined, programmable networks that exist in most public cloud providers, but now vSphere supercharges the application and workload performance through enhancements made with DPUs.

vSphere 8 leans into a new era of heterogeneous computing by introducing DPUs to enterprises through VMware vSphere Distributed Services Engine. vSphere Distributed Services Engine is the next step in the evolution of cloud infrastructure for modern applications leveraging a cloud proven technology and architecture that enables running core infrastructure services in a distributed fashion between the CPU and the DPU to enable maximum performance, better scale and efficiency of resources, and better security by separating tenant/provider domains and moving network security controls onto a dedicated processor, outside of the workload/application attack surface.

vSphere Distributed Services Engine modernizes cloud infrastructure into a distributed architecture enabled by DPUs to:

- Meet the throughput and latency needs of modern distributed workloads by accelerating networking functions
- Deliver best infrastructure price-performance by providing more CPU resources to workloads
- Reduce operational overhead of DPU lifecycle management with integrated vSphere workflows

Available today with the latest versions of vSphere, Distributed Services Engine offloads and accelerates vSphere Distributed Switch and networking and security with NSX on the DPU.

VMware vSphere®

- **Performance**
  - **EDP (Enhanced Datapath) - Standard**
- **Management**
  - **Unified Installer across the CPU and DPU**
  - **Seamless Lifecycle Management of the DPU via vLCM**
  - **DPU Performance & Health Monitoring via vCenter**

VMware NSX®

- **Performance**
  - **EDP (Enhanced Datapath) - Performance**
  - **VMDirectPath (Uniform Passthrough) with vMotion and DRS support**
  - **NSX Overlay and VLAN offload**
- **Observability**
  - **IPFIX, Port mirroring, packet capture**
- **Security**
  - **Distributed Firewall**

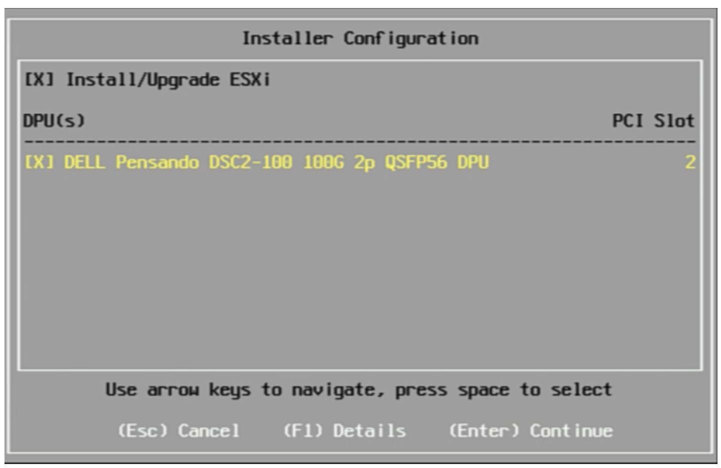
Key DPU capabilities within vSphere 8 and NSX 4
(see VMware vSphere on DPUs)

Simplified Management with vSphere 8

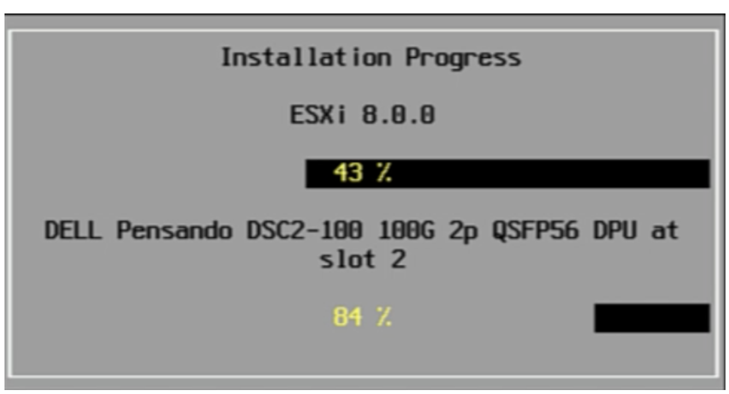
Because vSphere Distributed Services Engine preserves current VMware vSphere operators' existing Day 0, Day 1, and Day 2 vSphere experience, customers who are already familiar operationally with vSphere and have built automation and orchestration frameworks around vCenter APIs will not have to make changes to their tooling or applications and workloads to take advantage of the benefits that DPUs bring to the vSphere platform. From a configuration standpoint, enabling network offloads is a simple checkbox in the VDS creation wizard, which is the first step to supporting infrastructure services offloads (such as networking, security, telemetry, and storage) within vSphere/vCenter on the AMD Pensando DPU.

Unified Installer across CPU and DPU

The vSphere on DPU installation process has been enhanced and updated to support installing ESXi on x86 and ESXi on the DPU in sync, by selecting one within the ESXi 8.0 installer:

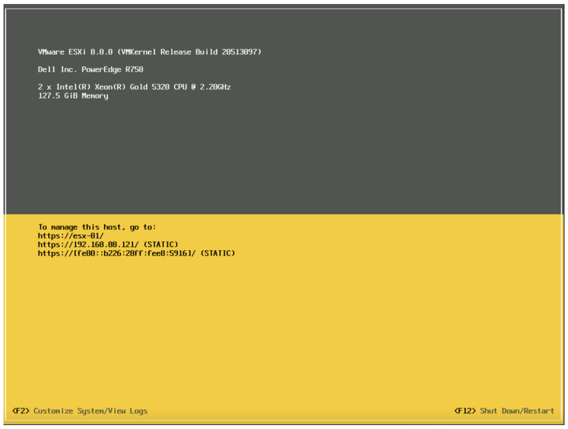


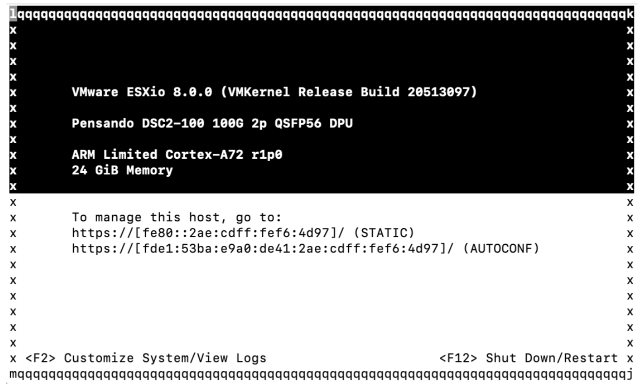
This process leverages the server's management controller (such as iLO or iDRAC) and Redfish API calls to install an ESXi image from the installer onto the AMD Pensando DPU. This process helps ensure that the version of ESXi running on the server and the DPU is always in sync. As the installer images ESXi on x86 to the local SSD that was selected when running the installer (top progress bar), it will in parallel load ESXi onto the DPU (bottom progress bar):



VMware vSphere Distributed Services Engine with AMD Pensando DPU

Upon completion of the installer process and a server reboot, the system will boot with an instance of ESXi on the x86 CPU(s) and an instance of ESXi running on the DPU. Once this process is complete, users can start taking advantage of networking offloads in vSphere 8:

