High-Performance Storage for NVIDIA Cloud Partners

Reference Architecture

NVIDIA validated and optimized High-Performance Storage (HPS) solutions for cloud service partners featuring NVIDIA HGX™ H100 and H200 based servers.

Executive Summary

The WEKA® AI-native data platform software, combined with class-leading storage hardware, provides a ready-to-use, purpose-built environment specifically designed for artificial intelligence (AI) applications.

WEKA empowers hundreds of the world's leading enterprises and premier research organizations to overcome complex data challenges, enabling them to achieve discoveries, insights, and outcomes more rapidly. This includes 12 of the Fortune 50 companies that rely on WEKA to enhance their data-driven initiatives.

This document outlines fully validated reference architectures for NVIDIA® Cloud Partners (NCPs). The solutions tested for the WEKA Data Platform's NCP Reference Architecture certification for the High-Performance Storage (HPS) tier, integrating the WEKA storage servers and WEKA Management Station (WMS) software with HGX servers.

With the completion of this NCP HPS Reference Architecture, WEKA and NVIDIA extend their deeper technical integration partnership that includes development and certification with NVIDIA Magnum IO™ GPUDirect Storage® (GDS) protocol, NVIDIA DGX BasePOD™, NVIDIA DGX SuperPOD™, NVIDIA OVX™ systems, NVIDIA networking solutions including ConnectX® NICs, NVIDIA Quantum and Spectrum switches, and NVIDIA BlueField-3 DPU/SuperNICs.

1. The WEKA Data Platform for NVIDIA Cloud Partners

HGX servers are engineered to deliver unparalleled performance for AI, machine learning (ML), and HPC workloads. These servers maximize computational efficiency and throughput, making them ideal for the most demanding applications. The integration of the WEKA Data Platform with NVIDIA HGX H100 and H200 servers elevates performance to new heights, offering NVIDIA customers an unmatched data processing and storage experience for their AI, ML, and HPC workloads.

Starting with a base configuration of eight servers, the WEKA Data Platform scales linearly and seamlessly to hundreds of servers, providing capacity, throughput, and IOPS expansion as needed while integrating effortlessly with NVIDIA Base Command™ Manager Essentials for centralized observability. The WEKA Data Platform offers the best read/write and mixed-workload performance available and supports small and large files simultaneously, with both mixed random and sequential I/O patterns. It achieves application-level 4kB I/O at sub-200 microsecond latency and tens of millions of IOPS.

The WEKA Data Platform is designed to complement the computational power of HGX servers, providing seamless, high-speed data access and storage. NVIDIA customers benefit from the WEKA Data Platform's rich enterprise feature set, including local snapshots, automated tiering, dynamic cluster rebalancing, private cloud multi-tenancy, backup, encryption, authentication, key management, user groups, quotas, and more.

This section describes the advanced features of the WEKA Data Platform for NCPs.

1.1 WEKA Data Platform Architecture

The WEKA Data Platform is purpose built to deliver high-performance file services leveraging NVMe flash in a single namespace for easy access and management.

WEKA's unique architecture is radically different from legacy storage systems, appliances, and hypervisor-based software-defined storage solutions because it not only overcomes traditional storage scaling and file sharing limitations but also allows parallel file access via POSIX, particularly using the Data Plane Development Kit (DPDK). DPDK is a set of libraries and drivers for fast packet processing. Using DPDK, WEKA can achieve high-speed data processing and low latency, making it ideal for performance-critical applications. This allows for efficient parallel file access, significantly enhancing the system's overall performance and scalability. It provides a rich enterprise feature set, including local snapshots and remote snapshots to the cloud, clones, automated tiering, cloud-bursting, dynamic cluster rebalancing, private cloud multi-tenancy, backup, encryption, authentication, key management, user groups, quotas with advisory, soft and hard parameters, and much more.

The WEKA Data Platform's patented data layout and virtual metadata servers distribute and parallelize all metadata and data across the cluster for incredibly low latency and high performance no matter the file size or number.

At the core of the WEKA Data Platform is WekaFS™, a distributed, parallel file system that eliminates the traditional block-volume layer managing underlying storage resources. WekaFS is an advanced file system designed to overcome the limitations of traditional storage solutions by offering superior performance, scalability, and data integrity. It features adaptive caching, which optimizes the use of local Linux data and metadata caches to significantly reduce latency and improve performance. WekaFS ensures full coherence and consistency across the shared storage cluster, automatically managing caching to eliminate the need for specific configurations or administrative interventions. This makes it an ideal choice for performance-critical applications, ensuring data integrity and ease of management.

The WEKA core components, including the WekaFS unified namespace and other functions such as virtual metadata servers (MDSs), execute in user space in a Linux container (LXC), effectively eliminating time-sharing and other kernel-specific dependencies. The exception is the WEKA Virtual File System (VFS) kernel driver, which provides the POSIX filesystem interface to applications. Using the kernel driver provides significantly higher performance than what can be achieved using a FUSE user-space driver, and it allows applications that require full POSIX compatibility to run on a shared storage system.

1.2 Performance Optimized Networking (DPDK)

To achieve optimum performance, the WEKA Data Platform does not use standard kernel-based TCP/IP services but instead employs a proprietary networking stack based on DPDK maps the network device in the user space, allowing the network device to be used without context switches or copying data between kernels. Bypassing the kernel stack eliminates the consumption of kernel resources for networking operations and efficiently scales across multiple hosts. It applies to both backend and client hosts and enables the WEKA system to fully use 400 GbE links.

Using DPDK provides operations with high throughput and extremely low latency. Low latency is achieved by bypassing the kernel and sending and receiving packets directly from the NIC. High throughput is achieved because multiple cores in the same host can work in parallel, eliminating common bottlenecks.

