

Support Predictive Analytics with a Dell AI Starter Kit

Core Infrastructure Components

Dell PowerScale



Scale-out NAS storage

- Support any workload
- Scale to petabytes
- Protect your data

Dell PowerEdge



Compute to accelerate AI innovations

- Performance and scale
- Designed for sustainability
- Reliability and security

Dell PowerSwitch



Networking modernization

- Software-defined networking
- SONIC – Software for Open Networking in the Cloud

Deep Learning (DL) techniques have provided great successes in many fields such as computer vision, natural language processing (NLP), autonomous driving and predictive maintenance (PdM) by enabling a model to learn from existing data and then make corresponding predictions. This is the heart of PdM, a technique designed to help determine the condition of in-service equipment in order to estimate when maintenance should be performed to avoid unplanned outages, increase worker safety, reduce downtime and more.

The success of DL in PdM is due to a combination of improved algorithms, access to larger datasets and increased computational power. Choice and design of the system components, carefully selected and tuned for DL use-cases, can have a big impact on the speed, accuracy and business value of implementing AI techniques for PdM.

In such a complex environment, it is critical that organizations can rely on vendors they trust. Over the last few years, Dell and NVIDIA have established a strong partnership to help organizations fast-track their AI initiatives. This document demonstrates how the Dell PowerScale all-flash scale-out NAS, Dell PowerEdge servers with NVIDIA GPUs and Dell PowerSwitch networking can be used to provide an excellent environment for small teams performing data science, AI and deep learning for PdM. The kit described provides a quick starter architecture to experiment and tune models for small production environments or singular solutions within larger production data centers.

Dell PowerScale

An efficient data science team often needs to share massive amounts of data while providing high performance, reliability and seamless access from multiple operating systems. Dell PowerScale scale-out NAS (network attached storage) provides this critical capability. With its ability to easily scale capacity and performance, PowerScale allows data science teams to effectively collaborate and share data across different applications and systems. Its flexible deployment options, including on-premises, hybrid cloud and multi-cloud, provide the necessary agility to adapt to changing business needs and future-proof the deployment.

PowerScale also offers advanced security features, such as file system and volume-level encryption and secure access zones, ensuring the confidentiality and integrity of sensitive data. These features make PowerScale an essential

component in any data science team's toolkit, enabling them to efficiently manage, analyze and derive insights from large amounts of data.

PowerScale all-flash storage platforms – powered by the OneFS operating system – provide a powerful yet simple scale-out storage architecture to speed up access to massive amounts of unstructured data while dramatically reducing cost and complexity. They deliver extreme performance and efficiency for your most demanding unstructured data applications and workloads.

- **Support any workload.** Choose from all-flash, hybrid and archive nodes for the best fit for your data. Run multiple data protocols with simultaneous access to avoid storage silos. Deploy as an on-prem NAS appliance, in APEX or in the cloud.
- **Scale to petabytes.** Scale up, down or out non-disruptively to tens of petabytes. Manage your storage infrastructure with a single UI with CloudIQ. Manage your datasets across your enterprise.
- **Protect your data.** Receive built-in availability, redundancy, security, data protection and replication with OneFS. Protect from cyber-attacks with integrated ransomware defense and smart Air Gap. PowerScale is designed for 6x9s availability.

Dell PowerEdge

The latest generation of Dell PowerEdge servers enhance both business agility and time to market, with the ability to support transformational workloads such as databases and analytics, virtualization, software-defined storage, virtual desktop infrastructure (VDI), containerization, HPC, AI and ML. PowerEdge systems can draw from NVIDIA's full AI stack – including GPUs, DPUs and the NVIDIA AI Enterprise software suite – providing enterprises the foundation required for a wide range of AI applications, including speech recognition, cybersecurity, recommendation systems and a growing number of groundbreaking language-based services.

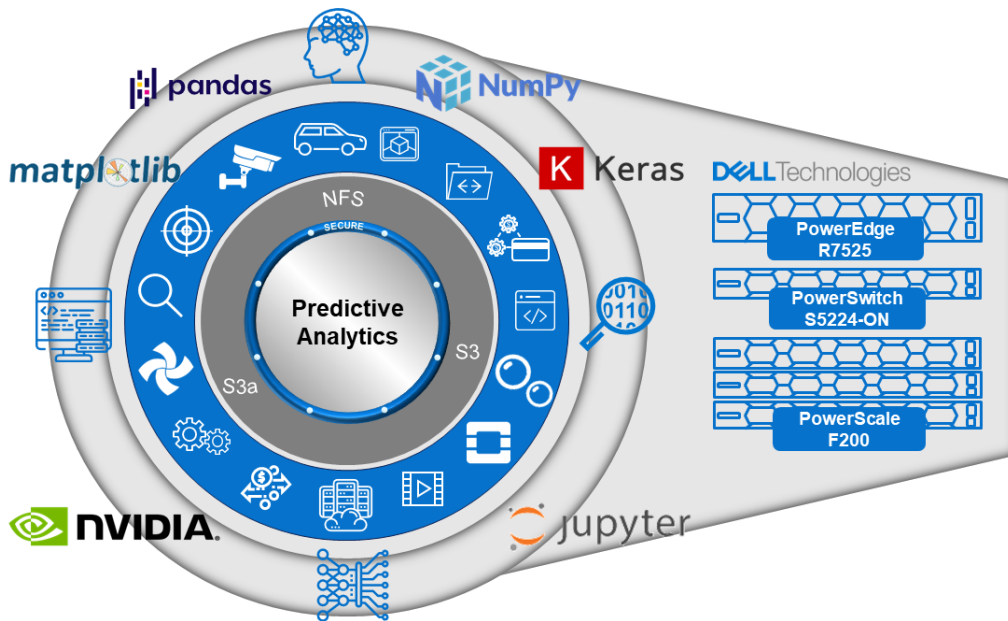
- **Performance and scale.** Next-generation PowerEdge servers provide improved performance, delivering greater AI inferencing. PowerEdge systems may be ordered with NVIDIA Bluefield data processing units to provide additional offload, acceleration and workload isolation capabilities ideal for power efficiency for private, hybrid and multicloud deployments.
- **Designed for sustainability.** Dell Smart Flow design, a new feature within the Dell Smart Cooling suite, increases airflow and reduces fan power by up to 52% compared to previous-generation servers. The Smart Flow design supports greater server performance with less power required to cool systems for more efficient data centers. An array of air movers are available from turnkey to high-end to best meet server cooling needs.
- **Reliability and security.** Dell's cyber-resilient architecture is a layered security approach consisting of a web of security solution elements designed to protect, detect and recover from threats. We enhance supply chain security with the Secured Component Verification (SCV) offering. SCV allows customers to cryptographically verify that the components set in the factory match what was delivered to them.
- **Accelerate Zero Trust adoption.** PowerEdge constantly verifies access, assuming every user and device is a potential threat. At the hardware level, silicon-based hardware root of trust, with elements including the Dell Secured Component Verification (SCV), helps verify supply chain security from design to delivery.