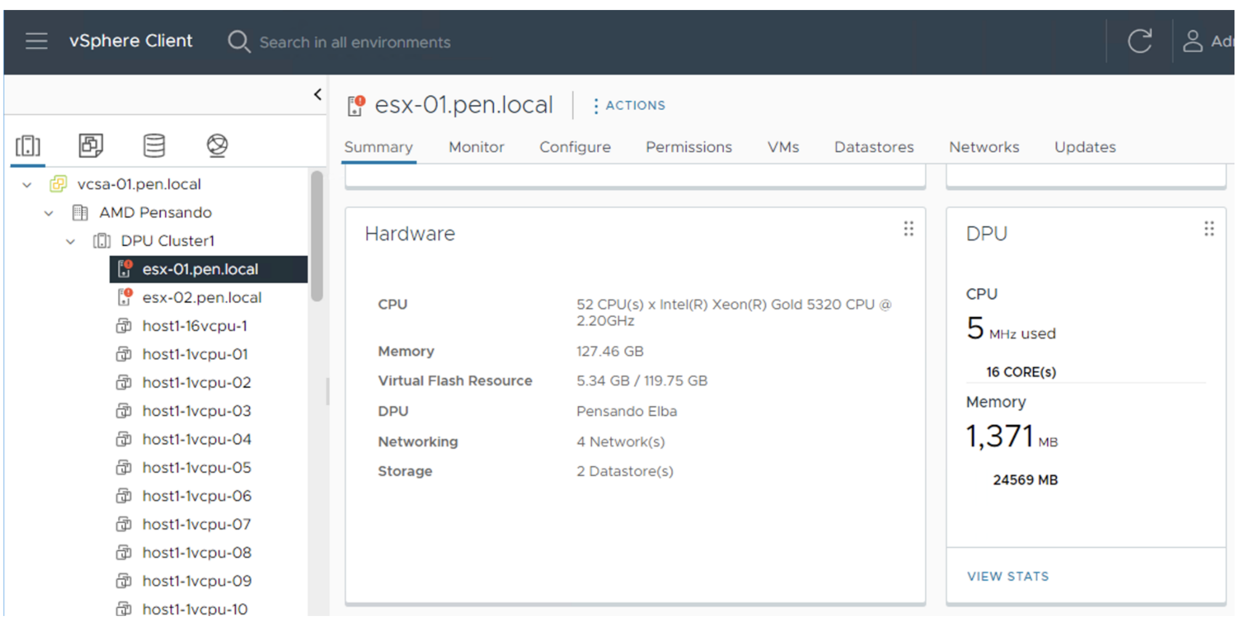




DPU Performance and Health Monitoring via vCenter

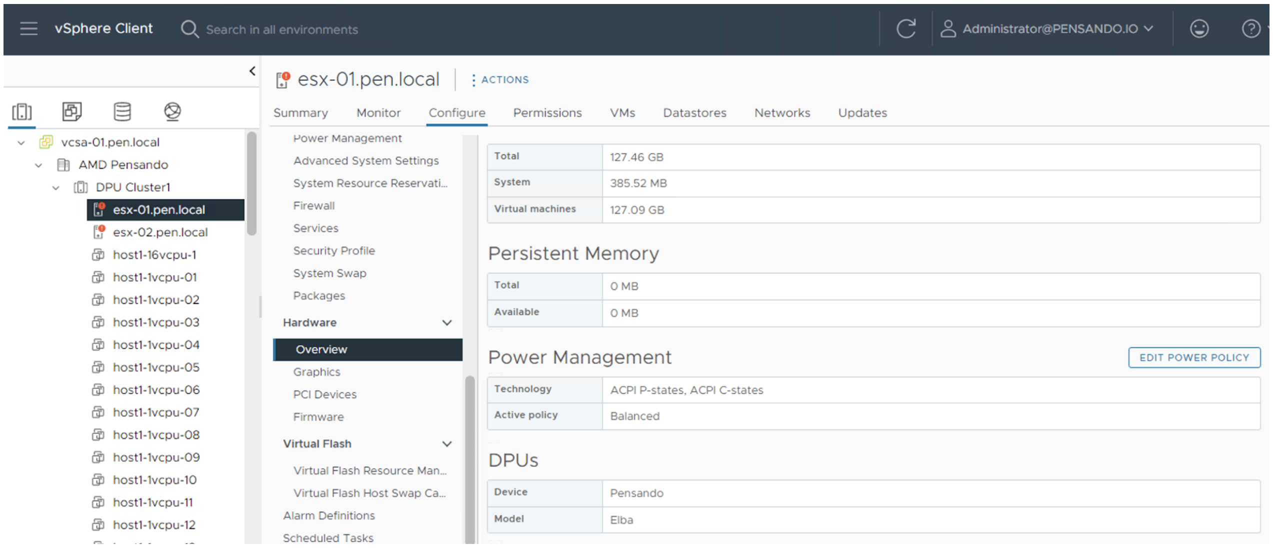
The vSphere-on-DPU architecture promotes DPUs into a first-class citizen within the vCenter management framework. From the UI and API, detailed information can be viewed and abstracted to address a variety of environmental and management operations (such as checking DPU link status, verifying firmware data, alarming/triggering on high DPU/CPU utilization, exporting system/DPU logs for troubleshooting).

General DPU Information in vCenter

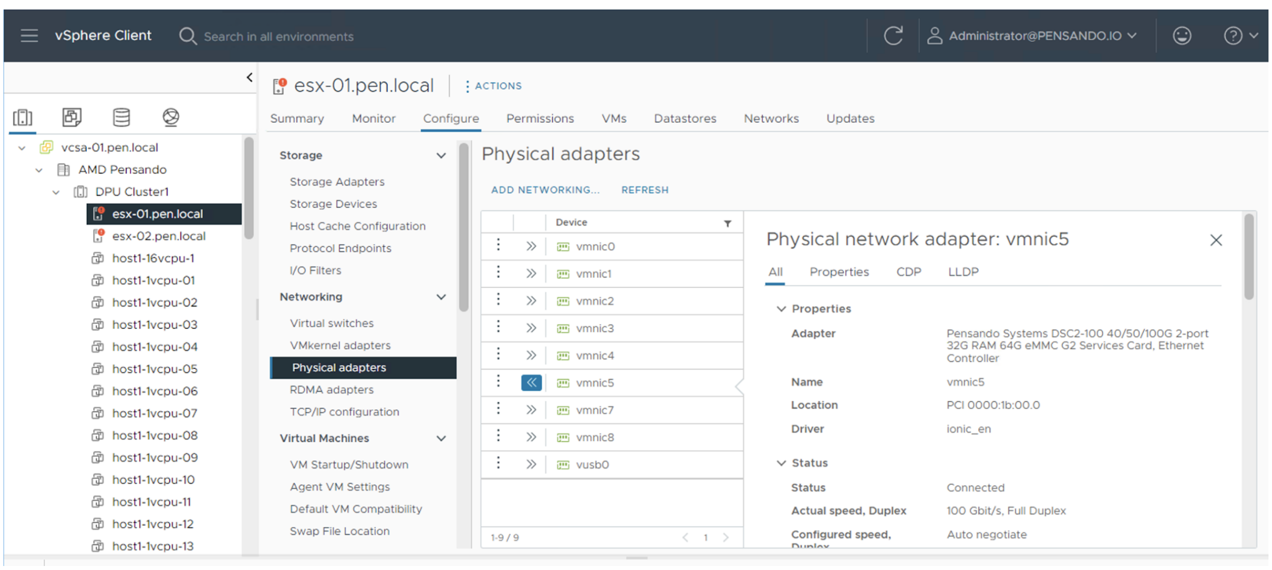


VMware vSphere Distributed Services Engine with AMD Pensando DPU

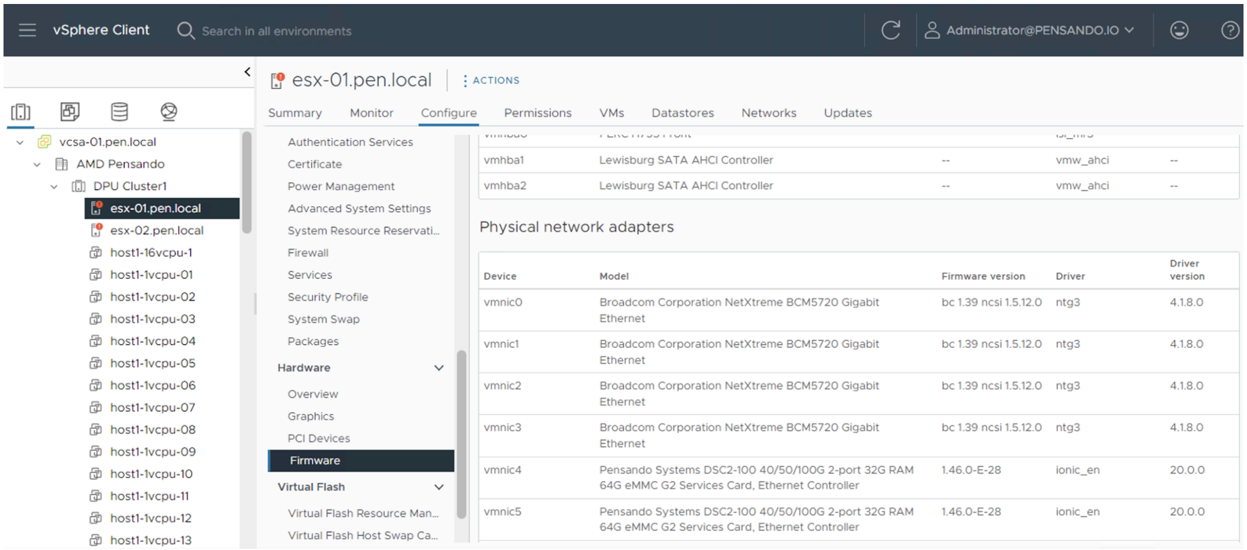
Detailed DPU Information in vCenter



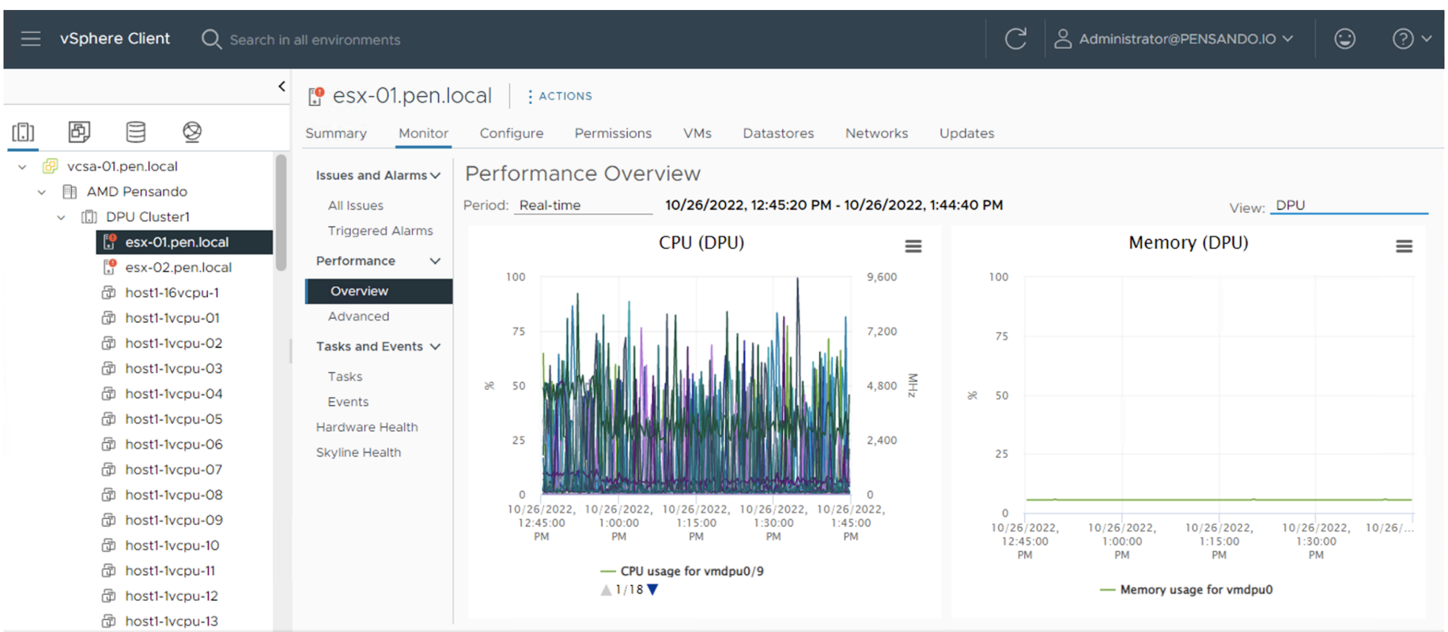
Detailed DPU Adapter Information in vCenter



