Artificial Intelligence (AI) and Deep Learning (DL) are transforming every industry around the world. They create the toughest workloads in modern computing history, and pose design challenges to compute, storage, and networking infrastructure that can be complex and time consuming. DDN A³I® solutions fully-integrate DDN Storage appliances and HPE Apollo 6500 servers for rapid and simple deployment of an AI-optimized infrastructure. DDN A³I solutions deliver the fastest application performance, and scale seamlessly to meet the demands of evolving AI workloads.

**DDN A³I with HPE Apollo 6500: Flexible Configurability and Maximum Acceleration**

DDN A³I (Accelerated, Any-Scale AI) solutions break new ground for AI and DL. Engineered from the ground up for the AI-enabled data center, DDN A³I solutions with HPE Apollo 6500 servers accelerate end-to-end data pipelines for AI and DL workloads of any scale. They are designed to provide extreme amounts of performance and capacity backed by a jointly engineered, validated architecture. DDN A³I with HPE Apollo 6500 is optimized at every layer of hardware and software to ensure data delivery and storage is fast, responsive and reliable.

The DDN shared parallel architecture enables full utilization of the Apollo 6500 GPU resources and maximizes developer productivity from every compute cycle on the Apollo 6500 server. DDN A³I solutions streamline DL by enabling all phases of the workflow to happen concurrently and continuously—including data ingest, curation, training, validation, inference, and simulation. With DDN A³I solutions, data is accessible simultaneously by multiple Apollo 6500 servers through a fully optimized and unified interface that is easy to use directly from applications.

**Easy to Deploy, Manage, and Use**

DDN A³I with HPE Apollo 6500 offers a turnkey and pre-configured solution, that is easy to deploy, shortening the timeline from AI concepts to business insights in a production setting. Based on comprehensive DDN A³I reference architectures, these solutions eliminate design guesswork, and have been validated in collaboration with HPE to ensure the highest performance, optimal efficiency, and flexible growth for Apollo 6500 servers.
The HPE Apollo 6500 Gen10 System is an ideal AI and DL platform providing unprecedented performance supporting industry leading GPUs, fast GPU interconnect, high bandwidth fabric, and a configurable GPU topology to match any workloads. With eight GPUs, up to 125 TFlops of single precision compute performance is available. The Apollo 6500 Gen10 can support multiple accelerator topologies and GPUs with NVLink 2.0, which enable dedicated GPU-to-GPU communication and increased bandwidth, or traditional PCIe GPUs. With efficient system management and security, comprehensive choice of enterprise options and operating system choices, the HPE Apollo 6500 Gen10 is a powerful, industrial-strength building block for AI infrastructure.

To meet the requirements of a variety of workloads, DDN A³I with HPE Apollo 6500 leverages the DDN AI200™ and AI7990™ storage appliances. The AI200 is an all-NVME, fully-integrated parallel file storage appliance that delivers 23GB/s of throughput and over 395K IOPS to applications, accelerating even the most I/O intensive workloads. The AI200 is specifically optimized to keep GPU computing resources fully utilized, ensuring maximum efficiency while easily managing tough data operations. The AI7990 is a hybrid, parallel file storage appliance that integrates both flash and deeply expandable capacity disk in a unified system for simplicity and flexibility. This integration makes it easy to collocate both hot training data and large libraries while maintaining optimal system efficiency. The AI7990 outperforms competing platforms and delivers the economics of capacity disk for your growing data library.

The DDN architecture allows for instant provisioning of new resources to applications, and easy deployment of additional Apollo 6500 servers. Advanced monitoring tools embedded within all components of the solution provide easy data management capabilities and extensive metrics for comprehensive optimization of live workloads.

Performance scaling test results for distributed training using multiple Apollo 6500 servers are available in the DDN A³I Solutions Guide.
High Performance AI at-Scale

DDN A³I solutions can scale seamlessly supporting additional Apollo 6500 servers and AI200 and AI7990 appliances for more performance and capacity through a single unified namespace. Storage and compute are fully-interconnected through a high-performance, low-latency network using EDR InfiniBand™ or 100Gbps Ethernet.

Fast and flexible access to a large and diverse data set is key to successful AI and DL efforts. The DDN high-performance parallel architecture delivers data to AI workloads with the highest bandwidth, lowest latency and maximum concurrency, ensuring full GPU resource utilization even for distributed DL training running across multiple Apollo 6500 servers simultaneously.

The DDN intelligent parallel client for Apollo 6500 server engages multiple high-speed data paths to the storage and delivers the full performance of NVMe flash directly to the application. The DDN shared parallel architecture provides predictable and linear scaling of performance for training across multiple GPUs on a single Apollo 6500 server or multiple Apollo 6500 servers simultaneously. With true end-to-end parallelism, DDN A³I with HPE Apollo 6500 eliminates the bottlenecks associated with legacy platforms.

Extensive performance and interoperability testing on widely-used DL frameworks demonstrate that standalone and containerized applications can now engage the full capabilities of the data infrastructure. All applications demonstrate significantly higher performance and shorter run times with DDN A³I than with legacy competing solutions.

Comprehensive application performance and interoperability test results are available in the DDN A³I Solutions Guide.
Backed by Deep AI Expertise from DDN and HPE

For 20 years, DDN has designed, developed, deployed and optimized solutions that enable organizations to generate value by accelerating time to insight from their data, both on-premises and in the cloud. In combination with HPE’s leadership and expertise in AI and DL, DDN has successfully deployed data-at-scale solutions for customers in all industries seeking to accelerate their business with AI and DL.

Fully-integrated and optimized, DDN A³I solutions with HPE are delivered and supported by expert partners worldwide, certified to provide comprehensive value-add services for enterprise.

DDN A³I solutions with HPE enable users and applications to easily harness the massive AI power of the Apollo 6500 server for the most efficient utilization of GPU resources, and the fastest time-to-insights. DDN A³I solutions make AI-powered innovation easier, with faster performance, effortless scale, and simplified operations, backed by the AI infrastructure experts.

FOR MORE INFORMATION VISIT:
ddn.com/products/a3i-accelerated-any-scale-ai/