Dell PowerSwitch

Dell Technologies is keenly aware of the challenges that exist in the networking space and what must be done to address the limitations brought on by slow-moving legacy, proprietary networks and their impact on AI initiatives. With Dell Technologies Open Networking, we offer a complete strategy that combines networking scalability and agility with standards-based hardware and innovative, best-in-class software solutions and the automation tools to streamline a large amount of manual intervention. You'll be in a better position to meet workflow and application demands with greater network flexibility and control.

- **Software-defined networking.** Dell Networking Operating Systems deliver full-featured software-defined networking functionality with Layer 2 and Layer 3 connectivity that meets your needs with software from Dell and Open Networking ecosystem partners.
- **Software for Open Networking in the Cloud (SONiC).** Dell Technologies offers a finely-tuned, enterprise-ready and globally supported distribution of SONiC – called Enterprise SONiC Distribution by Dell Technologies – to help bring the benefits of hyper-scale-focused SONiC contributions to rest of the Enterprise and Telco markets and use cases.

Predictive Analytics Stack Architecture



Architecture components

Predictive Analytics architectures incorporate a variety of hardware and software components. Dell Technologies offers a large selection of hardware to build such architectures, starting from compute servers with the PowerEdge family, PowerSwitch for networking and PowerScale for distributed storage. In this solution we used PowerEdge servers populated with NVIDIA GPUs running Ubuntu 20.04 LTS release.

To leverage NVIDIA GPUs, we used the NVIDIA Container toolkit, which allows users to build and run GPU accelerated containers. For more details about this tool please visit [NVIDIA's website](https://nvidia.com/en-us/gpu-compute/container-toolkit/). Finally, we used a customized docker container based on NVIDIA's TensorFlow docker image (available at nvcr.io/nvidia/tensorflow:22.12-tf2-py3). This image provides a large ecosystem of tools that allows engineers and data scientists to develop ML applications using JupyterLab, TensorFlow, Keras, RAPIDS cuDF libraries and many more. This methodology offers the flexibility of Docker: Users can build and customize their own images and deploy specific Docker containers based on their needs.

Core hardware and software components

Component	Description
PowerScale F200	PowerScale F200 delivers the performance of flash storage in a cost-effective form factor to address the needs of a wide variety of workloads. Each node allows you to scale raw storage capacity from 3.84 TB to 30.72 TB and up to 7.7 PB of raw capacity per cluster. The F200 includes in-line compression and deduplication. The minimum number of PowerScale nodes per cluster is three while the maximum cluster size is 252 nodes.
PowerSwitch S5224-ON	The S5200-ON is a complete family of switches:12-port, 24-port and 48-port 25GbE/100GbE ToR switches, 96-port 25GbE/100GbE Middle of Row (MoR)/End of Row (EoR) switch and a 32-port 100GbE Multi-Rate Spine/Leaf switch.

Component	Description
PowerEdge R7525	The Dell PowerEdge R7525 is a 2-socket, 2U rack-based server designed to run complex workloads using highly scalable memory, I/O capacity and network options. The system is based on the 2nd Gen AMD EPYC processor (up to 64 cores), has up to 32 DIMMs and PCI Express (PCIe) 4.0-enabled expansion slots, and supports up to three double-wide 300W or six single-wide 75W accelerators.
NVIDIA Container toolkit	The NVIDIA Container toolkit allows users to build and run GPU accelerated containers. The toolkit includes a container runtime library and utilities to automatically configure containers to leverage NVIDIA GPUs.
JupyterLab	JupyterLab is the latest web-based interactive development environment for notebooks, code and data. Its flexible interface allows users to configure and arrange workflows in data science, scientific computing, computational journalism and machine learning.

References

- [Dell PowerScale Scale-Out File Storage](#)
- [Dell PowerEdge Servers](#)
- [Dell PowerSwitch Networking](#)



Learn more about the core products.



Connect with a Dell expert.



Follow us for product news.