1.3 Multitenancy, Security, and QoS

A collection of capabilities provides multi tenancy for the WEKA Data Platform. Tenant isolation is both physical and logical, with each tenant being a self-contained WEKA Data Platform. Tenant creation is governed by the WEKA Kubernetes (K8s) Operator ("Operator") through composable sets of resources—these composable clusters are deployed adjacently to each other in collections of physical servers. The fundamental resource allocation method for tenants is dedicated components and physical isolation, allowing for naturally gated performance and strong security boundaries without extensive management and QoS overhead. The tenants do not share flash drives, drive memory, processor cores, or main memory.

Each tenant uses unique encryption keys and authentication credentials and is unaware of each other despite being adjacent. When encryption is enabled, all cluster data is encrypted. The work of encryption is done by clients, with the native data structures in WEKA containing wrapped key material. The KMS configuration for the cluster enables a key for each file system, all file data, and sensitive file metadata such as the file name. The only thing not encrypted is the size.

For service providers, this is a departure from the traditional storage approach to multi tenancy, leveraging hierarchical relationships and weaker process isolation with shared memory and contextual permissions by the tenant. While the Organizations feature of WEKA Data Platform also supports this model for trusted tenants and logical delegation hierarchies, dedicating resources allows for strong performance guarantees and security domain separation between tenants throughout the control and data paths, a necessary capability for service providers with untrusted tenants.

Logical isolation includes the network, as it is the only shared part. It can be changed per tenant using virtual functions to control bandwidth consumption by cluster and QoS settings by client. Using the Operator, a cluster can be expanded incrementally to add drives, cores, or a combination of both. The multi-tenancy capability scales up to 150,000 total tenants.

WEKA Data Platform deployments leverage templates with composable resource sets for achieving desired capacity and performance in a tenant. Every drive and core will produce a certain amount of performance, and the arithmetic is handled by the Operator. The mechanics are that the Operator allocates cores and memory in physical servers to Kubernetes containers, which contain LXC containers used by the WEKA Data Platform. Software inside the LXC containers forms a cluster and leverages all the previously mentioned performance mechanisms like DPDK. Next, the Operator adds drives to the tenant WEKA Data Platform.

Creation is fast, with most tenant clusters being able to perform I/O in single-digit minutes. Administrators can deploy both small and large sets of capacity with guaranteed performance as shown in Figure 1.

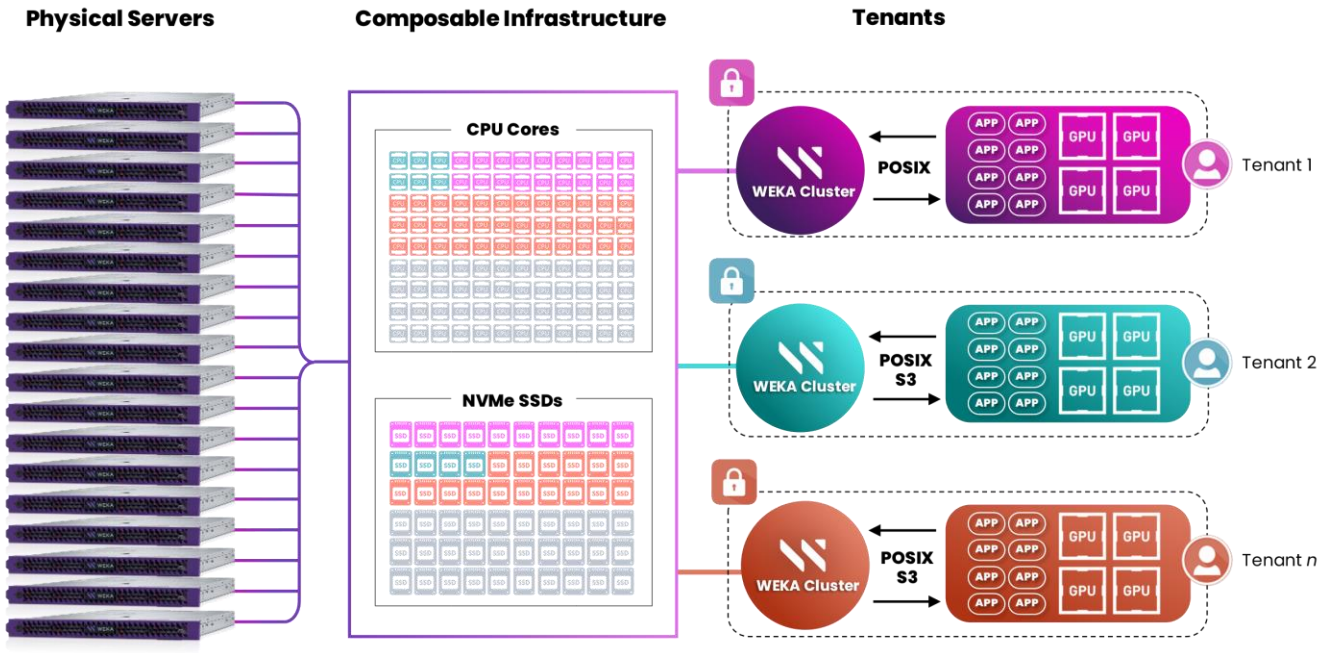

Figure 1. WEKA multi-tenant cluster architecture design

2. WEKA Data Platform

The WEKA Data Platform's AI-native architecture delivers the world's fastest AI storage and exceptionally high performance for AI data pipelines. It ensures low latency regardless of file size and provides the high levels of write throughput required by checkpointing operations to ensure business-critical continuity.

Additionally, the WEKA Data Platform offers efficient storage access for NCPs, ensuring that computational tasks, such as training and inference processes, are executed with maximum performance, reducing idle time and overall energy consumption. Performance scales linearly as more WEKA storage servers are added to a storage cluster.

During the NCP Reference Architecture validation, WEKA used and tested the WEKApod™ Appliance (Figure 2).

Figure 2. WEKApod™ Data Platform Appliance specifications.

