Scalable VMware vSAN Storage Architectures on Lenovo ThinkAgile VX

Solution Brief

Business Trends

Software Defined Datacenter Architecture: The growth of technology, modern applications, cloud, artificial intelligence and data-driven ecosystems bring the need for on-premises infrastructure to meet high performance and high-density workload requirements across datacenters. Servers need to have software defined solutions to build robust storage systems and integrated software stack for cloud and cloud native technologies and third-party software to build flexible and scalable architecture to meet any workload. This trend means businesses need faster servers and storage to achieve the performance and low latency required.

Infrastructure and Workloads Modernization: Adoption of hybrid cloud, artificial intelligence and machine learning and open source technologies drives enablement for tens of software and applications to coexist on the shared infrastructure and need increased processing and data storage capabilities. Ever growing data and AI/ML workloads require storage system with low latency and high capacity drives, high speed network adapters and interconnects that are pre-tested and pre-validated engineered solution to reduce deployment complexities. Business needs for consolidation enterprise workloads and rapid provision end-end infrastructure and software stack without compromising performance and integration capabilities.

Lenovo ThinkAgile VX Solutions for VMware vSAN

Lenovo ThinkAgile VX systems are the perfect choice for hyperconverged infrastructure and provide an outstanding platform to support the different VMware vSAN™ architectures. Lenovo and VMware's over 20 year partnership and collaboration continues to strongly drive innovation and technical enablement for vSAN-based storage solutions. This includes validation, certification, configuration and support for ThinkAgile VX systems.

Lenovo ThinkAgile VX V3 hyperconverged systems are equipped with 4th generation Intel® Xeon® Scalable processors and VMware vSAN 8. They are **Accelerated by Intel** offerings that drive greater performance for CPU and IO intensive workloads.

Lenovo ThinkAgile VX servers are available as Integrated Systems and Certified Nodes. Both are factory integrated, pre-configured systems with Lenovo hardware, VMware software, and deployment services. VX Integrated Systems provide a quick and convenient path to implement a hyperconverged solution powered by VMware vSAN and a single point of contact provided by Lenovo for purchasing, deploying, and supporting the solution. VX Certified Nodes come with optional VMware software and services.

ThinkAgile VX Integrated Systems can also be up and running quickly with a web-based deployment wizard. The installer can install and configure VMware ESXi, vCenter Server and Lenovo XClarity Integrator and either create or join a cluster.

Lenovo ThinkAgile VX systems support all vSAN architectures:

Table 1. VMware vSAN Storage Architectures

Original Storage Architecture (OSA)	• Tiered architecture with flash cache tier and capacity tier • Support NVMe, SSD and HDD drives • Minimum 2 drives and 10GbE • Disk group uses one cache and multiple capacity drives
Express Storage Architecture (ESA)	• Single tier storage pool architecture with all NVMe drives • Minimum 1 drive and 10 GbE • Minimum 3 nodes
vSAN Max Architecture	• Disaggregated storage based on Express Storage Architecture (ESA) • Compute nodes are not part of the cluster • Minimum 6 nodes and 100 GbE is recommended • Ready Node configurations available

Comparison of vSAN OSA and ESA Solutions

The table below shows a basic comparison of vSAN OSA and ESA solutions.

Table 2. Comparison of VMware vSAN OSA and ESA

Features	OSA	ESA
Disk Type Support	Cache tier - NVMe, SSD Capacity tier - NVMe, SSD, HDD	All NVMe drives (Minimum 1 drive per host)
Network (Minimum)	10 GbE	10 GbE
Compression Encryption	Datastore level	Policy driven virtual machine level enabled by default
Deduplication	Supported (Full Clone VDI)	Not supported
Data transfer over network	Not compressed	Compressed
vSAN File Services	Supported	Supported (vSphere 8.0 U2)
Default Storage Policy	RAID 1	Auto Policy Management RAID 5 for small clusters and RAID 6 for larger clusters
Maximum VMs per host	200	500 (vSphere ESXi 8.0 U2)

ThinkAgile VX and vSAN Storage Architectures

ThinkAgile VX650 V3

ThinkAgile VX630 V3

Applications where the servers would excel include:

