Lenovo High Performance File System Solution with WEKA Data Platform

Solution Brief

Accelerate High-Performance and Technical Compute Workloads

Big data has created new opportunities for innovation but also unforeseen problems for data-intensive applications, particularly in artificial intelligence (AI), machine learning (ML), deep learning (DL), financial analytics, genomics, life sciences, and high-performance computing (HPC). These compute and data-intensive applications demand maximum performance even while their datasets scale to extreme capacities. Legacy storage and data management solutions were designed to solve yesterday's problems and cannot take on the storage challenges that organizations running leading-edge applications are facing now or will in the future. These demanding new workloads can place significant stress on legacy solutions, which were not originally designed to handle these new challenges.

The Lenovo High Performance File System Storage Solution with the WEKA® Data Platform, built on Lenovo ThinkSystem SR630 V2, SR635, SR650 V2, and SR655 servers, is a modern data infrastructure solution that is uniquely built to solve big problems that previously had no solution. It eliminates the complexity and compromises associated with legacy data storage and offers unmatched performance at ANY scale while still providing all the necessary enterprise features and benefits at a fraction of the cost. By meeting the stringent storage demands of key industry workloads, the Lenovo and WEKA solution accelerates the process of obtaining insight from mountains of data and enhances the promise of discovering breakthrough innovation and helping solve humanity's greatest challenges.
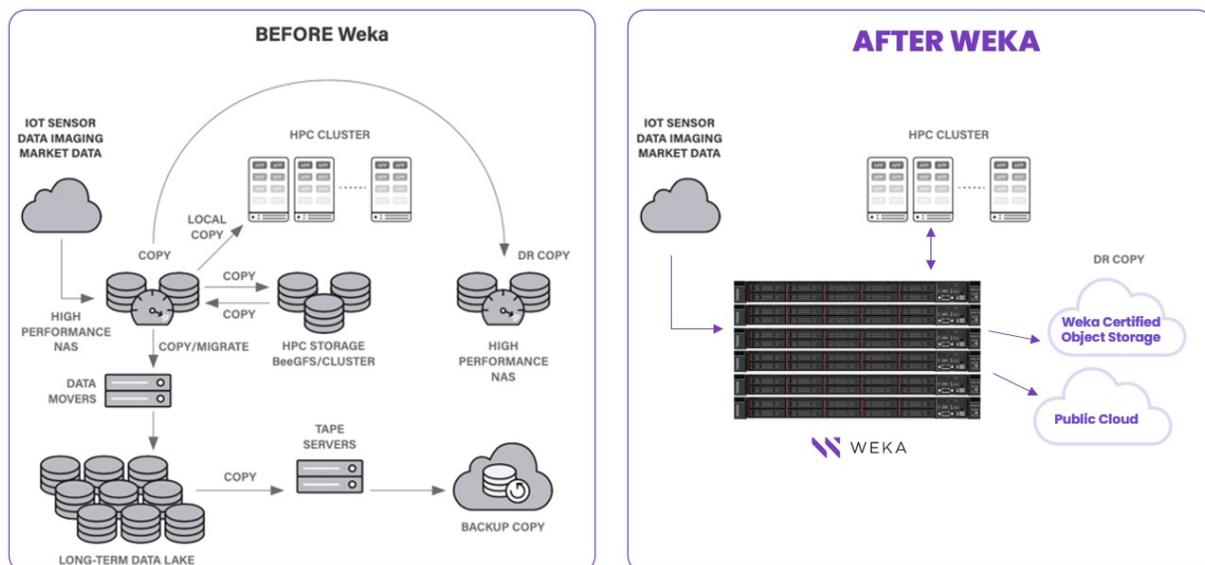

Figure 1. WEKA Unified Data Platform

The workloads that address these challenges require a modern infrastructure with extreme performance in compute, storage, and networking resources. The engineered solution from Lenovo and WEKA addresses these needs and helps you gain the insights needed to gain a competitive advantage, deliver customer value, and grow your business.

