Performance Meets Redundancy: The Dual DPU Revolution on Lenovo ThinkAgile VX Series

Article

Data Processing Units (DPUs) are rapidly becoming vital in modern data centers due to their ability to offload workloads from CPUs, enhancing efficiency and performance. They boost processing power by handling network functions, allowing CPUs to focus on application processing. DPUs also improve performance with hardware accelerators, reduce latency, and handle complex, data-intensive tasks like AI and deep learning.

Last year, we showcased how enterprises can benefit from a fully integrated solution with a single NVIDIA BlueField-2 DPU on Lenovo ThinkAgile VX Series. Today, we are happy to announce the first-ever fully integrated solution with dual DPU support for VMware Cloud Foundation, on Lenovo ThinkAgile VX Series.

Eliminating Single Points of Failure

Customers are on the lookout for technology trends such as faster DPU port speeds and dual-active DPUs. To achieve faster business benefits, enterprises are seeking to streamline development, testing and deployment of generative AI applications.

As the DPUs are intended to work as “smart NICs”, the need for failover redundancy and high availability is a desired feature that drives adoption of DPUs among customers. Additionally, as performance optimization continues to be a priority, having an additional DPU working independently offers increased offload capacity for additional throughput for workloads.

With Lenovo ThinkAgile VX Series, customers can now configure up to 2x NVIDIA BlueField-2 DPUs, that operate independently and seamlessly within a single host.

Infrastructure Reliability with Accelerated Workload Performance

The initial release of the single DPU support on Lenovo ThinkAgile VX Series introduced an innovative approach to accelerate and optimize IT infrastructure performance while embracing a zero-trust security with distributed firewalls offloaded to the NVIDIA BlueField-2 DPU.

In addition to these key features, the introduction of dual DPU support with NVIDIA BlueField-2 DPUs on Lenovo ThinkAgile VX Series provides:

- **High Availability and Redundancy -** Customers can configure the DPUs in an Active-Standby state to enhance redundancy. In this mode, if the active DPU fails, all traffic seamlessly fails over to the standby DPU, ensuring that there is no disruption in service. This configuration is perfect for mission-critical applications where uptime is paramount.
- **Enhanced Performance -** Both DPUs operate independently, providing maximum throughput. This configuration is ideal for environments where performance is critical, as it takes full advantage of the increased bandwidth with twice the capacity.

These are shown in the following figure.

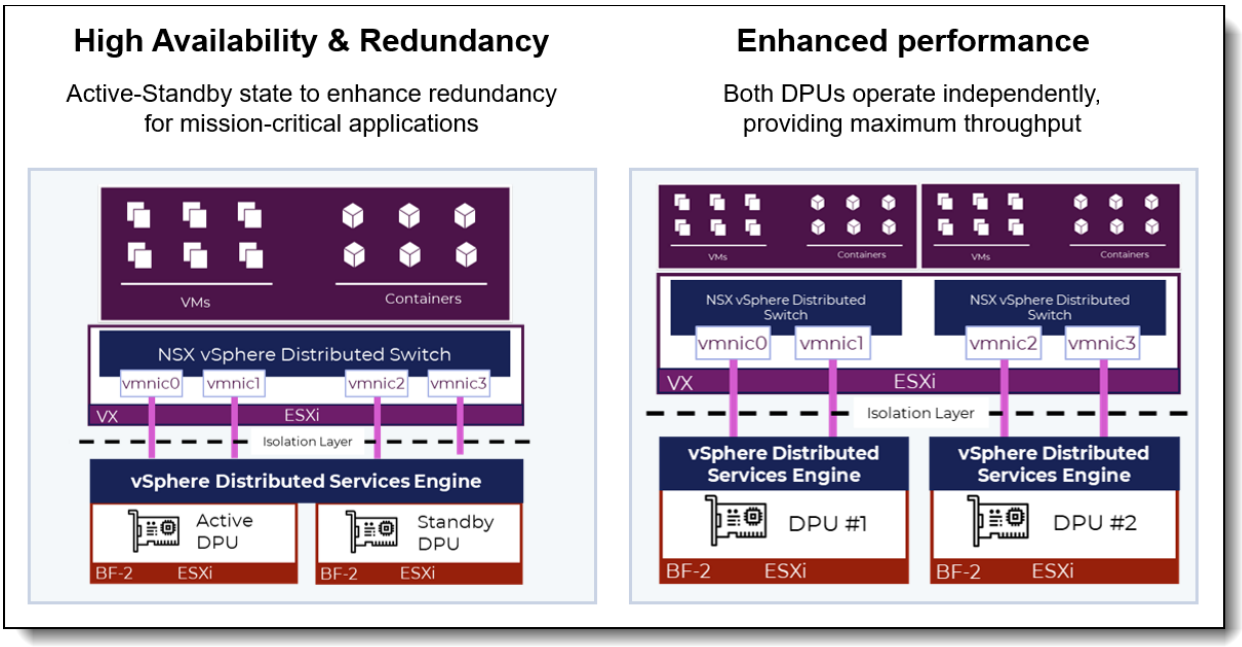

Figure 1. Dual DPU support on ThinkAgile VX Series

Industry-First Synchronous Lifecycle Management of Dual DPUs with ThinkAgile VX

As the configuration involves the operation of three ESXi instances on a single server (primary host instance and instances on each of the two DPUs), it is critical to ensure that all the ESXi instances are remediated and kept up to date with the same version.

Together with VMware and NVIDIA, Lenovo introduces the industry-first synchronous lifecycle updates of dual DPUs with VMware vSphere Lifecycle Manager (vLCM). This simplified, integrated support ensures that both DPU ESXi versions are remediated and kept up to date with the same version as the host ESXi instance, thereby simplifying management and ensuring consistency across your infrastructure.

Dual BlueField-2 DPU Availability on ThinkAgile VX Series

The dual 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 vSphere® 8.0 U3 or later, and is available on the Lenovo ThinkAgile VX650 V3-DPU model.

This solution caters to a wide range of HCI use cases in hybrid cloud environments, edge deployments and major applications include AI/ML modeling and inferencing, manufacturing and automation, business-critical applications, and large analytical databases.

Interested in Learning More?

Check out this VMware vSphere Breakroom Chat session featuring Neeraj Kuppam, General Manager Software Defined Infrastructure at Lenovo, with VMware and NVIDIA, where you will learn about the use cases and business requirements driving the adoption of dual DPUs and what the three organizations are doing together, to enable our customers meet these industry demands. Please contact your Lenovo Sales Representative, for more details.

Authors

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

Catherine Maina is the Worldwide Senior Product Manager for Lenovo's ThinkAgile VX Series, a hyper-converged engineered solution powered by VMware's virtualization software ecosystem. Catherine is responsible for identifying and commercializing opportunities that leverage joint Lenovo and VMware technology in the Software Defined Infrastructure (SDI) space.

Related product families

Product families related to this document are the following:

- ThinkAgile VX Series for VMware
- VMware Alliance

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.

This document, LP2008, was created or updated on August 23, 2024.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at:
  https://lenovopress.lenovo.com/LP2008
- Send your comments in an e-mail to:
  comments@lenovopress.com

This document is available online at https://lenovopress.lenovo.com/LP2008.

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®

The following terms are trademarks of other companies:

Revolution® is a trademark of Microsoft Corporation in the United States, other countries, or both.

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