- Mission Critical applications
- SAP HANA and ERP applications
- CRM
- Business Intelligence (BI)
- Databases and Data Warehouse
- Analytics
- Virtual Desktops
- Artificial Intelligence
- Server Consolidation
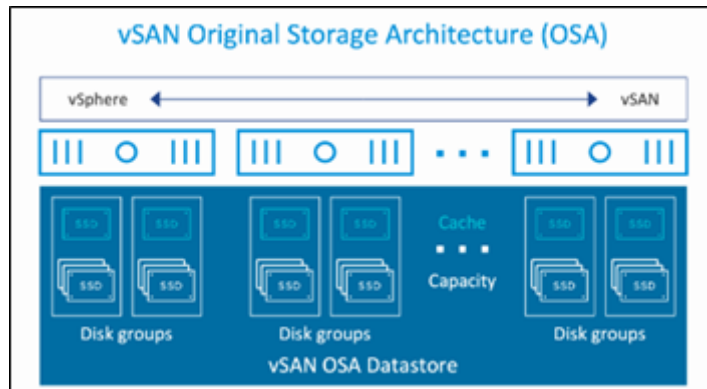- Virtualization

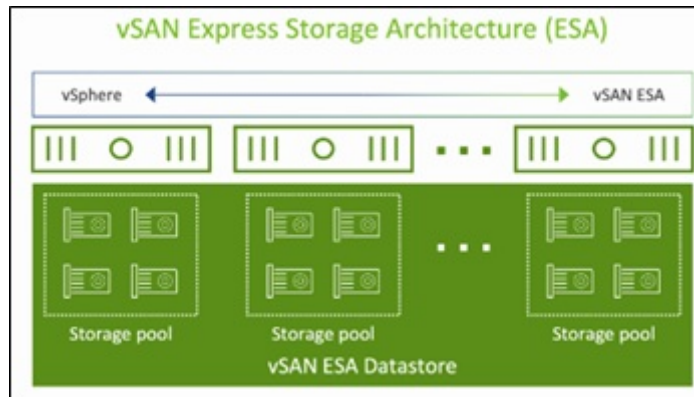Figure 1. vSAN Original Storage Architecture (OSA)

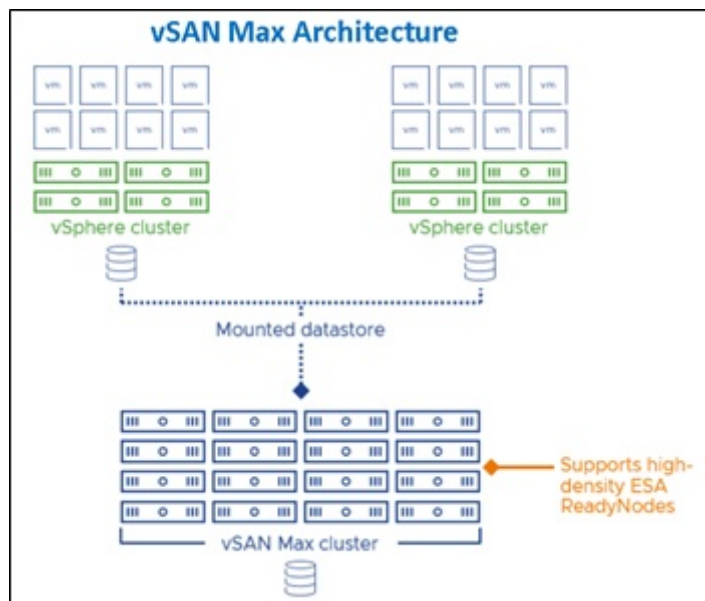Figure 2. vSAN Express Storage Architecture (ESA)

Figure 3. vSAN Max Architecture

The following sections provide benchmark results for Login Enterprise (VDI) and HCIBench (FIO) benchmark done on ESA and OSA configurations.

VMware Horizon VDI Performance on vSAN ESA vs. OSA

The Login Enterprise VDI benchmark was performed for the "knowledge worker profile" on a 4 node ThinkAgile VX650 V3 Integrated System cluster with 2x Intel® Xeon® Platinum 8468 48C processors for both vSAN OSA and vSAN ESA with 1500 users. The test was done with the following software versions and configuration:

- LoginEnterprise 5.1.2
- VMware ESXi, 8.0.1-21495797, Instant Clone
- VMware Horizon 8.9.0 build 21593375 v2303
- Microsoft Windows 10 Enterprise Edition
- Microsoft Office Professional 2007, Microsoft Edge
- 3 VCPU + 2 GB Memory + 40 GB disk

The knowledge worker profile generates 10-15 IOPS per user during steady state and is not a storage intensive benchmark.