2.1 WEKA Storage Server Qualified Configuration

If the NCP prefers another server vendor system, it must be a WEKA-qualified configuration and meet the following hardware specifications:

Table 1: WEKA storage server specifications used during NCP validation

Hardware specifications:

- AMD EPYC 9454P 48C 2.75GHz
- At least 384GB (12x32GB DIMMs, 1 DIMM per Memory Channel). DIMM type is DDR5 RDIMM 4800MT/s Dual Rank
- 2 x NVIDIA ConnectX-7 Single-Port 400Gb/s InfiniBand OSFP PCIe 5.0 x16 (requires 2x PCIe 5 x16 I/O Slots)
- 10 or more Samsung PM1743 RI E3.S TLC (3.84T, 7.68T, 15.36T SSD density), 1 or 3 DWPD. A minimum of 40 PCI-eE Gen 5 lanes to the drives

The WEKA storage server integrates WEKA's fully distributed parallel file server architecture and includes WEKA's AI-native data platform software, as described in Section 1, providing a ready-to-use, purpose-built environment for AI applications.

Each WEKA storage server runs both metadata and data services, simplifying NCP deployments. WEKA interfaces with HGX servers using DPDK, a highly efficient and low-latency packet processing technology, over ConnectX-7 interfaces. Each WEKA storage server provides two ConnectX-7 400Gb interfaces to connect to the storage fabric.

2.2 WEKA Data Platform Management GUI

WEKA provides three quick and easy ways to manage the WEKA file system, either through a Graphical User Interface (GUI), or a Command Line Interface (CLI), or REpresentational State Transfer API (REST).

Reporting, visualization, and overall system management functions are accessible using the REST API, CLI, or the intuitive GUI-driven management console (see Figure 3).

Point-and-click simplicity allows users to rapidly provision new storage, create and expand file systems within a global namespace, establish tiering policy, data protection, encryption, authentication, permissions, NFS, SMB and S3 configuration, read-only or read-write snapshots, snapshot-to-objects, and quality of service policies, as well as monitor overall system health.

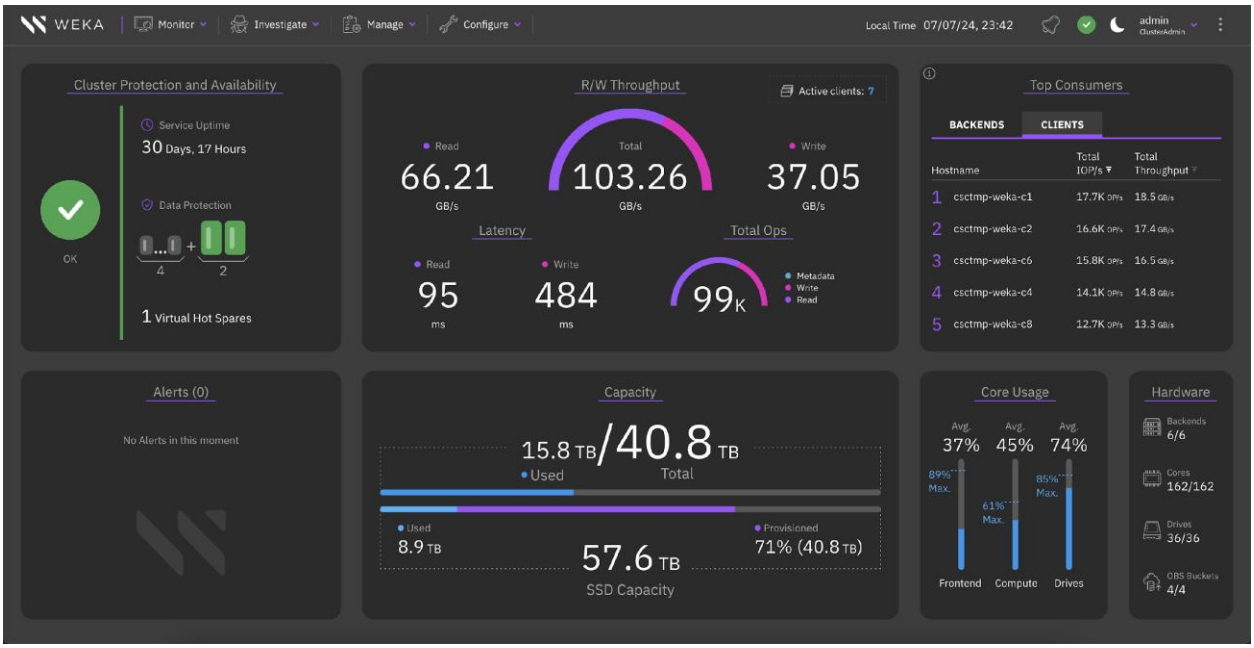

Figure 3. WEKA Management Software User Interface.

Detailed event logging provides users with the ability to view system events and status over time or drill down into event details with point-in-time precision via the time-series graphing function (Figure 4)