The Lenovo ThinkSystem server platforms and the WEKA Data Platform work together to maximize acceleration, reduce training times, and deliver unmatched performance and reliability at scale. This solution is ideal for all of today's data-intensive applications. This storage solution has proven scalable performance, delivering over 10x more performance than all-flash scale-out NAS and 3x more than locally attached NVMe SSDs. Together, Lenovo and WEKA support high-performance workloads in AI, finance, genomics, and technical computing. Because these workloads consist of billions of files, small and large, the storage system must deliver high throughput, low latency, and excellent metadata performance at extreme scale to keep these high-performance applications saturated with data. This is unattainable by legacy storage systems, but it is made possible with WEKA's modern NVMe-optimized, distributed, and parallel file system running on reliable and performance optimized Lenovo servers.

This software defined storage solution leverages WEKA software to support a hybrid cloud model, allowing enterprises to leverage on-demand public compute resources for cloud-bursting, remote backup, and disaster recovery.
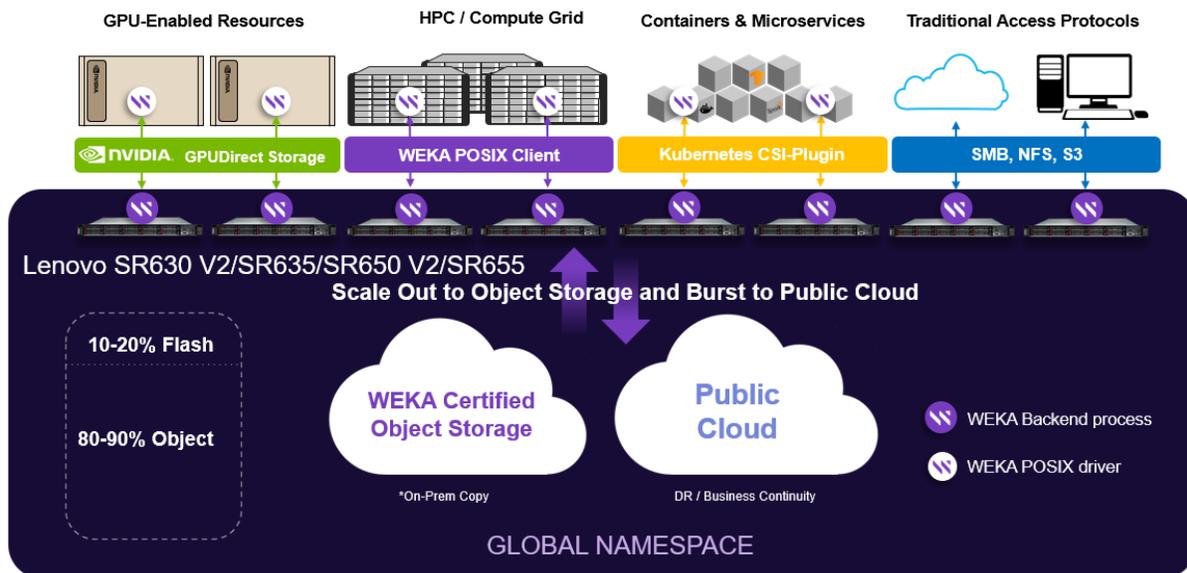

Figure 2. WEKA Data Platform for AI/ML and HPC

Choose A Solution That Delivers Performance, Flexibility, And Value

With Lenovo ThinkSystem servers and WEKA Data Platform software, you can take advantage of a plug-and-play solution that helps extract greater value from data. The solution allows you to start small and flexibly scale to meet your performance and capacity requirements.

The Lenovo ThinkSystem servers offer uncompromising performance, scalability, reliability, and density in your choice of a 1U or 2U form factors. At the core of the WEKA Data Platform is WekaFS, a POSIX-compliant file system that distributes both data and metadata evenly across the entire storage cluster to ensure massively parallel access. This enables the WEKA platform to reduce time to innovation by delivering more data to the applications that need it, faster than any other available solution. With a single namespace that offers on-premises storage and cloud connectivity, the software delivers simplified storage management and data protection. ThinkSystem adapters provide Ethernet and InfiniBand connectivity to deliver up to 200 Gb/sec bandwidth in a single network port, the lowest available latency, and application hardware offloads, satisfying even the most demanding application requirements.