OSA	ESA
2 disk groups Cache tier - 2 x ThinkSystem P5620 1.6TB Mixed Use NVMe PCIe 4.0 Capacity tier - 4 x ThinkSystem S4520 1.92TB Read Intensive SATA 6Gb HS SSD drives	8 x ThinkSystem P5620 3.2TB Mixed Use NVMe PCIe 4.0 drives
RAID 1	RAID 5
25 GbE	25 GbE
RDMA Enabled	RDMA Enabled
VSIMAX = 1428	VSIMAX = 1395

The Login Enterprise VSIMax provides desktop density at 100% CPU utilization. The performance of knowledge worker profile looks similar for both ESA and OSA architectures and ESA provides slightly higher density. The office applications response time and latency with ESA is better than OSA tests. The spike in response time at 190th minute is due to saturation of CPU and new users are trying to login and launch applications. It is recommended to design and size CPU capacity at no more than 70-80% to leave headroom to accommodate peak load and failover scenarios.
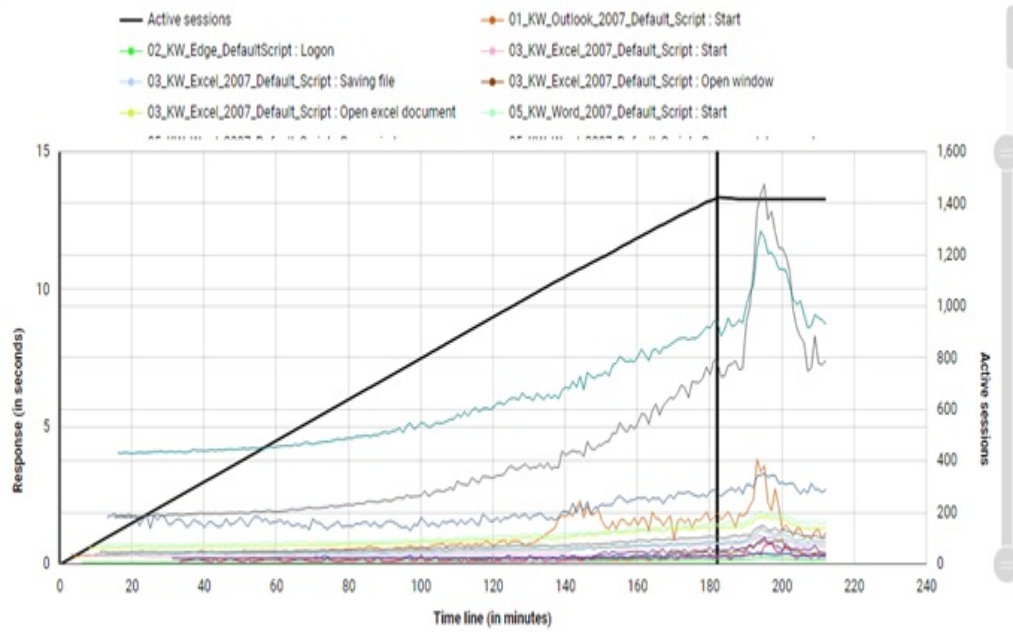
Figure 4. Login Enterprise Test Results with vSAN ESA and OSA

HCIBench FIO Benchmark on vSAN OSA and ESA

The HCIBench tool (version 2.8.2 and FIO 3.3) was used for measuring performance for OSA and ESA. The constant load scenario objective is to measure actual I/O latency on the storage during normal I/O operations and it does not let system reach maximum IOPS which can result in increase in latency.

VM Configuration	Constant Latency Scenario
Warm up before testing	5 mins
VCPU	4
Memory	8 GB
No. of data disks	8
Disk size	50 GB
IO rate	100%
FIO settings	ioengine=libaio iodepth= (1-32, refer next table)
Total number of VMs	16 (4 per node)

Table below shows different IO scenarios tested with HCIBench and each scenario matches with one or more real world use cases.