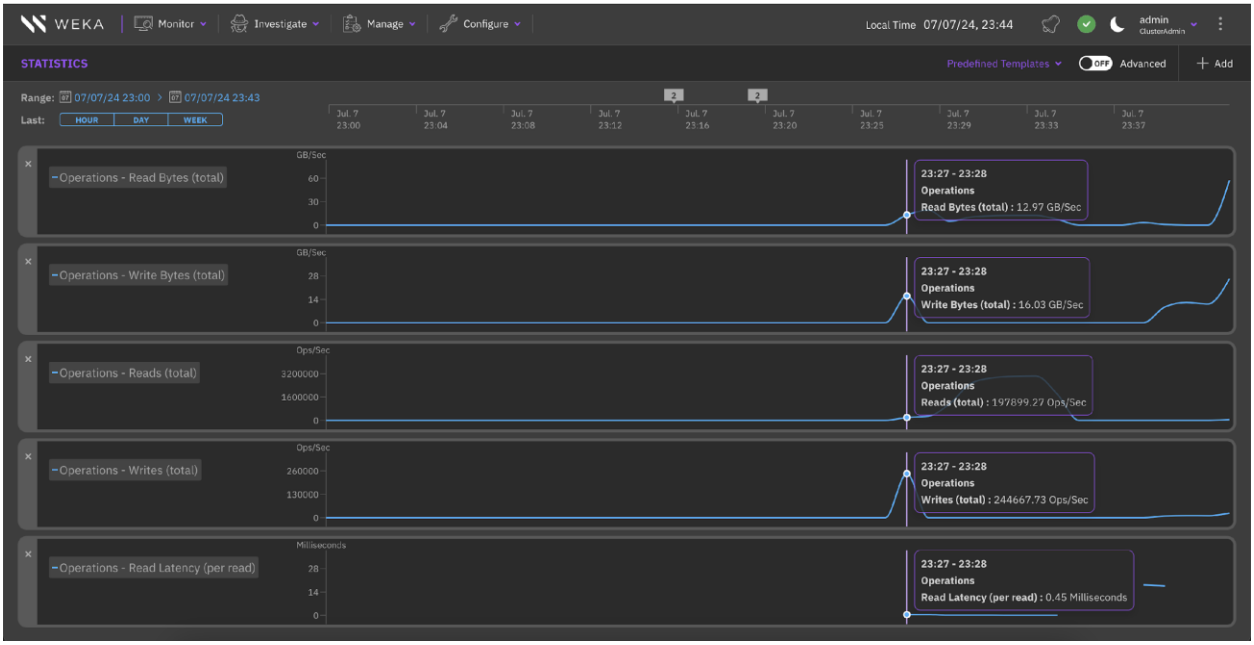

Figure 4. Time-Series Charts for Event Monitoring

3. WEKA Reference Architectures for NCP Deployments

WEKA qualified storage servers deliver all the capabilities of WEKA's Data Platform software in a predictable scale-out reference architecture that's easy to deploy. WEKA offers an exceptional, highly performant data management foundation for NCP deployments. Configured and tested with eight WEKA storage servers.

WEKA also provides efficient write performance for AI model checkpointing, delivering superior training efficiency, scalability, enterprise-grade resiliency, and support for real-time applications. WEKA storage servers provide both metadata and data services, streamlining the design, deployment, and management of a cloud environment, offering predictable performance, capacity, and scalability.

WEKA proposes the following recommended architectures for NCP deployments.

3.1 Recommended Storage Sizing

The storage performance needed to maximize training performance can vary depending on the type of model and dataset.

The guidelines in Table 2 and Table 3 provide guidance and targets to decide the I/O levels required for different types of models.

Table 2. Guidelines for storage performance requirements.

Performance Level	Workload Examples	Dataset Size
Good	Classical Natural Language Processing (NLP) workload, such as BERT	Datasets generally fit within local cache
Better	Training compressed images, compressed audio and text data, such as Large Language Model (LLM) Training	Many to most datasets can fit within the local system's cache, or datasets are distributed across different nodes
Best	Extract, transform, and load (ETL), training with large video and image files such as AV replay, generative networks such as stable diffusion, 3D images such as Medical uNet, genomics workload and protein prediction such as AlphaFold	Datasets are too large to fit into cache, massive first epoch I/O requirements, workflows that only read the dataset once, or in-place data processing

Table 3. Guidelines for HPS aggregate system performance

Compute Components			Operation	Good (GB/s)	Better (GB/s)	Best (GB/s)
GPUs	Servers	SUs				
1,024	128	4	Read	60	160	500
			Write	30	80	250
2,048	256	8	Read	120	320	1,000
			Write	60	160	500
8,192	1,024	32	Read	480	1,280	4,000
			Write	240	640	2,000
16,384	2,048	64	Read	960	2,560	8,000
			Write	480	1,280	4,000

While NLP cases often do not require as much read performance for training, peak performance for reads and writes is needed for creating and reading checkpoint files. This is a synchronous operation, and training stops during this phase. This Reference Architecture is sized for **Better**.

When looking for the best end-to-end training performance, I/O operations for checkpoints are important to the HPS design. Modern LLM workloads have a significant requirement in write performance not to consume too much time in writing checkpoints.

As a reference, Table 4 has hypothetical calculations for the required write rate for LLM training. It has these constants:

- Number of bytes per parameter: **14**
- Total write time percentage of total training time: **1%**
- Seconds per checkpointing interval: **3,600s**

Table 4. Estimates of LLM checkpoint size

Number of Parameters (in Billions)	Size (TB)	Tensor Parallel Domain Size	Pipeline Parallel Domain Size	Number of Files	Total Write Rate (GB/s)	Per Node Write Rate (GB/s)	Per GPU Write Rate (GB/s)
175	2.4	8	16	128	68.06	4.25	0.53
530	7.2	8	32	280	206.11	5.89	0.74
1,000	13.7	8	64	512	388.89	6.08	0.76

The metrics account for diverse workloads, datasets, and the necessity for local training directly from the HPS system. It is advisable to evaluate workloads and organizational needs before finalizing performance and capacity requirements.

Table 5. WEKA recommended storage sizing for NCP deployments

		SU Groups			
		4	8	32	64
Compute components	NVIDIA HGX Systems	127	255	1023	2047
	NVIDIA GPUs (Total)	1K	2K	8K	16K
WEKA Storage components	WEKA Storage Servers	8	10	32	64
WEKA Storage Specifications	Number of namespaces	1	1	1	1
	Min. aggregate capacity 3.84 TB SSD density	241 TB	311 TB	1,333 TB	2,709 TB
	Min. aggregate capacity 7.68TB SSD density	483 TB	622 TB	2,666 TB	5,419 TB
	Min. aggregate capacity 15.36 SSD density	967 TB	1,244 TB	5,333 TB	10,838 TB
	Aggregate usable inodes based on 15.36 TB SSDs	84 billion	105 billion	335 billion	670 billion
	Physical rack units	8	10	32	64
	Power, nominal	6.4 kW	8.0 kW	25.6 kW	51.2 kW
	Cooling, nominal	22 kBTU/hr	27 kBTU/hr	87 kBTU/hr	175 kBTU/hr

For configurations not listed in Table 6, additional WEKA storage servers will be determined by the aggregate read and write throughput based on each storage server's bandwidth.