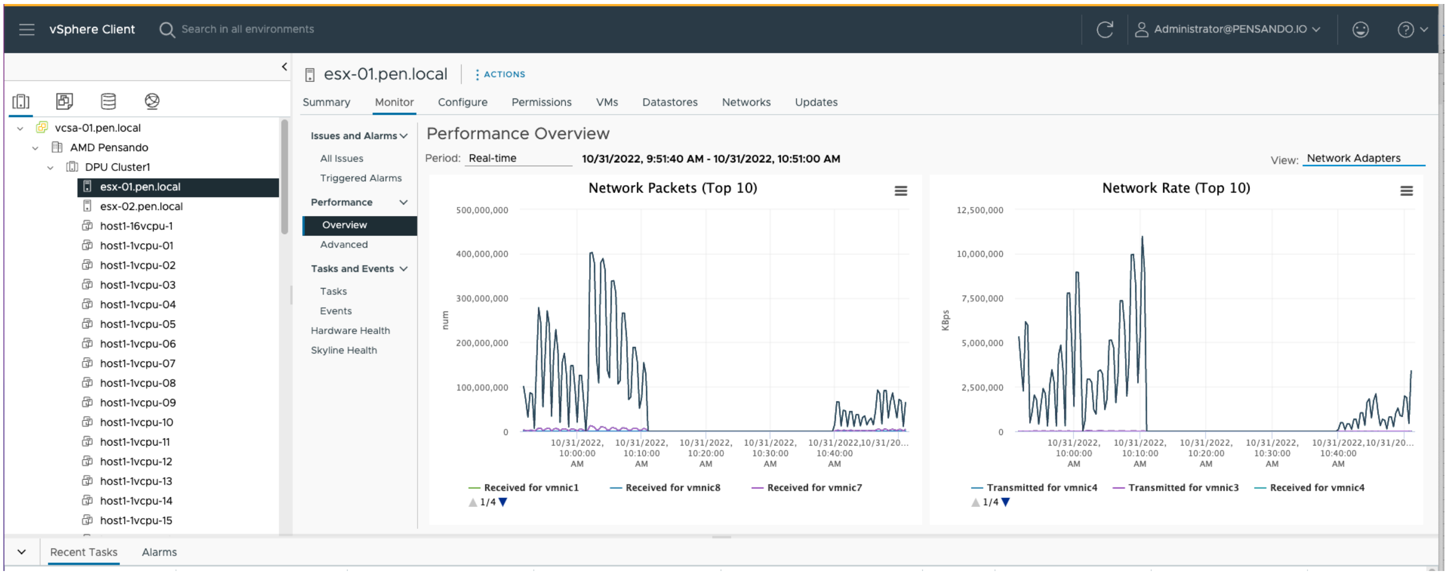
Detailed DPU Firmware Information in vCenter