Scenario	Block size	Iodepth	Read %	Write %	Random %	Workload Category
4K-Read 100%	4K	32	100%	0%	100%	Read-intensive workload
8K-Read 70%	8K	16	70%	30%	100%	NoSQL, Dey Value, Generic web workloads
16K-Read 70%	16K	8	70%	30%	70%	Files and relational databases
32K-Read 50%	32K	4	50%	50%	50%	SQL
128K-Write 100%	128K	1	0%	100%	0%	Write-intensive workload

Constant Latency Testing with 4th Gen Intel Xeon Processors

Lenovo performed HCIBench constant latency scenario with fixed IO depth value to measure OSA and ESA throughput on 25 GbE and 100 GbE networks. The benchmark was done on 4x Lenovo ThinkAgile VX650 V3 servers, each with Intel Xeon Platinum 8468 48C processors.

OSA	ESA
3 disk groups Cache tier - 3 x ThinkSystem P5620 1.6TB Mixed Use NVMe PCIe 4.0 x 4 HS SSD Capacity tier - 12 x ThinkSystem PM1653 1.92TB Read Intensive SAS 24Gb HS SSD	8 x ThinkSystem P5620 3.2TB Mixed Use NVMe PCIe 4.0 x 4 HS SSD
2 x ThinkSystem Mellanox ConnectX-6 Lx 10/25 GbE SFP28 2-port 2 x ThinkSystem Mellanox ConnectX-6 Dx 100 GbE QSFP56 2-port PCIe ethernet adapters	2 x ThinkSystem Mellanox ConnectX-6 Lx 10/25 GbE SFP28 2-port 2 x ThinkSystem Mellanox ConnectX-6 Dx 100 GbE QSFP56 2-port PCIe ethernet adapters
RAID 1	RAID 5
25 GbE, 100 GbE	25 GbE, 100 GbE
RDMA Enabled	RDMA Enabled
VMware ESXi, 8.0.1-21495797	VMware ESXi, 8.0.1-21495797

OSA vs ESA performance on 25 GbE and 100 GbE NIC

- vSAN ESA leverages 100 GbE network and provides up to 250% better performance than 25 GbE for many of the scenarios except 100% read scenarios. ESA can provide more throughput by increasing IO depth and removing constant latency factor. Refer to results for HCIBench maximum throughput scenario here.

- vSAN OSA performance on both 25 GbE and 100 GbE are similar for the scenarios 8K- Read70% and 128K-Write100%. The scenarios 16K-Read70% and 32K-Read50% provide 20-50% more throughput with 100 GbE. vSAN OSA with 3 disk groups perform better than ESA with 25 GbE network.
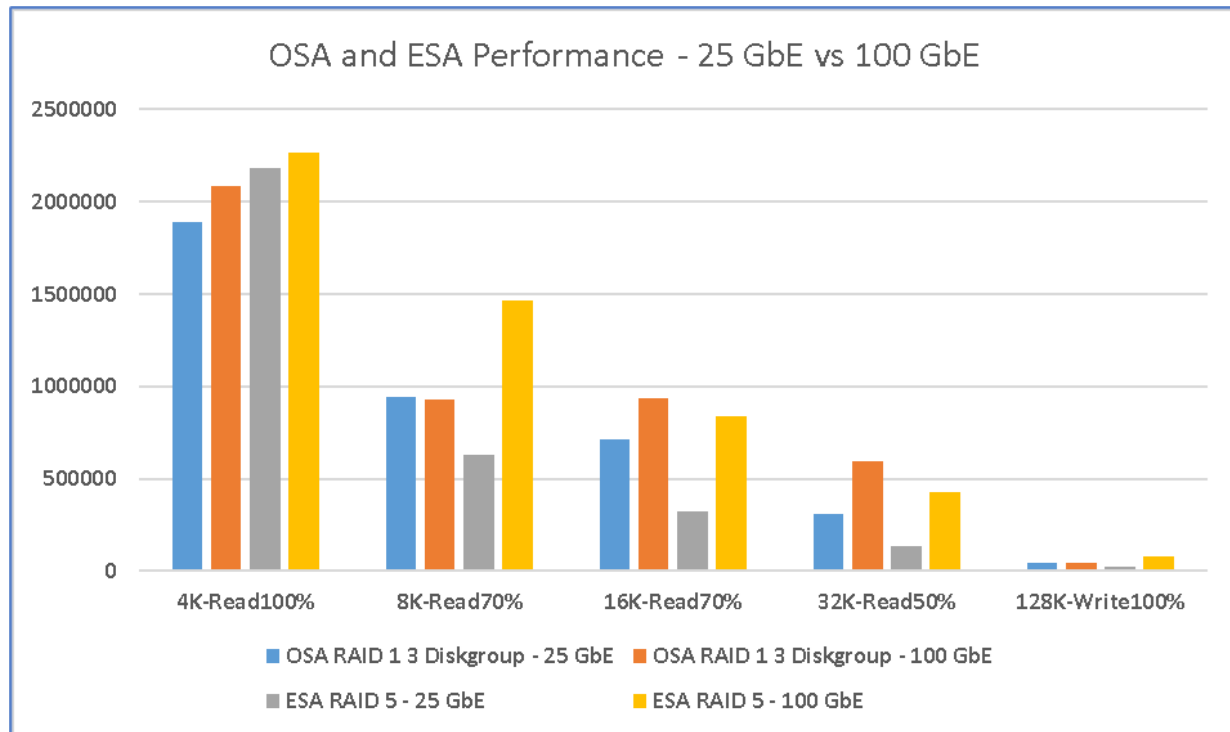