3.2 Recommended Storage Networking

The NCP reference design incorporates multiple networks, including the converged in-band and storage networks that the HPS system uses. This connects WEKA storage servers to compute servers, and management servers. It is engineered to fulfill the high-throughput, low-latency, and scalability demands of NCP deployments.

Figure 5. NVIDIA Spectrum SN5600 Ethernet switch.

The NVIDIA Spectrum SN5600 Ethernet switch (Figure 5) is required for NCP storage connectivity. The cables listed in Section 3.2.2 are validated with the SN5600 to work with WEKA storage servers.

3.2.1 WEKA Storage Server Leaf Switch Connectivity

WEKA recommends a splitter cable to connect a pairing of WEKA storage servers to the converged network, see Section 3.2.3 for splitter cables reference. The splitter cable should be connected to port 1 on a WEKA storage server pairing going to switch "A." The other splitter cable must be connected to port 2 on the same WEKA storage server pairing going to switch "B" to allow redundancy, see Figure 9 and Figure 10.

The management ("M") ports should be connected to the out-of-band management network.

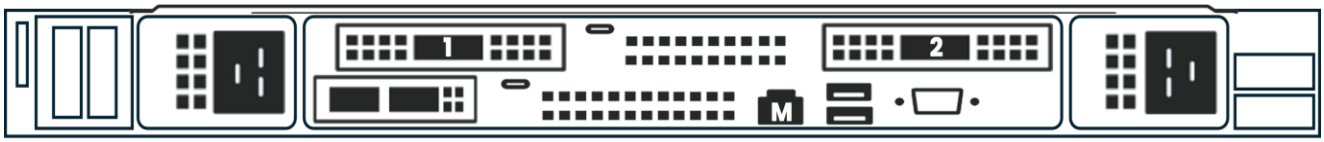

Figure 9. Recommended WEKA storage server network port connections.

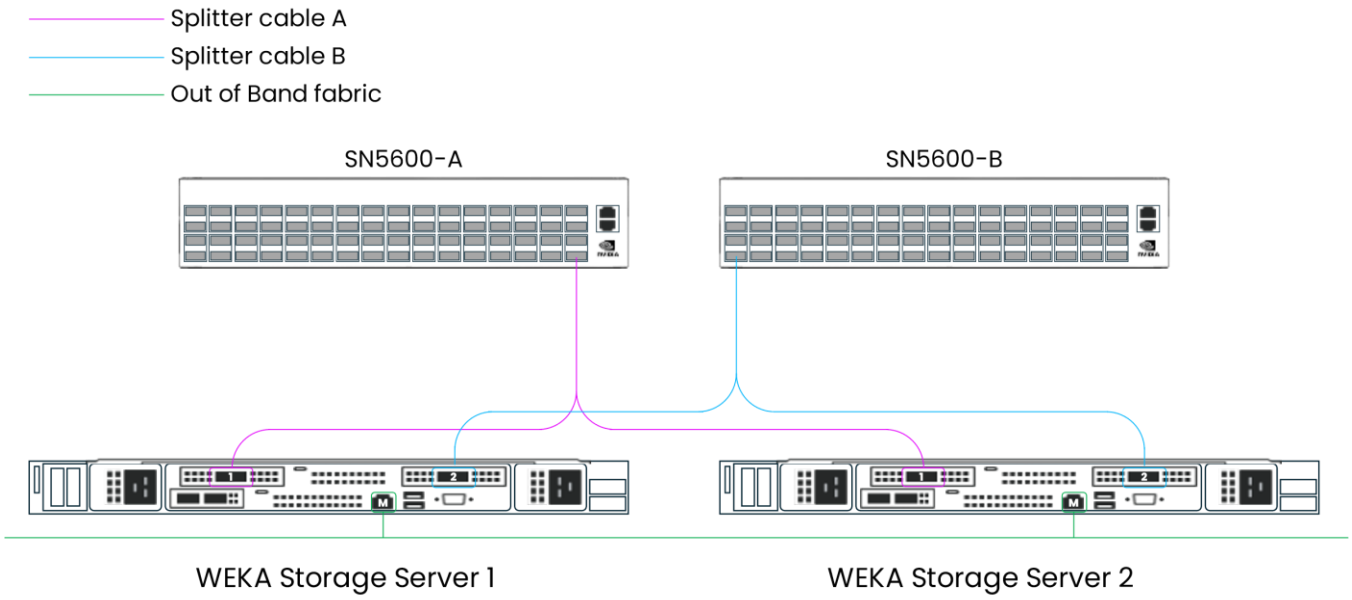

Figure 10. Recommended WEKApod network connection diagram with splitter cables.

Each WEKA storage server connects to the storage network with two 400GbE interfaces. Particular attention must be given to cabling selection to ensure compatibility between different Ethernet connectivity and data rates.

WEKA and NVIDIA have validated the cables in Table 7 to connect WEKA storage servers with SN5600 switches. Using splitter cables ensures the most efficient use of switch ports.

Table 6. DAC and ACC splitter cables

Cable Type	Part Number	Description
Direct Attach Copper	MCP7Y00-Nxxx	NVIDIA 800Gb/s Twin-port OSFP to 2x400Gb/s OSFP DAC Splitter Cable xxx indicates length in meters: 001, 01A, 002, 02A, 003
Active Copper Cable	MCA7J60-Nxxx	NVIDIA 800Gb/s Twin-port OSFP to 2x400Gb/s OSFP ACC Splitter xxx indicates the length in meters: 004, 005

3.3 NCP Deployments with 127 HGX Servers

Figure 11 illustrates WEKA's reference architecture for NCP deployments with 127 HGX servers and eight WEKA storage servers.