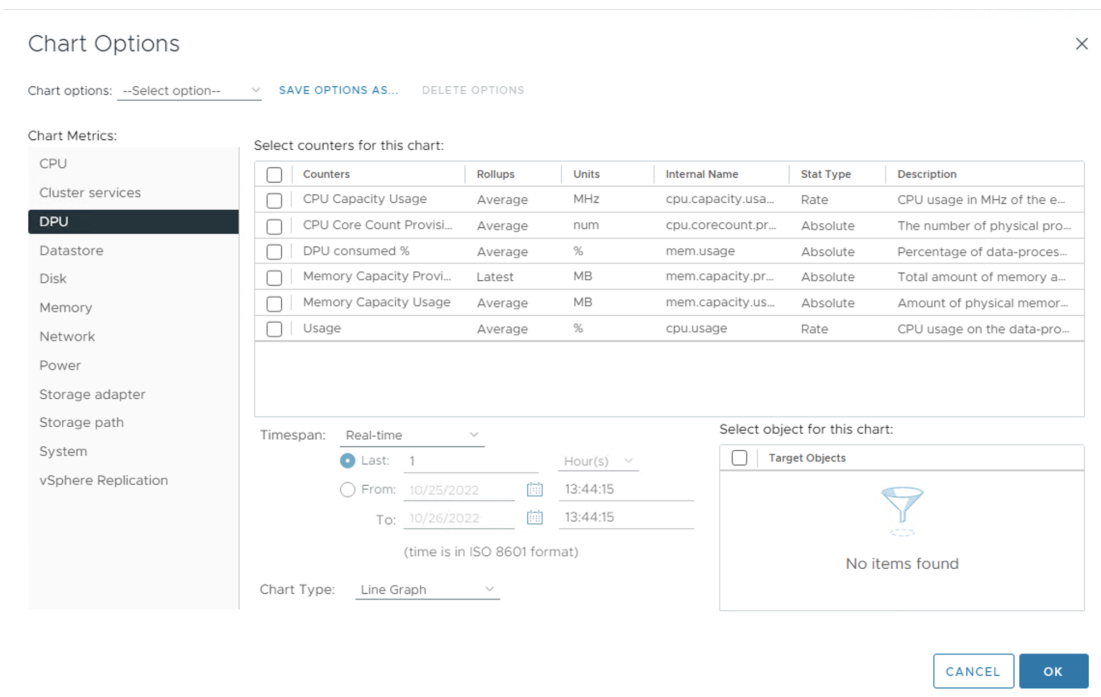
DPU Performance Stats in vCenter



VMware vSphere Distributed Services Engine with AMD Pensando DPU

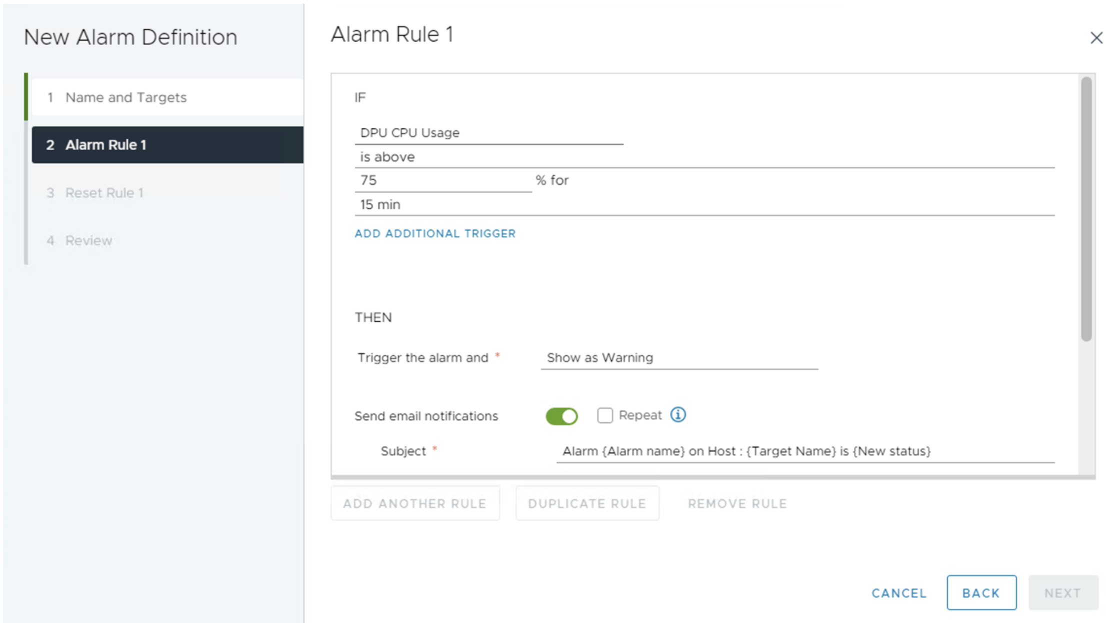


Detailed DPU Metrics for Performance in vCenter

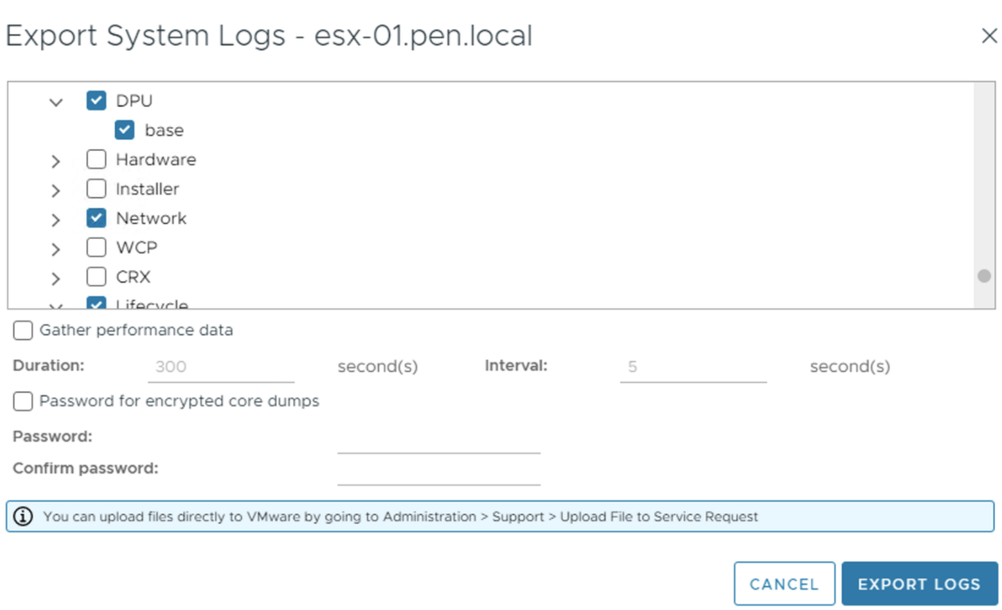


VMware vSphere Distributed Services Engine with AMD Pensando DPU

DPU Alarms in vCenter



DPU Logs in vCenter

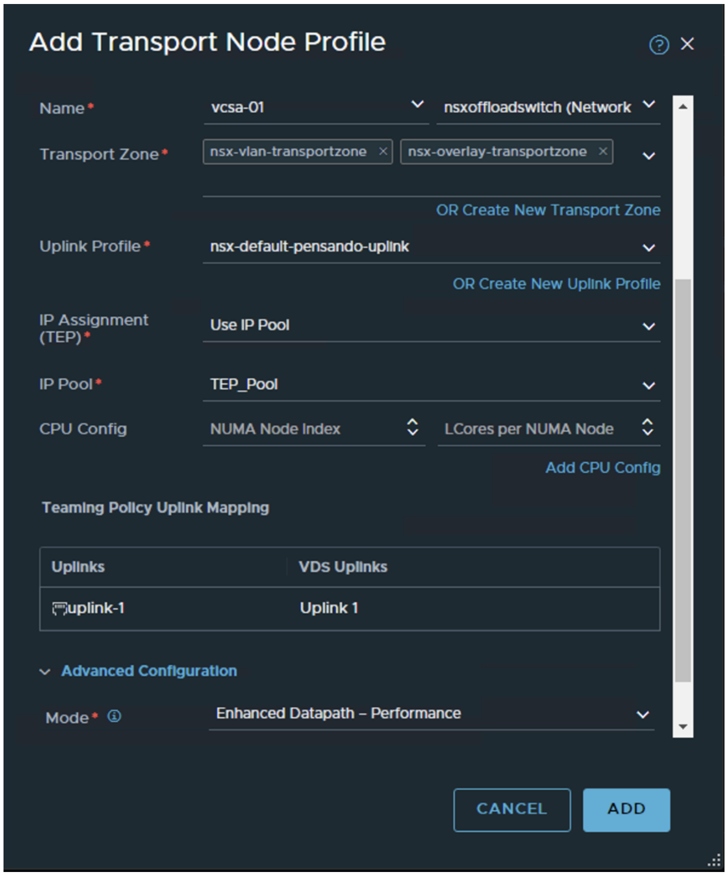


Seamless DPU Lifecycle Management via vLCM

vSphere Lifecycle Manager (vLCM) is a next-generation replacement for vSphere Update Manager (VUM) and is used for host and cluster management within vCenter. vSphere Lifecycle Manager is a comprehensive collection of abilities in vSphere 8 for consistency across ESXi hosts and an easier way to update hosts. vLCM is built on a desired-state, or *declarative*, model that provides lifecycle management for the ESXi hypervisor. These packages include a full stack of drivers and firmware for the servers powering a data center based on vSphere on DPUs.

Below is an example of the process of enabling a DPU-backed virtual switch in Enhanced Data Path (EDP) mode within NSX manager. All current NSX workflows remain the same (integration to vCenter, creating overlay and L2 backed transport zones, security and networking policy orchestration, etc.) and the only additional step necessary with any standard deployment of a transport node profile is the selection of *Enhanced Datapath**, enabling ESXi nodes connected to this NSX-managed DVS to run UPT/VMDirectPath for the application VMs running on the x86 hypervisor.

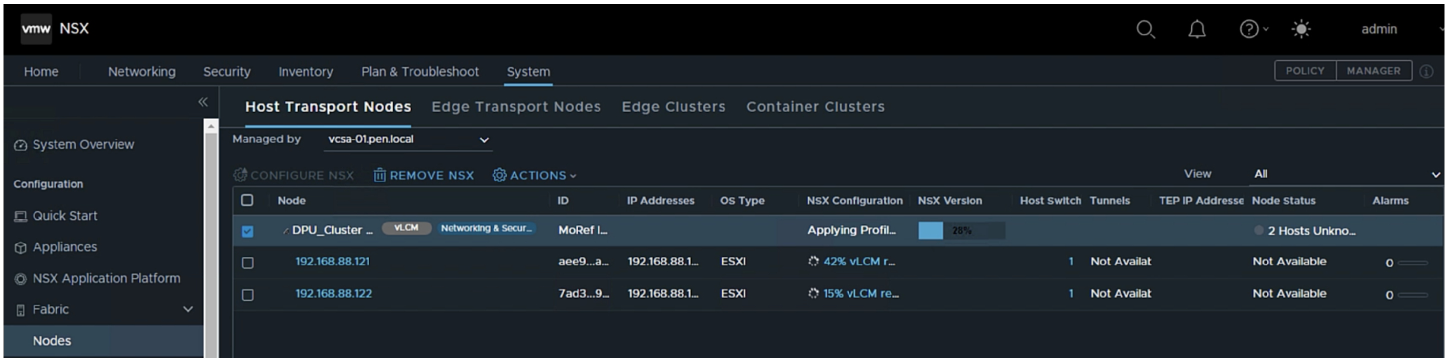
<u>vLCM to Install NSX on DPU and Enable Enhanced Datapath</u>

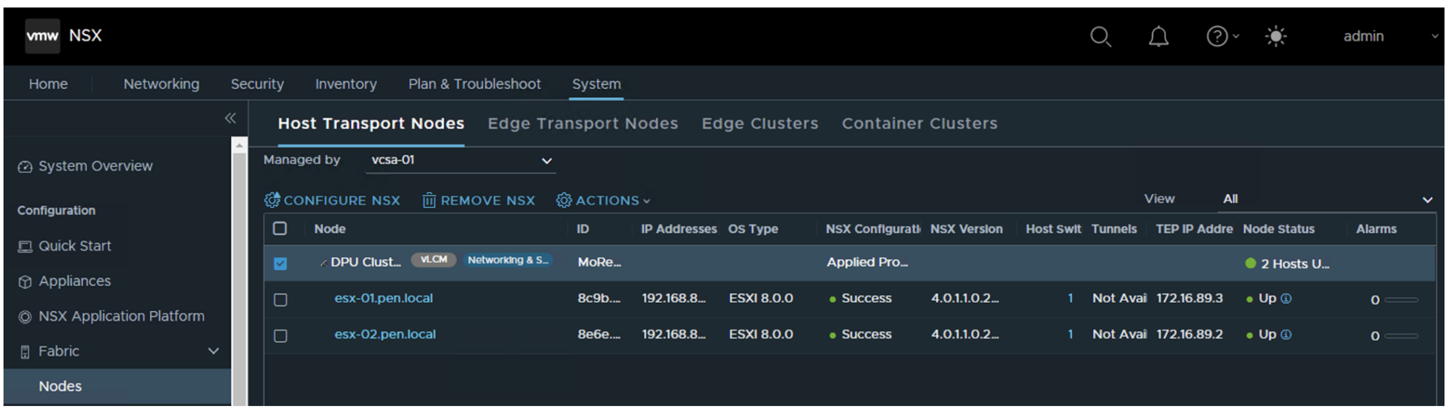


*VMware recommends **Enhanced Datapath – Standard** for current 8.0 builds

VMware vSphere Distributed Services Engine with AMD Pensando DPU

vLCM Installing and Configuring NSX on the DPU



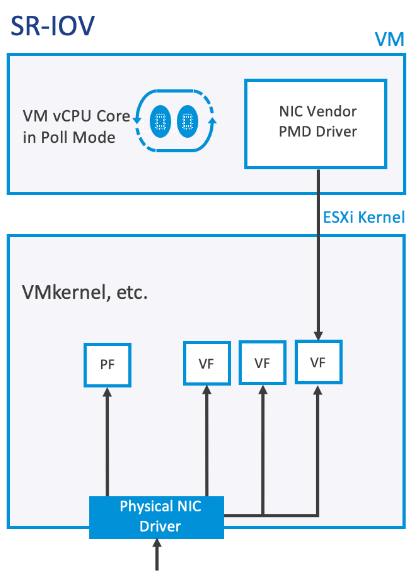


Network Performance with vSphere Distributed Services Engine

vSphere Distributed Services Engine offloads and accelerates vSphere Distributed Switch and NSX Networking to the AMD Pensando DPU. NSX provides networking services (for example, management of overlay or L2-backed port groups in vCenter, and load balancing) and security services (for example, distributed firewall, advanced threat protection via IDS/IPS, and network traffic analytics) to the VM or container-based applications/workloads running on the ESXi hypervisor.