Figure 5. HCIBench Constant Latency Load Results – OSA RAID 1 vs ESA RAID 5 – 25 GbE vs 100 GbE

vSAN OSA Disk Group Scalability

Figure below shows 25 GbE throughput results for different number of disk groups for vSAN OSA. The throughput increases by 10-30% range for a greater number of disk groups while the write-only scenarios do not show considerable benefits. Adding more disk groups does require more cache drives and capacity drives which increases TCO.
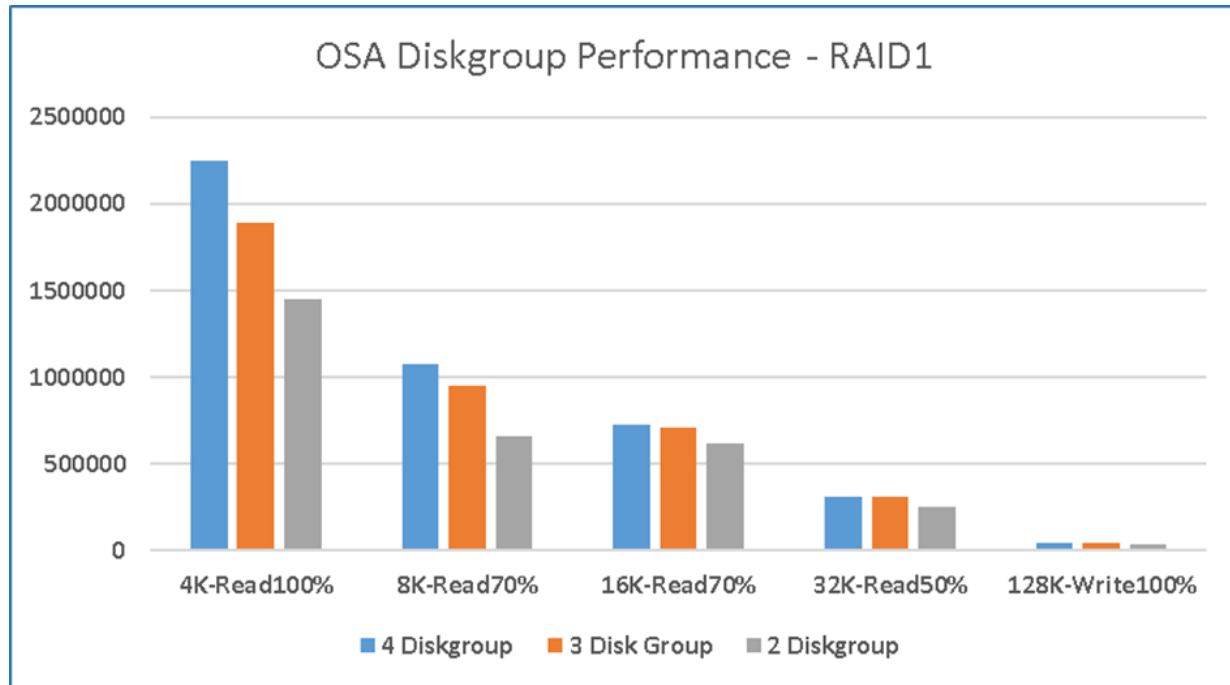

Figure 6. OSA Disk Group Scalability on 25 GbE and RAID 1

Design Guidelines for OSA and ESA

Both vSAN OSA and ESA are scalable solutions and ESA reduces storage footprint up to 40% and also it can accommodate 60% more virtual machines. vSAN OSA supports vSphere 7 & 8 and provides support for a wide number of options with NVMe, SSD and HDD and the cache tier provides improved performance for many workloads. The vSAN cluster configuration should be balanced for CPU, memory and storage requirements based on the workloads running on them and the cluster must be either scaled up or scaled out appropriately. These aspects need to be considered during initial assessment. The table below provides a comparison between vSAN OSA and ESA features and supported configurations.