Every WEKA storage server connects to the converged network with two 400GbE links using the appropriate cable type.

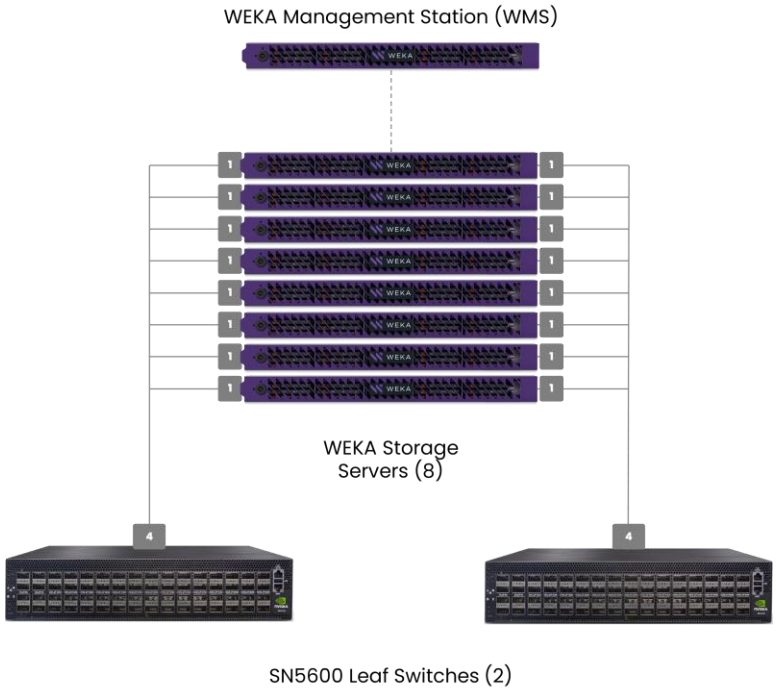

Figure 11. WEKApod reference architecture for NCP deployments with 127 HGX servers.

Description	Count
400GbE links (Storage)	16
OSFP ports (Switch) / Splitter cables (MCP700-Nxxx)	8
1GbE links (Management)	8

Table 7. Cable counts for NCP deployments with 127 HGX servers.

3.4 NCP Deployments with 255 HGX Servers

Figure 12 illustrates WEKA's reference architecture for NCP deployments with 255 HGX servers, ten WEKA storage servers. Every HGX server connects to the storage network with two 400GbE links using the appropriate cable type. Each WEKA storage server connects to the storage network with two 400GbE links using the appropriate cable type.

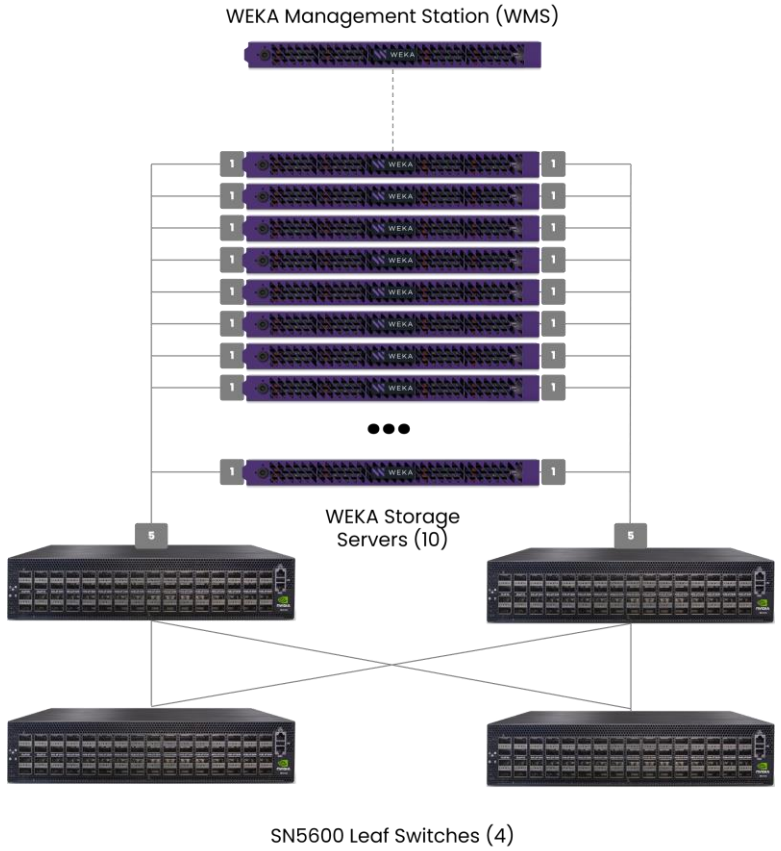

Figure 12. WEKApod reference architecture for NCP deployments with 255 HGX servers.

Description	Count
400GbE links (Storage)	20
OSFP ports (Switch) / Splitter cables (MCP700-Nxxx)	10
1GbE links (Management)	10

Table 8. Cable counts for NCP deployments with 255 HGX servers.

3.5 NCP Deployments with 1023 HGX Servers

Figure 13 illustrates WEKA's reference architecture for NCP deployments with 1,023 HGX servers and 32 WEKA storage servers. Every HGX server connects to the storage network with two 400GbE links using the appropriate cable type. Each WEKA storage server connects to the storage network with two 400GbE links using the appropriate cable type.