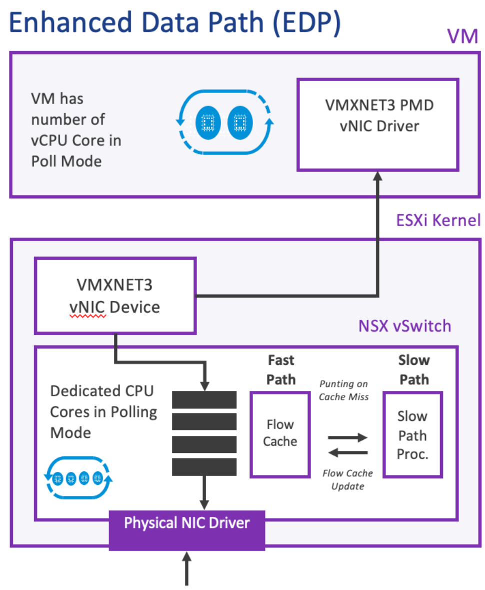
Historically, vSphere has offered several options to enable passthrough with a goal of boosting performance. Single root I/O virtualization (SR-IOV) is an option available in vSphere that can provide virtual machines shared access to physical network cards installed in the hypervisor. It is very similar to PCI passthrough, which grants exclusivity of a device to a single virtual machine; but SR-IOV allows a single physical PCIe® device to appear to be multiple separate physical PCIe devices. SR-IOV supports the concept of *physical functions* (PFs) and *virtual functions* (VFs). PFs are full featured PCIe functions; VFs are typically more lightweight functions. In a typical vSphere virtual environment, VMs attach to a VF on a physical network interface card with SR-IOV configured.

Any form of PCI passthrough, including SR-IOV, pins a VM to a specific ESXi host, meaning vMotion is not possible, nor is HA failover. Below is an example of SR-IOV on ESXi; notice there is no virtual switch on the hypervisor servicing the VMs:



NSX 4.x introduces two new switching modes: *Enhanced Data Path – Standard* and *Enhanced Data Path - Performance*, which offload and accelerate both networking and security services (ex. Distributed Firewall) on the DPU. Both enhanced data path modes can be enabled outside of servers with DPUs, but the DSE leverages these new modes to support UPTv2 on the guest VMs (discussed below).

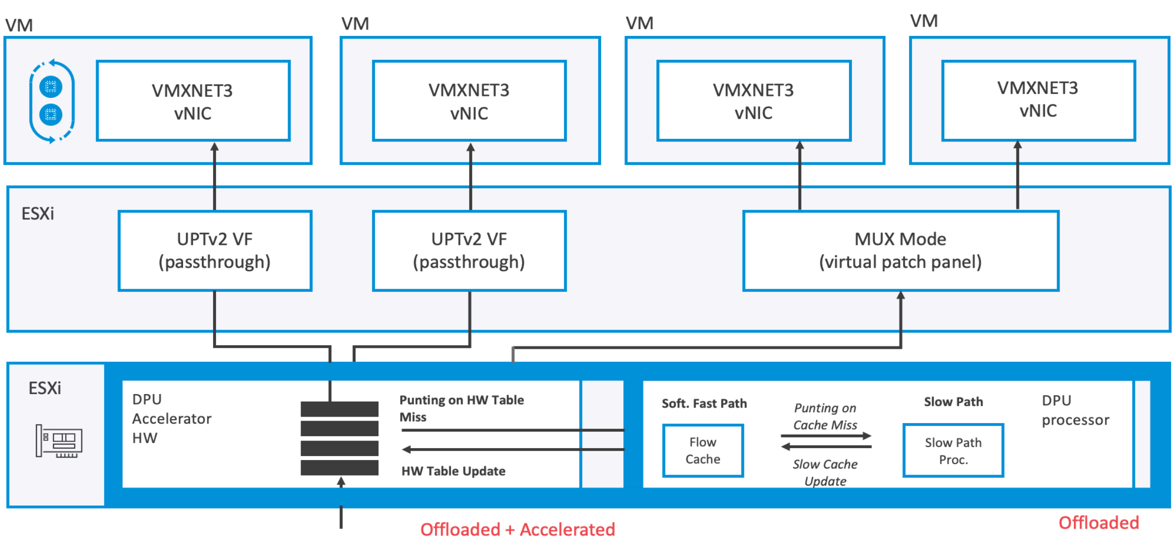
The new Enhanced Data Path modes in NSX 4.x help the system achieve higher throughput, at much lower latency, with minimal CPU overhead on the hypervisor by offloading to the DPU. These new DPU focused data path modes are flexible and configurable, allowing the administrator to customize and optimize the acceleration and offloading behavior at a per VM and per vNIC level. The example below, which depicts Enhanced Data Path - Performance (polling based), leverages the Fast Path in the vSwitch on the hypervisor to achieve very high throughput and packets per second, while delivering very low latency.



VMDirectPath with vMotion and DRS Support

In a typical vSphere environment today (for example, with vSphere 6.7 or 7.0), to get the best network performance for an application, SR-IOV is usually enabled on the ESXi hypervisor. However, this comes with a few trade-offs: first, this typically requires customizations such as running the NIC vendor's device drivers, increasing operational burden. Second, it is only available for core vSphere features (no NSX), and many core vSphere features aren't supported when a guest is configured to leverage SR-IOV on ESXi (e.g., vMotion, HA, or DRS).

The new Enhanced Data Paths, with VMs configured to leverage Uniform Passthrough (UPTv2) or VMDirectPath, is a complete passthrough mechanism, offloading all processing to the DPU. This results in the best networking performance the data path on the DPU hardware can derive. In addition, given the vSwitch in the hypervisor and the leveraging of VMXNET3 driver in the guest VM, vSphere and NSX can now enable all features like vMotion and DRS, along with security features like Distributed Firewall (DFW) to the application VMs running on the ESXi x86 hypervisor.



AMD Microbenchmark: vSphere Distributed Services Engine Performance on 100G AMD Pensando DPU

Looking at workloads, there have been significant changes in application architectures over the past decade: from the original monolithic approach, to more tiered VM-based applications, and to more modern cloud-native based applications leveraging Kubernetes®, applications are growing in complexity, and the way they are deployed is in turn changing drastically. At the same time, many new applications use more and more server CPU cycles.

Traditionally, the IT industry has relied on the CPU for everything—application business logic, specialized workloads like 3D modeling, processing network packets, and so on. To reclaim compute resources for applications, hardware accelerators have been developed, including GPUs, FPGA- based SmartNICs, and now DPUs. By leveraging these accelerators, organizations can both improve performance of the offloaded activities and free up CPU cycles for core app processing work.

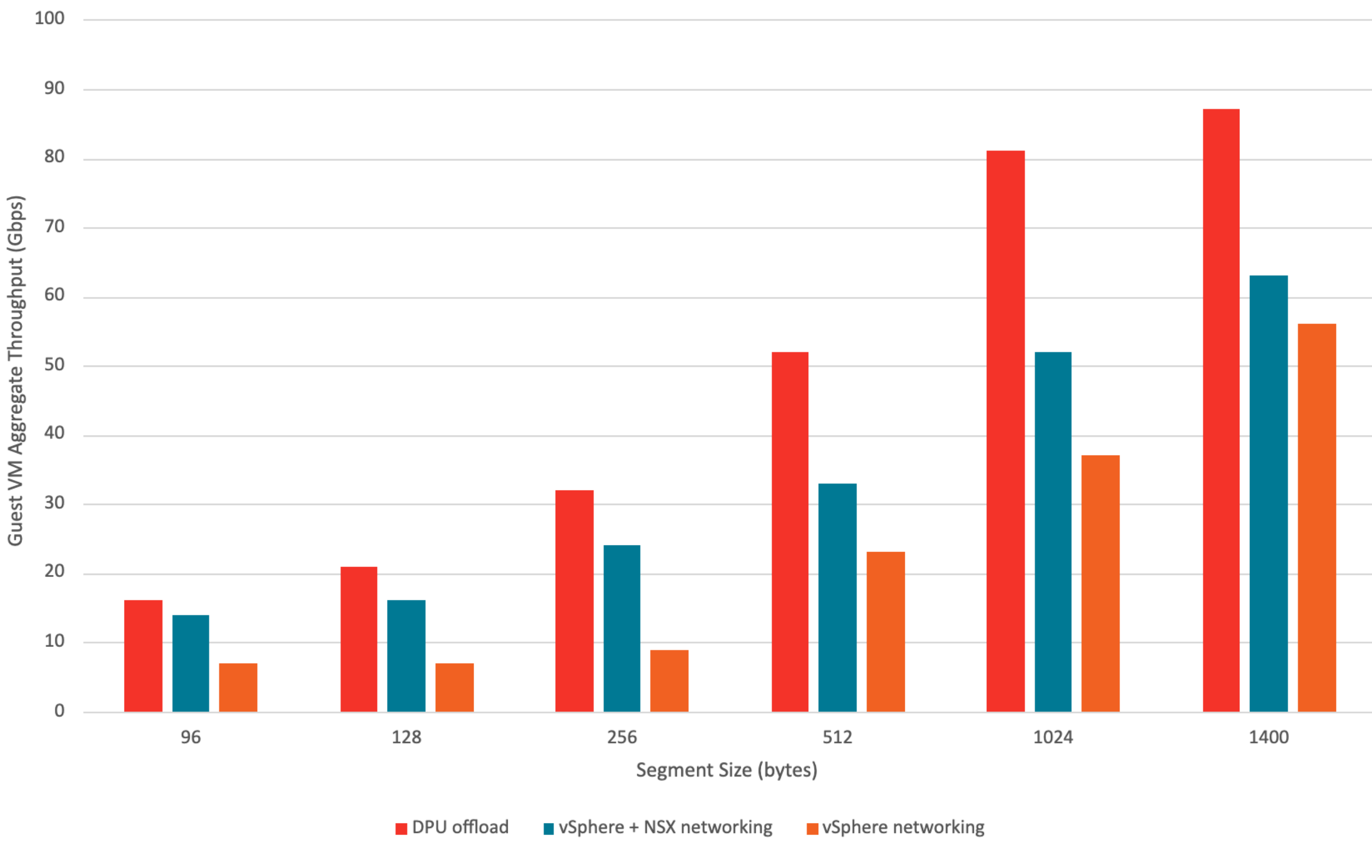
At the same time, security risks are continuing to proliferate, especially with applications that are distributed across many locations, which breaks apart the traditional perimeter security model. Security now needs to be distributed broadly yet enforced locally. Delivering and implementing this requires the proper hardware enforcement of those security policies and boundaries. As many organizations continue to drive toward a zero-trust security model, rather than loading up CPUs with yet another infrastructure workload demand, offload accelerators like DPUs can also help implement distributed security. Additionally, a DPU-based security architecture isolates security controls (e.g., App Firewall Rules) from the attack surface of the application. With vSphere Distributed Services Engine, where vSphere networking and NSX run on the DPU, if a breach does occur within a guest VM, there is a new layer of protection between the application and ESXi hypervisor on the server, and the ESXi software running on the DPU.

