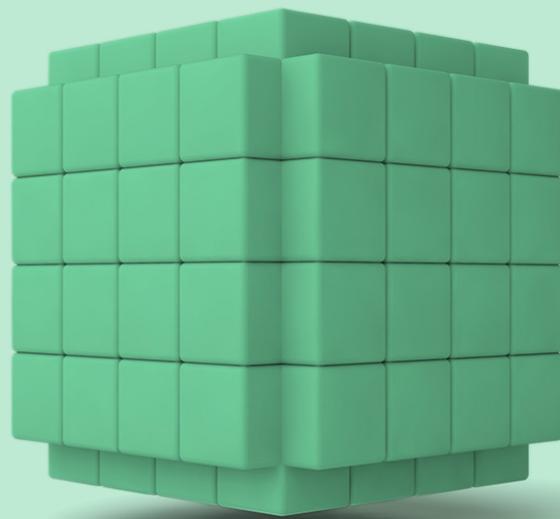# NVIDIA DGX Foundry and NetApp

The development hub
for the world's most
demanding AI enterprises

■ NetApp

The challenge

Many businesses find themselves sinking in artificial intelligence "model debt" with data science innovation languishing in prototypes and rarely realized in production. Typical challenges that constrain time to insights include data scientist productivity, workflow bottlenecks, inefficient access to resources and infrastructure, and lack of integration in IT and DevOps processes. Businesses need a solution that accelerates their AI initiatives with simplified experimentation and workflow, as well as streamlined management of users, jobs, and datasets, that is effortlessly supported by IT teams.

The solution

NVIDIA DGX Foundry, which includes NVIDIA Base Command™ software and uses NetApp storage, is a world-class infrastructure solution for businesses and their data scientists who need a premium AI development experience without the struggle of building it themselves. Offered as a hosted solution, it includes access to fully managed infrastructure based on the NVIDIA DGX SuperPOD™ architecture.

AI expertise from NVIDIA and NetApp

To succeed in AI, enterprises need flexible and simple ways for data scientists to define, configure, and execute their workloads; collect, store, and inspect results; and iterate for further innovation. At the same time, management demands comprehensive reporting on AI infrastructure utilization to set priorities and plan for the future. NVIDIA DGX Foundry is an immediately usable, cloud-based solution for AI development without the overhead and pitfalls of deploying and running a do-it-yourself platform.

Coordinated access to NVIDA DGX SuperPOD

Enterprises are starting to infuse AI into their businesses and investing in powerful infrastructure, but their IT organizations often struggle to provide simple but powerful tools to their researchers and scientists to share that infrastructure efficiently.

Key benefits

- **AI expertise from NVIDIA and NetApp**
  Shorten your cycle time from AI concept to production.
- **Coordinated access to DGX SuperPOD™**
  Experience it with a simple subscription model.
- **Comprehensive AI workflow management**
  Share across all your data science teams.
- **Integrated monitoring and reporting**
  Gain unprecedented insights to optimize utilization.

DGX Foundry is the easiest way to access premium accelerated computing infrastructure and subscription storage services, by providing a cloud-hosted solution that gives your data scientists seamless access to industry-leading AI infrastructure with NVIDIA DGX SuperPOD and NetApp.

Comprehensive AI workflow management

Deep learning and data science practitioners need to solve increasingly bigger and more complex data science problems faster. They need ready-to-run, optimized AI software that eliminates tedious setup and testing, and they want end-to-end AI experimentation and workflow management that enables them to take data science development from concept to production scale.