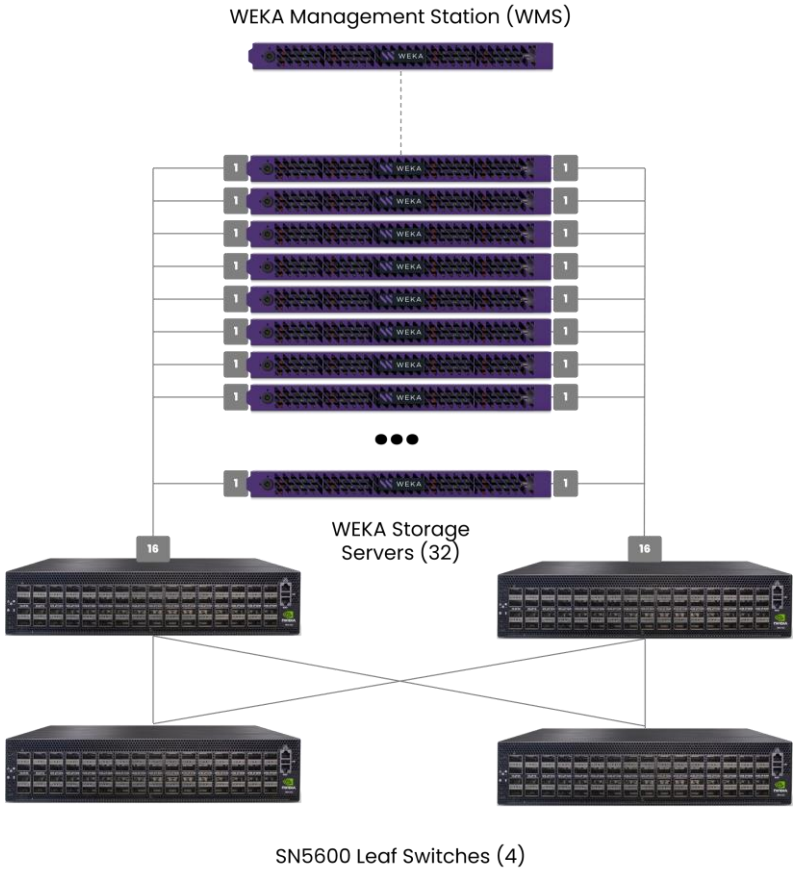

Figure 13. WEKApod reference architecture for NCP deployments with 1,023 HGX servers.

Description	Count
400GbE links (Storage)	64
OSFP ports (Switch) / Splitter cables (MCP700-Nxxx)	32
1GbE links (Management)	32

Table 9. Cable counts for NCP deployments with 1,023 HGX servers.

3.6 NCP Deployments with 2,047 HGX Servers

Figure 14 illustrates WEKA's reference architecture for NCP deployments with 2,047 HGX servers, sixty-four WEKA storage servers. Every HGX server connects to the storage network with two 400GbE links using the appropriate cable type. Each WEKA storage server connects to the storage network with two 400GbE links using the appropriate cable type.

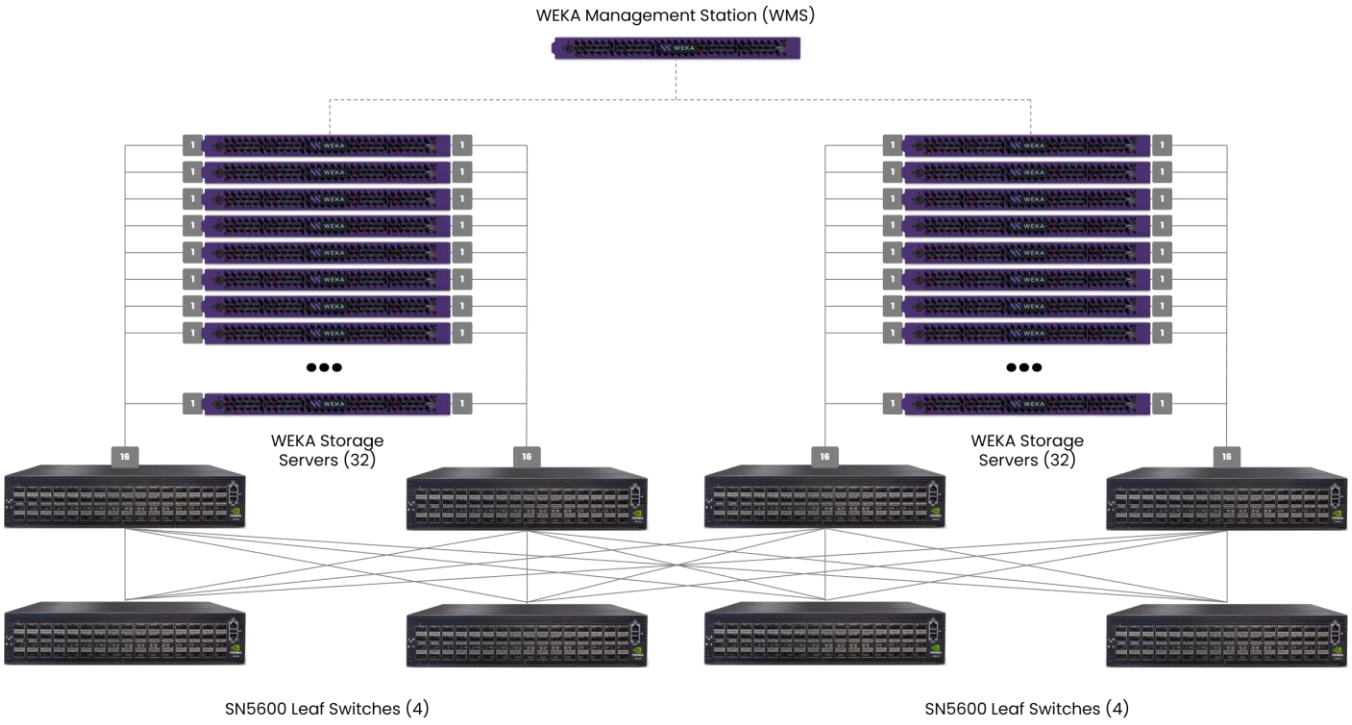

Figure 14. WEKApod reference architecture for NCP deployments with 2,047 HGX servers.

