



# Secure and Accelerate Your Infrastructure with VMware vSphere and NVIDIA BlueField-2 DPU on Lenovo ThinkAgile VX Series

## Positioning Information

Traditionally, the industry has relied on the CPU for everything – application business logic, processing network packets, specialized work such as 3D modeling, and more. As application requirements for compute continue to grow, hardware accelerators, including GPUs, FPGAs, etc., have been developed for processing workloads that could be offloaded from the CPU. By leveraging these accelerators, organizations can improve performance for the offloaded functions and free up CPU cycles for application processing.

The increasing demand for modern and distributed applications has given rise to a new foundation for a comprehensive and innovative security offering – a class of hardware accelerators, such as DPUs, that offload, accelerate, and isolate workloads from the CPU.

### DPUs and data-centric architecture

Data Processing Units (DPUs), popularly known as SmartNICs, are programmable integrated systems on a chip (SOC) that combine high-performance CPU, network interface, and data center function accelerators into a single application specific integrated circuit (ASIC). DPUs can improve performance for AI, machine learning, and other specialized tasks by freeing up CPU cycles for application processing.

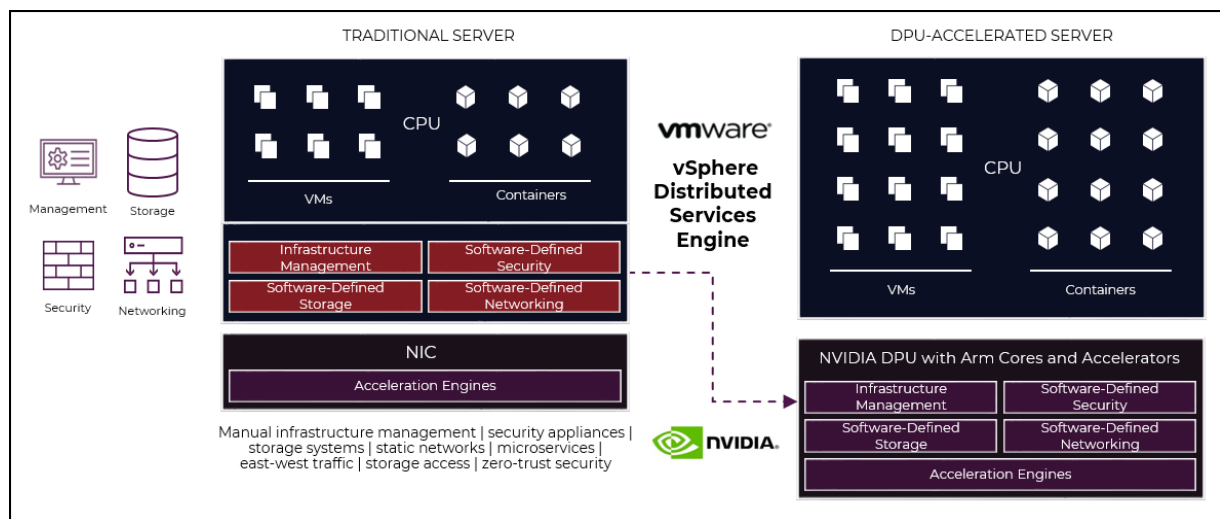


Figure 1. Data Processing Unit (DPU) – Software-Defined, Hardware-Accelerated Infrastructure on a Chip

vSphere on DPUs, formerly known as [Project Monterey](#), modernizes cloud infrastructure into a distributed architecture to meet the throughput and latency needs of modern distributed workloads by offloading and accelerating infrastructure services, such as networking and security, on NVIDIA BlueField-2 DPUs running on Lenovo ThinkAgile VX Series. This solution enhances performance by providing more CPU resources to workloads and reduces the operational overhead of DPU lifecycle management with integrated vSphere workflows.

DPUs provide a comprehensive security offering that utilizes programmable hardware and flexible acceleration engines to offload and improve application performance for zero-trust security, telecommunications, storage, and more.