Feature	OSA	ESA
vSphere Support	ESXi 7.x, 8.x	ESXi 8.x
Drive Options	- More choices for Cache and Capacity drives - Flash drives can be used for Cache, Capacity and All Flash configurations	All NVMe
TCO	- More TCO as more drives required for RAID 1 and cache & capacity tiers. 10/25 GbE NIC reduces cost	- Less storage TCO and less drives required for RAID 5/6 100 GbE NIC increases cost if workloads benefit from it.
Maximum Capacity Per Node	35 drives (5 disk groups, 7 capacity drive per disk group) - RAID 1 reduces usable capacity	32 drives - RAID 5 increases usable capacity
Scale Up	- Increase Disk Groups (Cache and Capacity) - Increase Capacity disks and ensure enough cache capacity	Capacity can be increased by adding more drives without any dependencies
Scale Out	- Add nodes and disks together - If total disk capacity reached 70% threshold and if no more drives can be added, then add more hosts	Add nodes and disk drives
Disk Maintenance	- Removing Cache drive requires removal of disk group and creation of new disk group. The larger the disk group the more time is required to evacuate data and rebalance after maintenance. - Adding/removing capacity drives are easy and data migration is at the disk level - Need a free host with enough drives to migrate data from the failed node	Drives can be added/removed easily, and data migration is at the disk level
Compression	Enabled at vSAN datastore level and causes CPU overhead and performance impact	Enable/disabled at storage policy level. No overhead on performance.
Optimal Performance	RAID 1	RAID 5 and RAID 6 based on cluster size

Accelerated by Intel

To deliver the best experience possible, Lenovo and Intel have optimized this solution to leverage Intel capabilities like processor accelerators not available in other systems. Accelerated by Intel means enhanced performance to help you achieve new innovations and insight that can give your company an edge.

Lenovo and VMware

With co-located engineering organizations and a history of technical collaboration, VMware and Lenovo consistently deliver innovative joint solutions for the data center. Lenovo's leadership in reliability, customer satisfaction, and performance, combined with VMware's leadership in software and cloud services, continues to deliver innovative data center solutions and lower TCO for our joint customers.

Why Lenovo

Lenovo is a US$70 billion revenue Fortune Global 500 company serving customers in 180 markets around the world. Focused on a bold vision to deliver smarter technology for all, we are developing world-changing technologies that power (through devices and infrastructure) and empower (through solutions, services and software) millions of customers every day.

For More Information

To learn more about workload solutions on Lenovo ThinkAgile VX630 V3 and VX650 V3, contact your Lenovo Business Partner or visit: https://www.lenovo.com/us/en/servers-storage/sdi/thinkagile-vx-series/

References:

Lenovo ThinkAgile VX630 V3 1U Integrated Systems and Certified Nodes:
https://lenovopress.lenovo.com/lp1672

Lenovo ThinkAgile VX650 V3 2U Integrated Systems and Certified Nodes:
https://lenovopress.lenovo.com/lp1673

VMware vSAN Design Guide: https://core.vmware.com/resource/vmware-vsan-design-guide

VMware ESA Ready Node Hardware Guidance:
https://www.vmware.com/resources/compatibility/vsanesa_profile.php

Related product families

Product families related to this document are the following:

- ThinkAgile VX Series for VMware
- ThinkSystem SR630 V3 Server
- ThinkSystem SR650 V3 Server
- 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, LP1872, was created or updated on May 28, 2024.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at:
  https://lenovopress.lenovo.com/LP1872
- Send your comments in an e-mail to:
  comments@lenovopress.com

This document is available online at https://lenovopress.lenovo.com/LP1872.

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®
ThinkSystem®
XClarity®

The following terms are trademarks of other companies:

Intel® and Xeon® are trademarks of Intel Corporation or its subsidiaries.

Microsoft®, Microsoft Edge, and Windows® are trademarks 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.