Addressing these changes in workloads and security demands make the performance and efficiency benefits of DPUs very attractive. The next section contains examples of performance-related improvements on a traditional vSphere x86 server (e.g. throughput, bandwidth, and latency) where everything is running on the main CPU, vs. an x86 server with the AMD Pensando DPU and vSphere Distributed Services Engine, where the networking and security functions are being both offloaded and accelerated using the software fast path backed by the AMD Pensando DPU's P4-enabled data path to deliver exceptional performance. The topology and methodology are simple: build a setup leveraging the VMware GA image for vSphere 8 and NSX-T 4 on the servers, deploy various 6 core (2.2 GHz) Ubuntu® 18.04 based VMs (VMXNET3) on ESXi, and use industry standard tools for benchmarking and performance data.

VMware vSphere Distributed Services Engine with AMD Pensando DPU

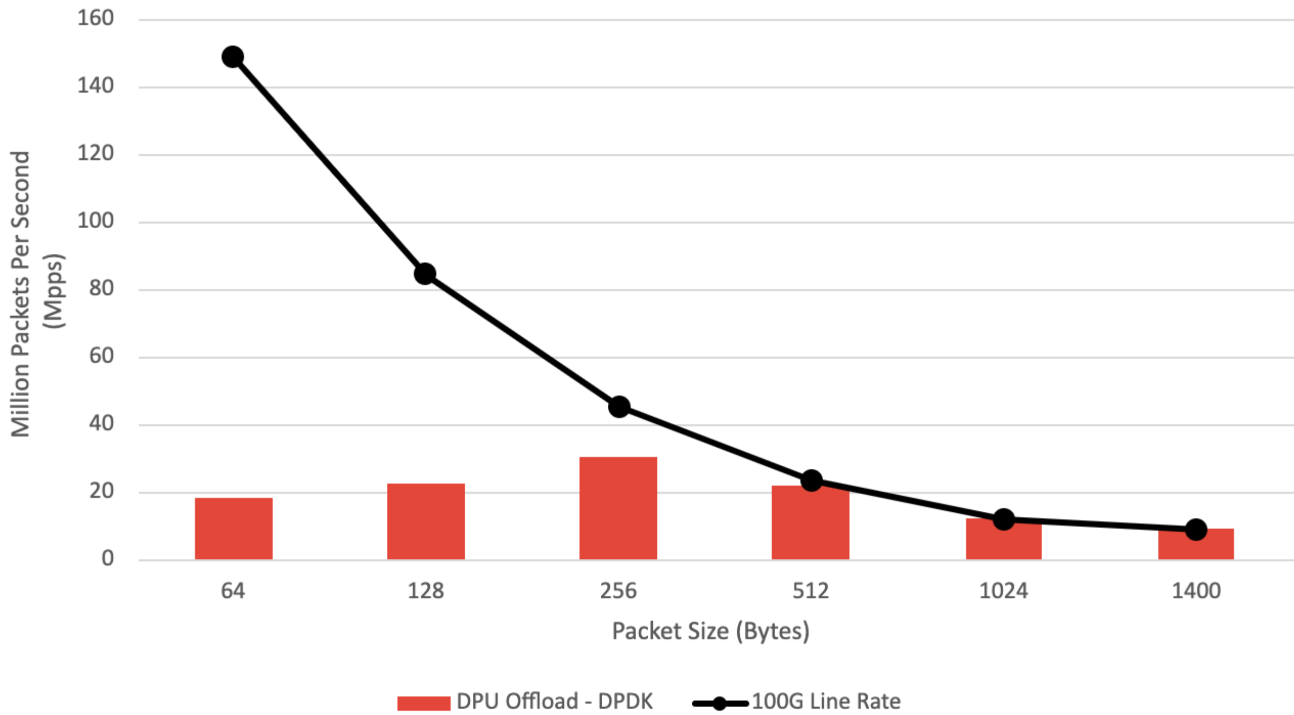
As shown in the diagram below, we used connected ESXi hosts with a modern low-latency data center switch. Results were generated with tools such as `iperf3` and `sockperf`.

Throughput Analysis: (6 vCPU 8 VM DPU Offload vs. Non-Offload)[2]



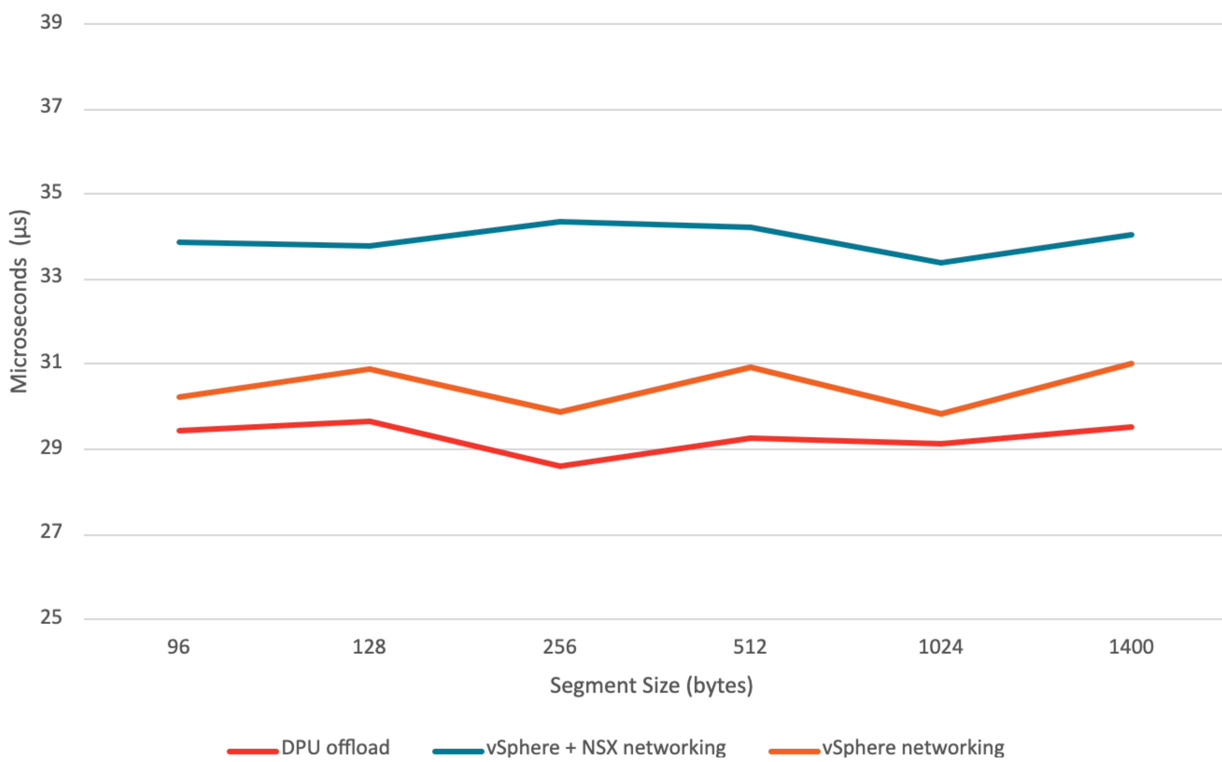
This first throughput performance test compares offloads running on the DPU (first bar) to two non-offloaded scenarios (nothing 'offloaded' on the DPU). Both non-offloading scenarios represent an industry-standard performance NIC, the vSphere + NSX bar is leveraging EDP-Standard configured on the DVS. Note that the 100G DPU approaches near line-rate performance with larger packet sizes (1024 and 1400 byte segments) and that there are scenarios where the DPU performance is 2x-3x the performance of regular vSphere-based networking with a 100G-capable standard/performance NIC (128, 256, 1024 byte segments).

Maximum Packet per Second (PPS) Analysis: 2vCPU 8VM DPDK[2]



The next test introduces a slight variation to the previous setup. Customized workloads and applications that require maximum performance typically are different from standard kernel-based applications. Specifically, they will leverage things like DPDK or SR-IOV to get the greatest performance possible on the ESXi hypervisor. For these specific test runs, we have built separate VMs and have enabled DPDK in the Ubuntu VMs to test the maximum packets per second (PPS) the DPU and fast-path can handle, removing the application layer from imposing any bottleneck. The rest of the topology and setup remains the same. The unidirectional result is that we observe over 30M PPS on the AMD Pensando DPU with 256 byte frames as we approach 100G line rate.