## An innovative approach to Zero-Trust Security

Businesses across industries are embracing a zero-trust philosophy to improve network security at every level. The security model of most organizations' data center networks is like the shell of an egg that is only protected from outside traffic using perimeter firewalls. However, if an attacker finds a way to penetrate, then the shell is cracked, and everything inside is exposed, and attackers can gain control of the infrastructure services and access company data

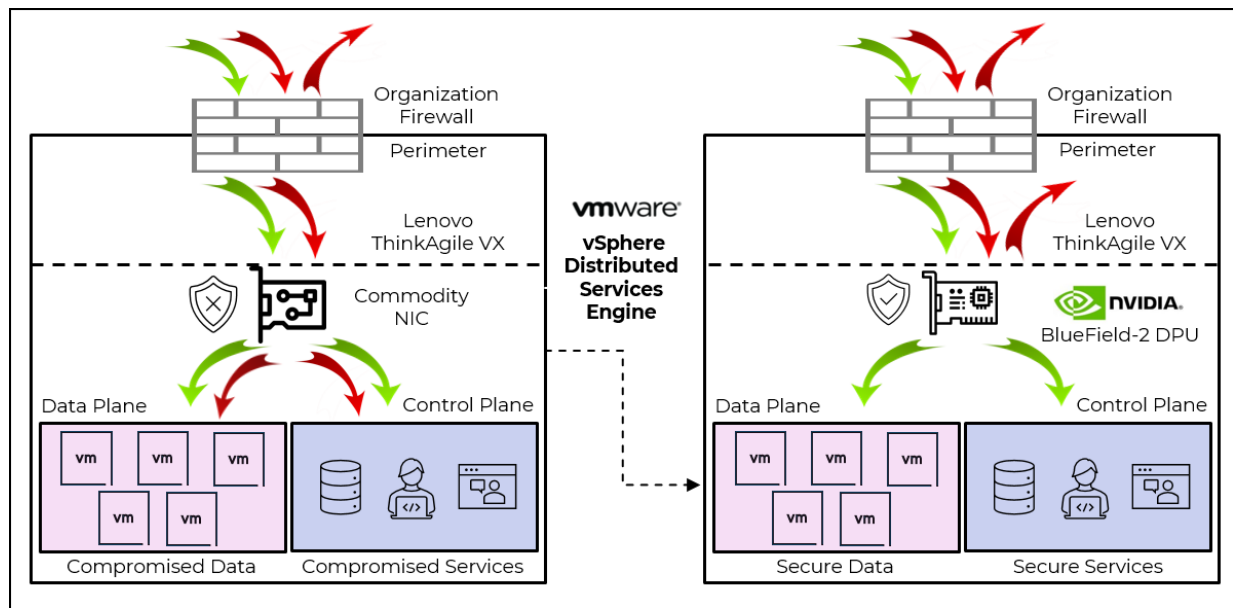


Figure 2. Zero-Trust Security – Combatting growing security threat

NVIDIA BlueField-2 DPUs on Lenovo ThinkAgile VX Series enables distributed firewalls deployments and micro-segmentation of the network layer, providing additional protection that is critical to implementing a modern, zero-trust security model.

Additionally, there is enhanced visibility and observability of the network traffic with better encryption, isolation, and protection through network acceleration with VMware NSX that enables a distributed firewall, with Layer 2 to Layer 7 network security that doesn't impact performance.

## Accelerate and optimize IT infrastructure performance

By offloading network traffic functions from host or server CPUs to DPUs, [VMware vSphere® Distributed Service Engine™](#) helps free up CPU cycles with up to 20% savings in processing cores<sup>1</sup>. The hosts and virtual machines (VMs) can take advantage of these free resources to serve compute-intensive applications. By offloading network and security functions to the DPU, customers can attain line rate performance without any core CPU overhead.

Recent benchmark tests were performed by VMware and NVIDIA with network acceleration on NSX using overlay networking with an L4 distributed firewall on 10,000 NVIDIA BlueField-2 DPU-enabled hosts. These tests showcased over 30% improvement in throughput with a 25% reduction in transaction latency when compared to standard NICs without any DPU offload or acceleration<sup>1</sup>. This allows businesses to achieve greater efficiency and productivity.