Description	Count
400GbE links (Storage)	128
OSFP ports (Switch) / Splitter cables (MCP700-Nxxx)	64
1GbE links (Management)	64

Table 10. Cable counts for NCP deployments with 2,047 HGX servers.

4. Summary

The WEKA Data Platform has pioneered an architecture that not only delivers outstanding performance and scalability but also ensures operational efficiency and ease of use. This document presents a validated reference design for WEKA Data Platform, specifically adhering to the NCP RA specifications for HPS. With flexible configurations supporting over 32K GPUs in a single cluster, NCPs can confidently pair WEKA Data Platform with large-scale AI infrastructure deployments. The jointly validated WEKA Data Platform reference architecture, based on the NCP RA and featuring HGX servers, meets and exceeds the rigorous demands of AI, HPC, and other data-intensive applications.

Appendix A: Data Services & Unified Security Model

The WEKA Data Platform supports multi-protocol and data-sharing capabilities across various protocols, allowing diverse application types and users to share a single pool of data. Unlike other parallel file systems, WEKA does not require gateway server infrastructure to offer this capability.

The currently supported protocols include:

- Full POSIX for local file system support
- NFS for Linux
- SMB for Windows
- S3 for Object access

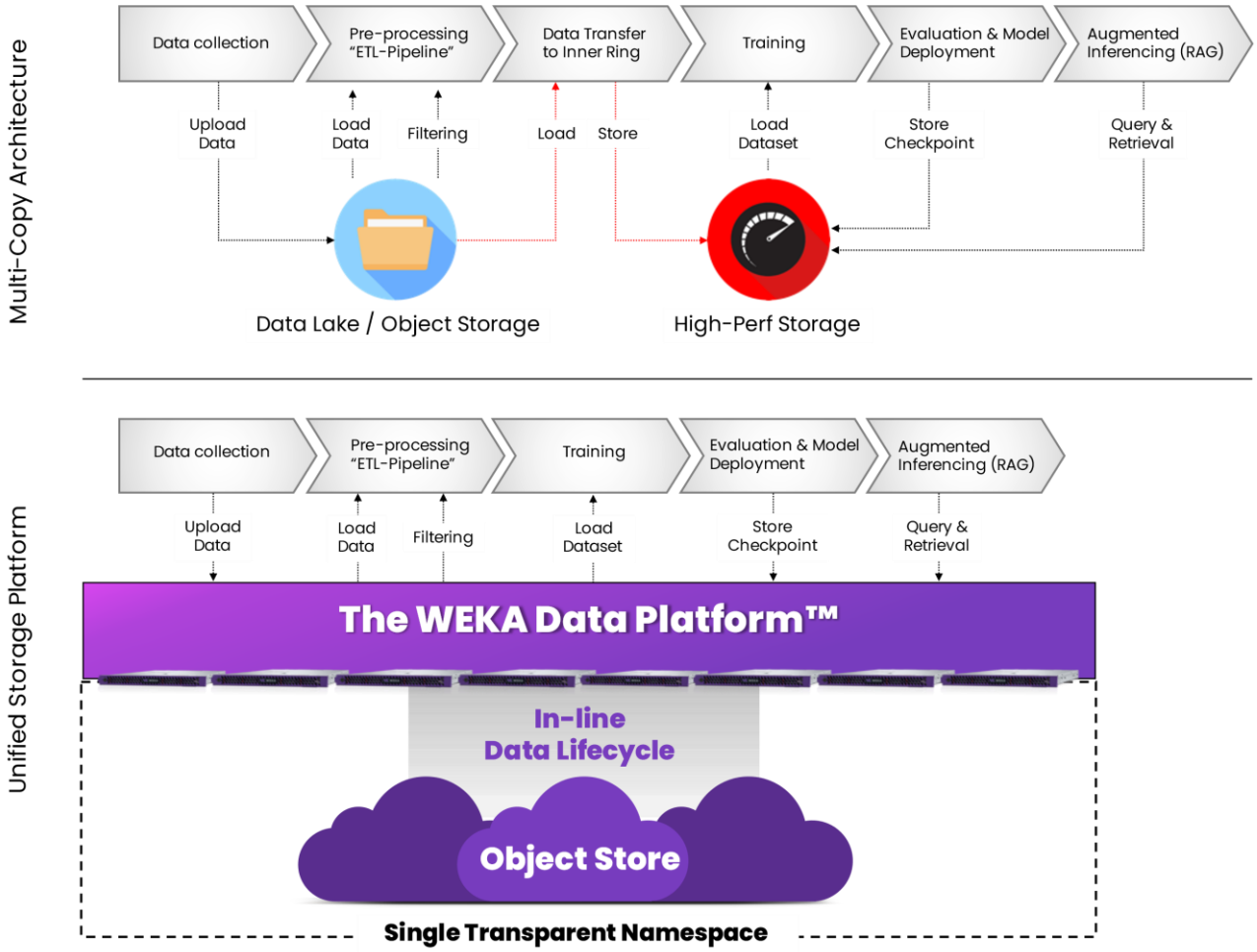

Figure 15. WEKA Unified Storage Platform provides ubiquitous data access across every stage of the AI pipeline.

By supporting POSIX and file/object protocols, the WEKA Data Platform allows organizations to consolidate HPS (training data) and data lake (data collection and preprocessing) storage needs onto a single platform. This consolidation eliminates the need for data transfers between separate storage systems, thereby simplifying AI training, inference, and model serving processes. This integration enhances speed and reliability, streamlining the entire AI data lifecycle. Consequently, organizations can focus on innovation rather than dealing with intricate multi-copy storage architectures and burdensome data transfer operations.

Furthermore, by unifying various storage systems into a single, cohesive data platform, WEKA empowers NVIDIA Cloud Partners to offer a highly differentiated solution that accelerates the entire AI pipeline. This capability enables NCPs to develop and deliver services to their customers more quickly and with reduced operational overhead, in contrast to traditional multi-copy architectures.