Base Command, included with DGX Foundry, offers users both a comprehensive cloud-based UI, as well as a complete command line interface, to efficiently execute AI workloads with right-sized accelerated resources, from a single GPU to a multi-node cluster, with datasets hosted on integrated dedicated NetApp storage. A vast library of pre-built containers with NVIDIA-optimized deep learning and data science frameworks and pre-trained models is instantly available via NVIDIA NGC, a catalog of curated sets of GPU-optimized software for AI, HPC, and visualization. This lets them deliver production-ready models and applications sooner, and eliminates the overhead of systems management.
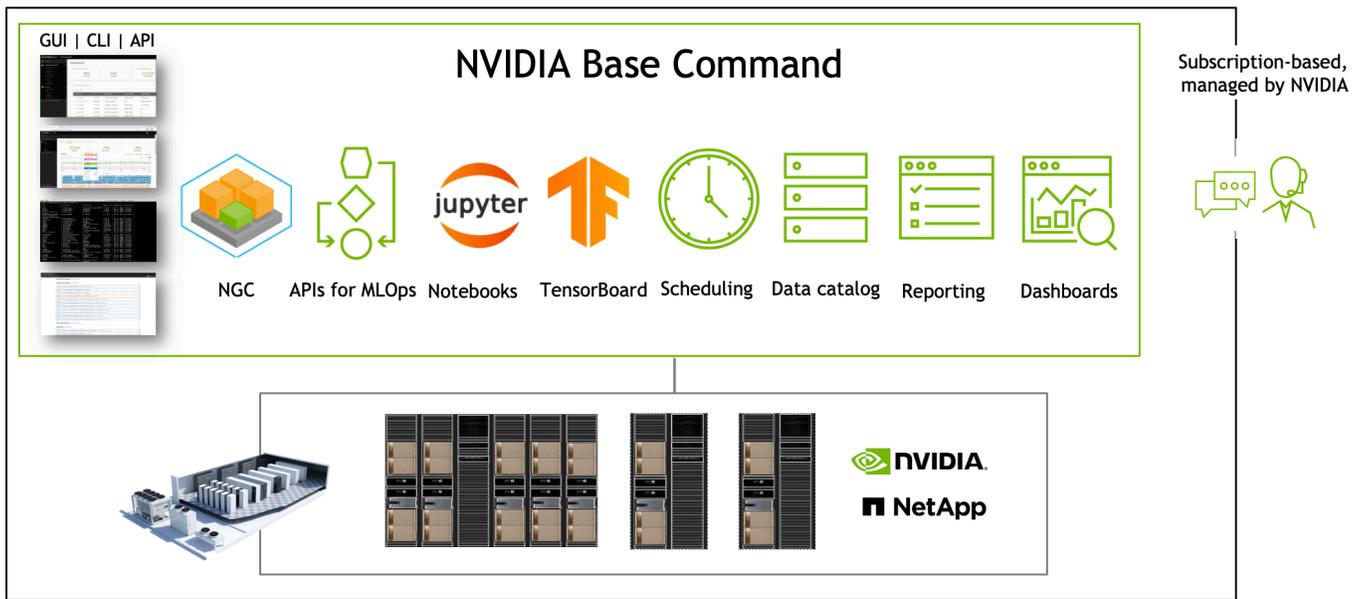
Figure 1) NVIDIA Base Command with NetApp storage.

Integrated monitoring and reporting

Strategically investing in AI to transform your business means increasing infrastructure utilization and developer productivity. Full visibility and accountability are required to achieve successful outcomes. AI projects are multifaceted, very iterative in nature, and require constant fine tuning. Base Command has built-in telemetry for end users, so that they can validate their deep learning techniques, workload settings, and resource allocations to constantly improve.

Base Command also includes reporting and showback capabilities for business leaders who want to measure AI projects against business goals. These capabilities are also helpful for team managers who need to set project priorities and plan for a successful future by correctly forecasting compute capacity needs. For system administrators, the solution provides extensive monitoring tools for the NVIDIA DGX A100 systems and NetApp® Keystone™ Flex Subscription services for the NetApp all-flash storage arrays. NetApp Keystone offers 24/7 proactive storage monitoring and support from NetApp.

The AI platform that NVIDIA uses

NVIDIA's own engineers and researchers rely on the same AI infrastructure every day, and have ever evolving demands as they build state-of-the-art AI. NVIDIA is continuously innovating for our own engineers, and is committed to provide these enhancements continuously to the Base Command software and NVIDIA DGX Foundry. Your end users get the new functionality that is vetted by hundreds of NVIDIA researchers first, with the efficiency and reliability that our researchers are accustomed to.

Solution components

- NVIDIA Base Command, a cloud-hosted AI development solution
- AI job execution on an NVIDIA DGX SuperPOD starting with a minimum of three dedicated NVIDIA DGX A100 systems
- NetApp Keystone Flex Subscription, file-based storage subscription service
- 24/7 support, infrastructure managed and hosted by NVIDIA
- Consumed in a subscription model

About NVIDIA

NVIDIA's invention of the GPU in 1999 sparked the growth of the PC gaming market and has redefined modern computer graphics, high performance computing and artificial intelligence. The company's pioneering work in accelerated computing and AI is reshaping trillion-dollar industries, such as transportation, healthcare and manufacturing, and fueling the growth of many others. More information at www.nvidia.com.

About NetApp

In a world full of generalists, NetApp is a specialist. We're focused on one thing, helping your business get the most out of your data. NetApp brings the enterprise-grade data services you rely on into the cloud, and the simple flexibility of cloud into the data center. Our industry-leading solutions work across diverse customer environments and the world's biggest public clouds.

As a cloud-led, data-centric software company, only NetApp can help build your unique data fabric, simplify and connect your cloud, and securely deliver the right data, services, and applications to the right people—anytime, anywhere. www.netapp.com.

∏ NetApp