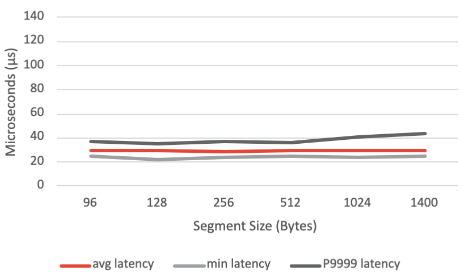
Latency and Jitter Analysis: 6vCPU One-Way Average Latency[2]



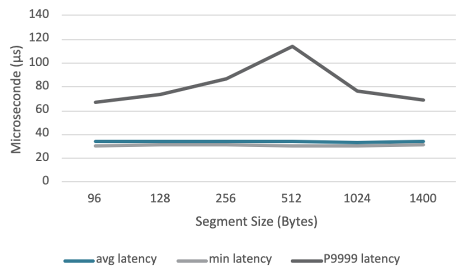
The above graph shows average latency plotted across the original three test scenarios. This demonstrates that even with vSphere and NSX networking running on the DPU, we see the lowest latency across the board. While introducing NSX onto the x86 CPU adds some overall latency (as would be expected with features like DFW and overlay enabled), offloading and accelerating these services on the DPU unlocks the lowest average latency for all packet sizes.

VMware vSphere Distributed Services Engine with AMD Pensando DPU

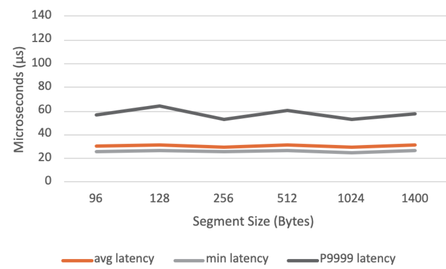
DPU Offload[2]



vSphere + NSX Networking[2]



vSphere Networking[2]



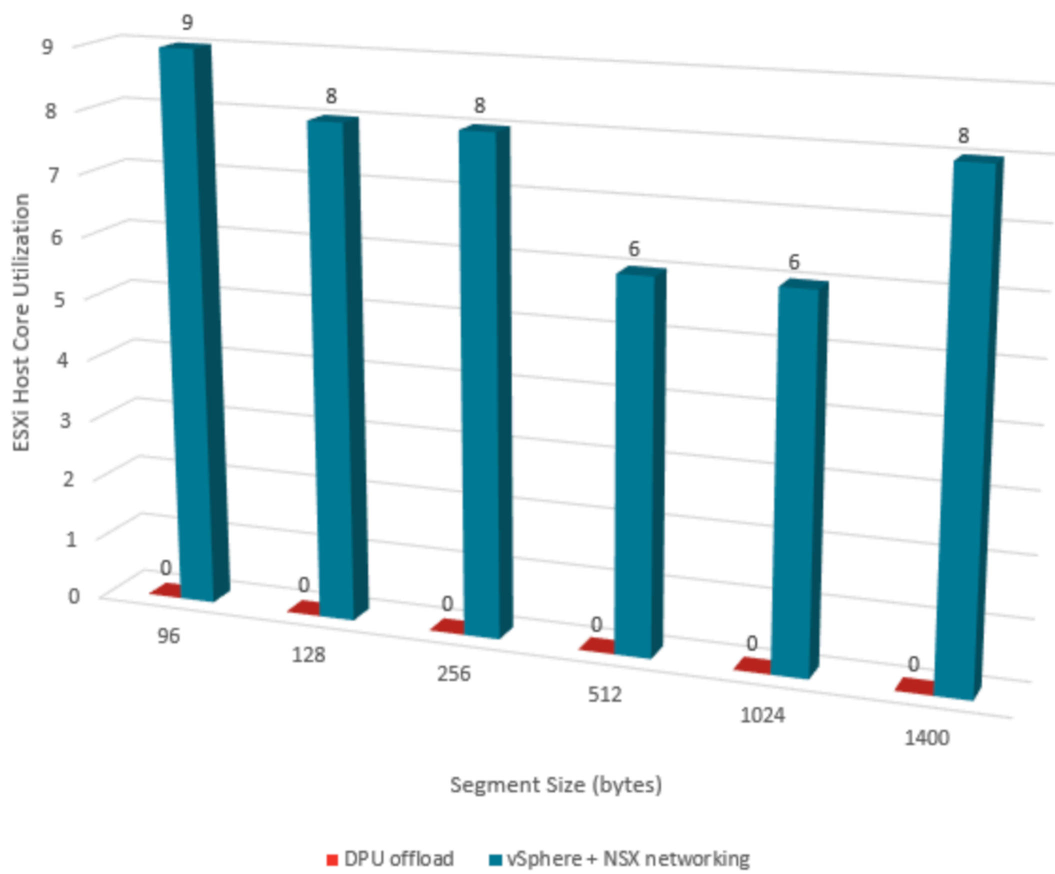
Another observed benefit of the DPU is related to jitter. We also observed when plotting the highest percentile data (P9999), which essentially shows us the spikes (or maximums) in the raw data results, the DPU offload data is very flat (i.e., deterministic) when compared to the spikes observed when relying only on networking on the x86 host. Tail latency, or high-percentile latency (P9999), is the indicator that best represents flow completion time and ultimately application performance. Comparing the P9999 latency between the “offload” result on the AMD Pensando DPU vs vSphere + NSX “non-offload” result at 512 bytes as an example, shows over a 3x reduction in tail latency.

In summary, the DPU data for all packet sizes shows lower latency and deterministic jitter when compared to running these services on the server’s x86 CPUs, which is very important for performance-focused applications and workloads where minimal delay matters. Even in applications where minimal delay isn’t as important, consistency of that delay is critical to almost every application, which is the result when networking and security are offloaded to the DPU.

CPU Utilization Analysis

In addition to raw networking performance gains, another aspect of offloading to the DPU is the freeing of critical x86 cores within a system’s primary CPU complex. This is important not only from an investment standpoint (in terms of both CapEx spend of the infrastructure itself, but also the OpEx spend when licensing your hypervisors on a per-core-count basis) as customers prefer to use their budgets for revenue-generating applications and workloads, not licensing CPU cores that wind up being consumed by basic infrastructure services. Below, we have measured the amount of ESXi on x86 cores consumed just to provide the basic networking for our Ubuntu VMs.

CPU Utilization Analysis: 6vCPU 8VM Host CPU Consumed[2]



Two noteworthy data points to consider: first, this validates the Distributed Services Engine claims that UPTv2 (DPU offload) delivers near zero CPU consumption by ESXi. Second, for the 100G test performed, this shows up to 9 cores cores returned to the x86 CPU when offloading just the networking and DFW services to the DPU.

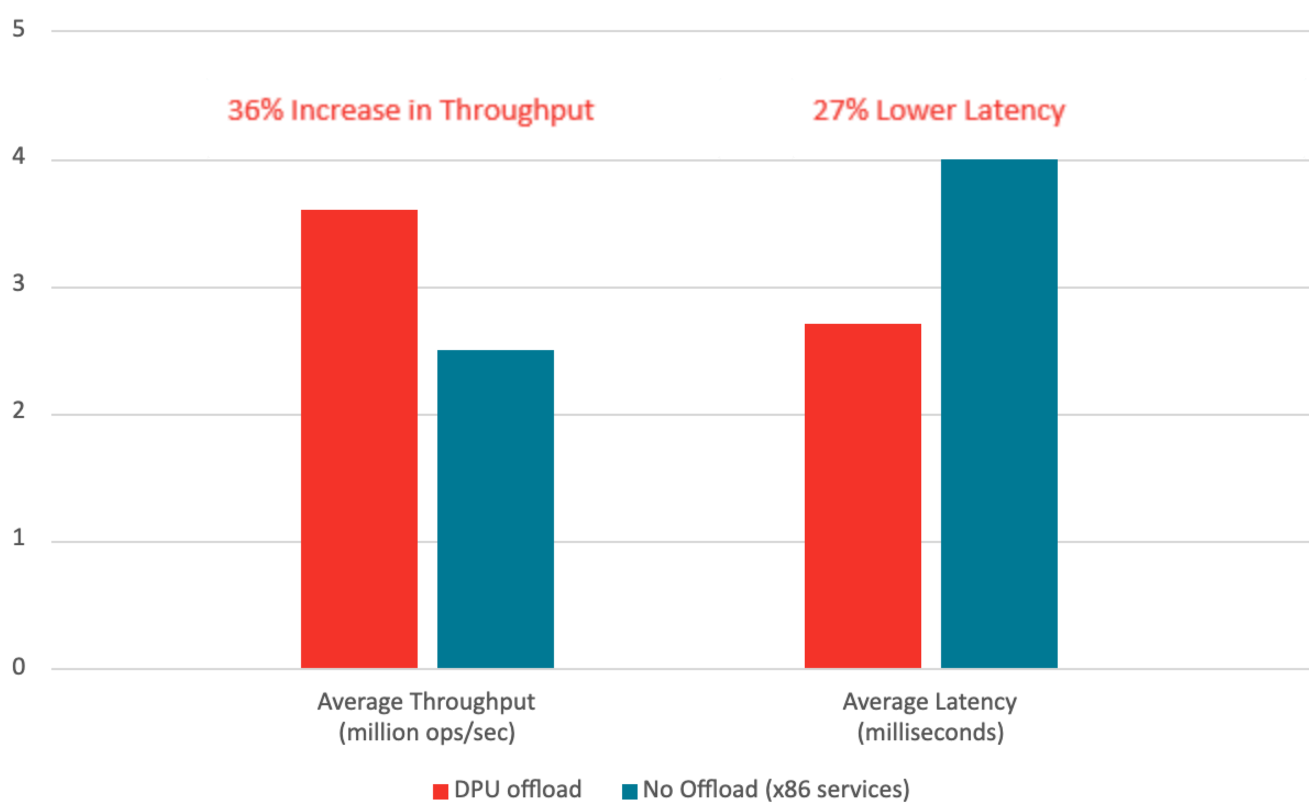
As the sweet spot for host/server bandwidth continues to increase from 10G to 25G/40G to 100G, the amount of cores needed to support higher bandwidth and next-generation security services will in turn increase significantly. Consider core consumption scenarios based on various networking and advanced security services today at 10G. Basic networking and DFW at L4 consumes a small number of cores, maybe 1-2 cores in total. For L7 DFW, which is more processing intensive, consumption becomes slightly higher, whereas more advanced security features like IDS/IPS could consume even more cores. Something like IPsec/TLS encryption could consume anywhere from 8 to 10 cores itself.