Derive Benefits For Your Business

This high-performance storage solution not only addresses your specific needs based on workload types and performance and capacity requirements, but goes beyond that to provide greater overall benefits for your business:

- Optimization of your IT environments with a modern storage infrastructure and simplified storage management
- Datacenter agility with faster data access and effective resource utilization
- Data transformation for AI, analytics, and technical computing with faster time to value and insight

Lenovo recommends the following minimum entry-level building block for your initial storage deployment, giving you the ability to scale out to greater capacities and performance levels by adding more servers.

6x Lenovo ThinkSystem SR630 V2 or SR635 servers, each configured with the following:

- 256GB or 512GB of memory
- 2x Mellanox ConnectX-6 network adapters
- 4-10x PCIe 4.0 NVMe SSD drives

Lenovo and WEKA deliver a differentiated solution that goes beyond the current market and performance standards for storage and deliver the best solutions for your IT and business challenges.

Minimum Recommended Configuration

The following table lists the minimum recommended configuration for the Lenovo Storage Solution with the Weka Data Platform.

Table 1. Minimum Recommended Configuration

Specification	Recommendation
Server	Lenovo ThinkSystem SR630 V2 (Intel) or Lenovo ThinkSystem SR635 (AMD)
Number of Servers	6
Drives/Server	10
SSD Capacity/Drive	3.84 TB
Usable Capacity (Note 1)	115 TB
CPU	2x Intel Xeon Gold 6326 16C or 1x AMD EPYC 7543P 32C
Networking	High-speed Ethernet or InfiniBand
Software	WEKA
Data Protection	Distributed Data Protection (N+2 or N+4) (Note 2) Drive Virtual Hot Sparing Error Detection: End-to-End Data Protection In-Flight and At-Rest Data Encryption
Protocols	POSIX, NFS, SMB, S3, NVIDIA GPUDirect, and Kubernets CSI Plugin
Snapshots and Clones	File System Level, up to 24K Snapshots (14K Writable)
Tiering	S3 Compatible Cloud Object Store (Public or Private)
Storage Monitoring	Cloud-based Monitoring and Analytics for Application Tuning and Remote Support
System Management	ThinkSystem XClarity Controller Enterprise
Dimensions per Server	1 Rack Unit, Max Depth: 715 mm
Support Team	Software: 3-Year, Including Same Day Support Hardware: 3-Year, Next Business Day Support
Estimated Performance (Note 3)	Up to 180 GB/sec of Read Throughput up to 14 Million Read IOPS

Notes:

1. 4+2 stripe size with one configured hot spare. Configurations may vary based upon tiering customer network configuration. Please speak with your Lenovo representative for solution guidance appropriate to your environment.
2. N+4 is available only for a cluster with 14 or more servers
3. Performance estimates are for a 6-server cluster on a 200Gb network

Configuration details are subject to change without notice. Additional capacity, drive, and server configurations are available. Please call your Lenovo representative for more information.

For more information

To learn more about Lenovo Data Management solutions, please visit our Data Management solutions web page. More details on the WEKA Data Platform software can be found at weka.io. You can also contact your Lenovo sales representative or authorized channel partner.

Related product families

Product families related to this document are the following:

- ThinkSystem SR630 V2 Server
- ThinkSystem SR635 Server

Notices

This document, LP1691, was created or updated on February 7, 2023.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at:
  https://lenovopress.lenovo.com/LP1691

- Send your comments in an e-mail to:
  comments@lenovopress.com

This document is available online at https://lenovopress.lenovo.com/LP1691.

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®
ThinkSystem®
XClarity®

The following terms are trademarks of other companies:

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

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