It is worth noting that the benchmark results also demonstrated that DPU offload with full acceleration mode in a Redis-on-vSphere deployment contributes to increased energy efficiency in data centers. This results in a potential reduction, of up to 18%, in the number of servers required for a workload of up to 80 Redis instances per ESX host<sup>1</sup>. Consequently, this leads to increased CapEx savings due to the reduced server requirement, as well as OpEx savings resulting from decreased electricity consumption, cooling, and power distribution costs. For more information on additional savings, refer to white paper “[DPU Power Efficiency](#)” published by NVIDIA, which covers the improvements in the Total Cost of Ownership (TCO) for the DPU solution.

<sup>1</sup>NVIDIA Developer Technical Blog, “[Accelerating Redis Performance Using VMware vSphere 8 and NVIDIA BlueField DPUs](#)”, May 2023. Actual results may vary.

## Explore the value of co-engineering with Lenovo, NVIDIA, and VMware

Lenovo ThinkAgile VX Series, in conjunction with NVIDIA BlueField-2 DPUs and vSphere Distributed Service Engine, enables next-generation business data centers with better infrastructure services and management focused on improved efficiency, performance, security, and total cost of ownership (TCO).

The solution is available to customers as a fully integrated product, reducing operational complexity and optimizing IT infrastructure performance.

The NVIDIA BlueField-2 Data Processing Unit (DPU) leverages existing tools and user experiences with VMware vCenter® to help integrate lifecycle management into current workflows. Lenovo XClarity Integrator automates non-disruptive DPU upgrades and patches with comprehensive lifecycle management and native integration with vCenter.

Lenovo delivers a single point of contact, simplifying support and enabling organizations to keep up with the demands of modern workloads and applications.

## BlueField-2 DPU availability on ThinkAgile VX Series

The DPU solution with Lenovo ThinkAgile VX Series features the [NVIDIA Bluefield-2 Data Processing Unit \(DPU\)](#) with dual ports of 25GbE, an array of Arm cores, purpose-built hardware-acceleration engines, and full software programmability to leverage DPU hardware accelerators, providing breakthrough data center performance, efficiency and security for software-defined storage, networking, and management of workloads.

The feature is factory-installed, requires the vSphere® 8 or later release, and is available in the following Lenovo offerings:

- ThinkAgile VX650 V3 DPU Integrated System
- ThinkAgile VX650 V3 DPU Certified Node
- ThinkAgile VX650 V3 DPU SAP HANA Certified Node
- ThinkAgile VX650 V2 DPU Certified Node

Lenovo ThinkAgile VX Series with vSphere on NVIDIA BlueField-2 DPUs power a wide range of HCI use cases with its flexibility to scale up or out for hybrid cloud environments and deployments at the edge. Some major applications include AI/ML modeling and inferencing, manufacturing and automation, business-critical applications, and large analytical databases, including SAP HANA.

## Experience vSphere on NVIDIA BlueField DPUs

[NVIDIA LaunchPad](#) offers a unique and immersive hands-on experience, giving you exclusive access to vSphere with NVIDIA BlueField-2 DPUs running on Lenovo hardware. With ready-to-use infrastructure and expert assistance, you can prototype and test next-generation applications and workloads over secured accelerated infrastructure, ensuring consistency across the data center, edge, and telco cloud. Take curated labs, learn from self-paced experiences, and access assistance from NVIDIA experts to confidently make design and purchase decisions, accelerating your journey.

## How Lenovo ThinkAgile VX Can Help

[Lenovo ThinkAgile VX Series](#) is a turnkey hyperconverged infrastructure (HCI) solution, tested and validated for VMware vSAN compliance. Powered by the most reliable and secure server platforms and VMware's virtualization software ecosystem, it features next-generation technology to future-proof infrastructures and enable deep integration across the stack. It helps organizations keep up with workload capacity and performance needs to drive revenue growth and efficiencies across the infrastructure.