These are estimates based on services at 10G (not precise measurements; environment/application variables may impact results) and will drastically increase when the services need to be provided at higher speeds such as 25G or 100G. As we observed above, our testing results yielded roughly 9 cores being saved when vSphere networking and NSX overlay with NSX DFW at L4 are offloaded to the DPU at 100G. Because these network and security services are offloaded and accelerated on the AMD DPU and no longer consuming these primary CPU cores, each server in the data center that has a DPU will be able to scale to higher VM density than previously available. Increased density and higher throughput with lower latency can bring significant application performance improvements.

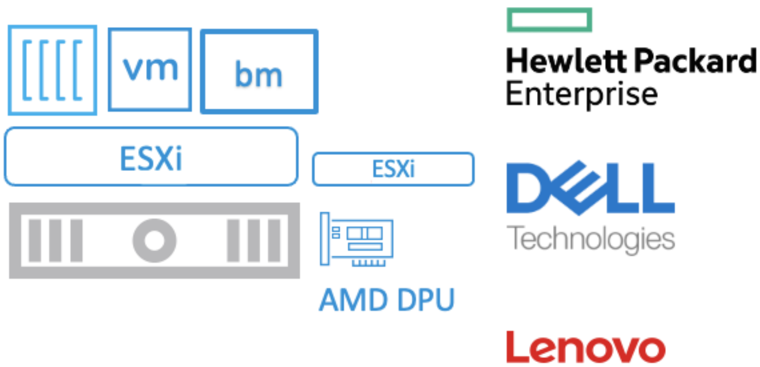
VMware vSphere Distributed Services Engine with AMD Pensando DPU

As more and more infrastructure services are offloaded from the x86 CPU and moved to the DPU, the benefit in terms of reclaimed CPU resources becomes clear. The cores being saved allow vSphere and application administrators to run additional workloads, which drives up VM density and provides better resource utilization. In addition, the increased VM density per physical server node can result in a reduced overall server footprint and lead toward a greater overall ROI model—specifically as it relates to CapEx (dollars spent on infrastructure) and OpEx (rackspace, software licensing, power, cooling, data center or colo footprint, admin costs, etc).

Application Performance: In-Memory Database Testing with vDSE Acceleration[3]



AMD: The Lead Platform for VMware vSphere Distributed Services Engine



vSphere Distributed Services Engine with AMD Pensando DPUs, in conjunction with system partners, delivers VMware's new vision for a software-defined, hardware-accelerated data center.

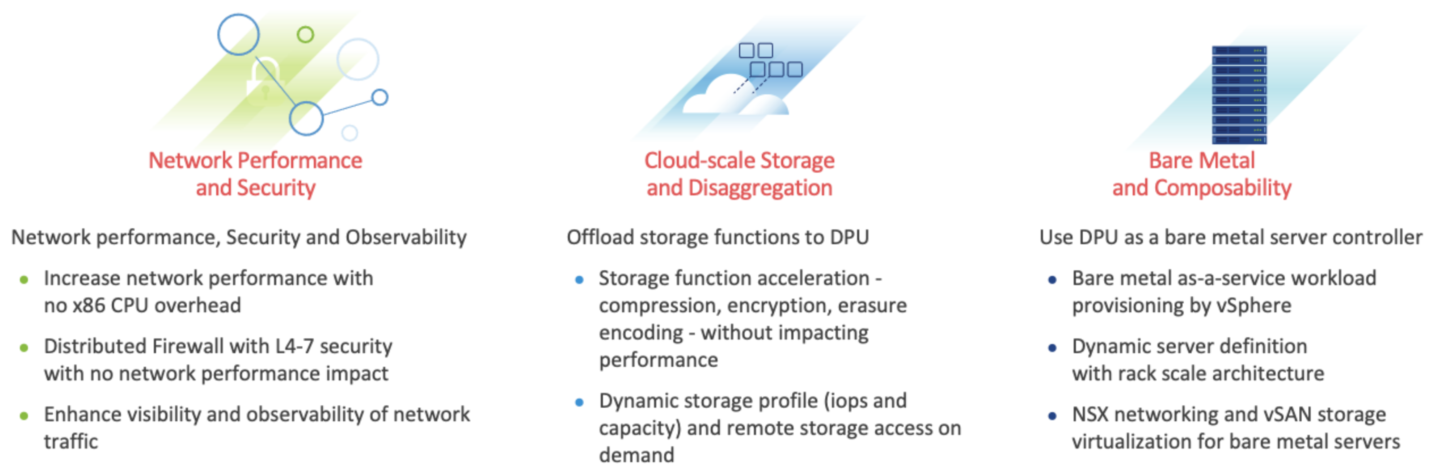
vSphere Distributed Services Engine gives VMware customers the ability to accelerate and secure workloads on best-of-breed next-generation composable systems enabled by Dell, HPE, Lenovo and AMD. These solutions can deliver exceptional vSphere performance, offer enhanced security over today's current model where security controls are not isolated from the applications they are intended to protect on x86, and provide the most efficient use of CPU cores for applications and workloads running on vSphere. AMD Pensando DPUs unlock infrastructure-bound CPU cycles, and help customers meet the throughput, latency, and security needs of modern distributed workloads.

VMware, server OEMs, and AMD have collaborated to transform the DPU into a first-class citizen within next-generation servers. iDRAC/iLO and VMware vCenter manage the entire lifecycle of the DPU, enabling operational simplicity and consistency for customers as they start to enable vSphere/VCF clusters with DPUs.

Previously known as "Project Monterey", VMware vSphere 8 introduces dozens of new features to help customers bring the benefits of the cloud to their on-premises infrastructure and will deliver vSphere on DPUs. This approach, pioneered by hyperscale and cloud service providers, offloads all infrastructure service functions to DPUs, delivering a single robust operating model across all workload types (including bare metal), helping ensure isolation of the workload domain from the infrastructure domain.

Networking and security offloads and acceleration is just the start. In subsequent releases of vSphere, VMware has the ability to deliver even more services enabled on the DPU, including storage offloads, extending infrastructure services to bare-metal instances, and full bare-metal lifecycle management under vCenter:

VMware vSphere Distributed Services Engine with AMD Pensando DPU



- **AMD Is the Clear Platform of Choice for vSphere 8:** Of the three DPU vendors chosen by VMware for the vSphere Distributed Services Engine (or vSphere on DPUs), AMD is the *only* vendor with supported solutions from both vDSE launch server partners (Dell and HPE).
- **AMD Offers More Flexibility:** AMD is the only vendor at launch with 10G/25G and 100G solutions on the market.
- **AMD Is the *Only* Proven DPU Platform:** AMD is the only vendor with large scale of DPU units deployed in production with hyperscalers, public clouds, and enterprise customers.

Disclaimer
The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. Any computer system has risks of security vulnerabilities that cannot be completely prevented or mitigated. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes.
THIS INFORMATION IS PROVIDED 'AS IS." AMD MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE CONTENTS HEREOF AND ASSUMES NO RESPONSIBILITY FOR ANY INACCURACIES, ERRORS, OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION. AMD SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY, OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT WILL AMD BE LIABLE TO ANY PERSON FOR ANY RELIANCE, DIRECT, INDIRECT, SPECIAL, OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION CONTAINED HEREIN, EVEN IF AMD IS EXPRESSLY ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

AMD, the AMD Arrow logo, Pensando, and combinations thereof are trademarks of Advanced Micro Devices, Inc. VMware ESXi™, VMware vSphere® vMotion® and VMware vCenter® are trademarks of VMware. Kubernetes is a registered trademark of The Linux Foundation. PCIe is a registered trademark of PCI-SIG Corporation. Ubuntu and the Ubuntu logo are registered trademarks of Canonical Ltd. Other product names used in this publication are for identification purposes only and may be trademarks of their respective companies.

© 2023 Advanced Micro Devices, Inc. All Rights Reserved.

PWP22007

amd.com/pensando

End Notes

[1] PEN-001: As of 8.1.2022, the AMD Pensando DPU is based on a 7nm process with support for 400G throughput, placing it at least one generation ahead in process node over any other GPU on the market; closest competitor is based on a 16nm process.

[2] PEN-008: Testing conducted by AMD Performance Labs as of November 20, 2022, on the AMD Pensando 100Gb/s DPU on a production system comprising of Intel Gold 5320 26-Core processors and 128GB of DDR4 memory running at 2933MT/s with Vmware ESXi version and NSX version with DPU offload in EDP-Standard mode. VMs were running Ubuntu 18.0.4, iperf version 3.1.3, and sockperf version 3.9, iperf3 and sockperf VMs were configured with 6vcpu, 16GB of memory and 25GB of disk space. Testing with iperf and sockperf used packet sizes ranging from 96-1400 bytes. Pktgen version 22.04.1 and DPDK version 22.07.0. Testing with pktgen used packet sizes ranging between 64-1400 bytes. PC manufacturers may vary configurations, yielding different results.

[3] Project Monterey SmartNIC Network Offload. VMware, August 30, 2022 (8:03 in video)