ThinkAgile VX provides virtualized compute and storage capabilities to run a variety of workloads and applications powered by industry-leading VMware software such as [VMware vSAN](#), [vSphere](#), and [VMware Cloud Foundation™](#). ThinkAgile VX is integrated with [VMware vSphere Lifecycle Manager™](#) for a single, unified platform for firmware and software updates. The ThinkAgile VX Integrated System is preinstalled with the VX Deployer software to simplify deployment and accelerate time to value while deploying new clusters or expanding existing clusters with additional nodes. In addition, it comes paired with ThinkAgile Premier support, providing a single source of L1/L2 support for both hardware and software, which enables faster recovery.

Customers can also leverage [Lenovo TruScale Hybrid Cloud with VMware](#), an infrastructure as a service (IaaS) offering, that gives enterprises access to the resources they need, when they need them. It provides the flexibility of a scalable pay as you go model that frees up capital for other business initiatives, and comprehensive management and support to accelerate their hybrid multi-cloud transformation.

## For More Information

For more information, see these pages:

- ThinkAgile VX Series home page:  
<https://www.lenovo.com/us/en/servers-storage/sdi/thinkagile-vx-series/>
- VMware solutions on Lenovo servers:  
<https://www.lenovo.com/vmware>
- vSphere on DPUs (vSphere Distributed Services Engine)  
<https://www.vmware.com/products/vsphere/distributed-services-engine.html>
- vSphere Distributed Services Engine Hands-on-lab  
<https://customerconnect.vmware.com/en/evalcenter?p=vsphere-hol-dse-ov-23>

## Authors

Brian Faleiro is the Worldwide Technical Product Manager for Lenovo's ThinkAgile VX Series of Hyperconverged Infrastructure (HCI) solutions based on VMware's vSAN software defined storage (SDS). Brian is responsible for showcasing the business value and differentiation of Lenovo's hybrid cloud solutions and contributing to the product lifecycle process.

Shobhit Bhutani is a Principal Product Marketing Manager in VMware and is responsible for vSphere, AI/ML and DPU initiatives. Shobhit has 20+ years of leadership experience in manufacturing, technical support, and marketing in Dell EMC, Pure Storage and Rackspace. Shobhit's passions are technology solutions & product marketing. Shobhit holds a Bachelor's in Mechanical Engineering and a Master's in Industrial Engineering.

Motti Beck is the Senior Director for Enterprise Market Development at NVIDIA Networking. Prior to that, he was a VP and General Manager at DuPont Photomask, and a founder and CEO of BindKey Technologies, a company that provided deep submicron semiconductors verification solutions and was acquired by DuPont Photomask. Motti holds Bachelor of Science in Computer Engineering from Technion - Israel Institute of Technology.

## Related product families

Product families related to this document are the following:

- [ThinkAgile VX Series for VMware](#)
- [VMware vSphere](#)

## Notices

Lenovo may not offer the products, services, or features discussed in this document in all countries. Consult your local Lenovo representative for information on the products and services currently available in your area. Any reference to a Lenovo product, program, or service is not intended to state or imply that only that Lenovo product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any Lenovo intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any other product, program, or service. Lenovo may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

Lenovo (United States), Inc.  
8001 Development Drive  
Morrisville, NC 27560  
U.S.A.  
Attention: Lenovo Director of Licensing

LENOVO PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary. Lenovo may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk. Any performance data contained herein was determined in a controlled environment. Therefore, the result obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

© Copyright Lenovo 2024. All rights reserved.

This document, LP1733, was created or updated on March 26, 2024.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at:  
<https://lenovopress.lenovo.com/LP1733>
- Send your comments in an e-mail to:  
[comments@lenovopress.com](mailto:comments@lenovopress.com)

This document is available online at <https://lenovopress.lenovo.com/LP1733>.

## Trademarks

Lenovo and the Lenovo logo are trademarks or registered trademarks of Lenovo in the United States, other countries, or both. A current list of Lenovo trademarks is available on the Web at <https://www.lenovo.com/us/en/legal/copytrade/>.

The following terms are trademarks of Lenovo in the United States, other countries, or both:

Lenovo®

ThinkAgile®

XClarity®

Other company, product, or service names may be trademarks or service marks of others.