Delivering Fully-Automated, High-Performance Storage Over NVMe-oF for AI & Data Analytics

On-demand instances of compute and storage to deliver the exact resources your applications need

EXECUTIVE SUMMARY

Organizations are realizing the economic value of data and analytics in gaining a competitive edge, which in turn is driving the growth of artificial intelligence (AI), deep learning (DL), and machine learning (ML). Data analytics places a higher demand on the underlying IT infrastructure. For example, data ingestion impacts capacity and throughput requirements, data streaming from databases requires incredibly high I/O and model training necessitates rapid reads of large volumes of random small blocks of data. To accomplish this, increased compute and GPU power are essential along with a high-performance fabric with accelerators and offloads. Without these, data analytic workloads suffer from reduced utilization and performance. To properly scale data analytic projects efficiently and provide proper performance, a dynamic and elastic infrastructure is required to handle the scale, performance, latency, and capacity demands.

INFRASTRUCTURE CHALLENGES

The rising demand for data analytics is causing a shift to infrastructure that can deploy underlying resources on the fly to match the requirements across the varying stages of the data pipeline (Figure 1.); ingest, preparation, training, and inference. Many of the hyperscalers have been reaping the benefits of data analytics through the means of AI for the last decade. The results they have shown through predictive analytics have optimized supply chain ordering and personalizing recommendations for online shoppers.

KEY BUSINESS BENEFITS

- Single pane of glass resource orchestration for all resources at each stage of the pipeline
- Ability to instantly attach/detach data sets and storage, and replace failed components in seconds
- Heterogeneous, commodity storage deployments with high performance and high availability, and without complexity
- Local-drive equivalent performance at scale and throughput that scales with demand
- Block storage foundation that supports any file system or distributed file system
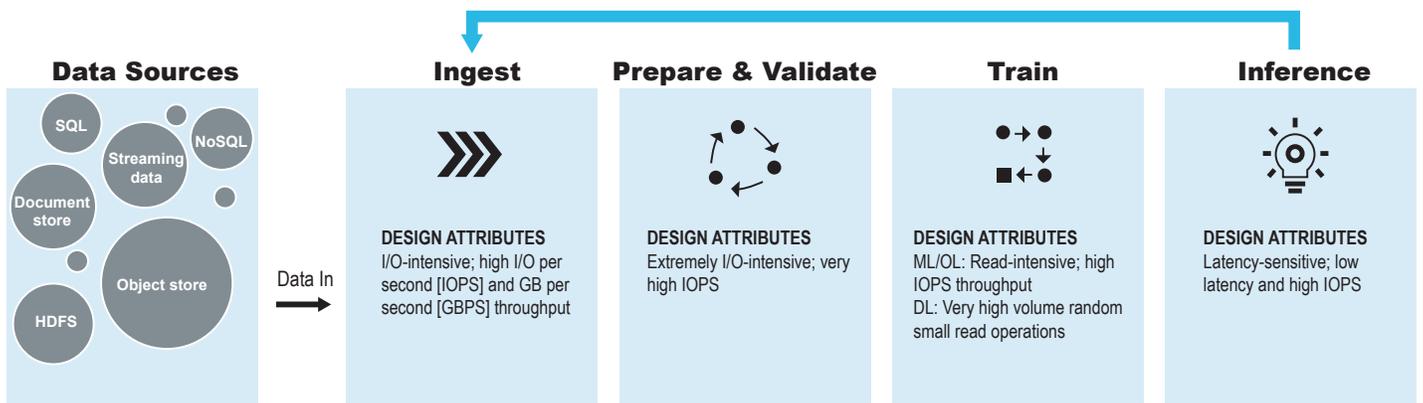- Scale-out architecture in any size increments

Data Sources

- SQL
- NoSQL
- Streaming data
- Document store
- Object store
- HDFS

Data In →

Ingest

DESIGN ATTRIBUTES
I/O-intensive; high I/O per second [IOPS] and GB per second [GBPS] throughput

Prepare & Validate

DESIGN ATTRIBUTES
Extremely I/O-intensive; very high IOPS

Train

DESIGN ATTRIBUTES
ML/OL: Read-intensive; high IOPS throughput
DL: Very high volume random small read operations

Inference

DESIGN ATTRIBUTES
Latency-sensitive; low latency and high IOPS

Figure 1. Data analytic pipeline and demands on infrastructure

However, most IT organizations don't have the budget to build a system of enormous scale, nor do they have the decade of experience to perfect it. For organizations in the early stages of deploying AI or cloud service providers who offer ML as a service, the cost and inflexibility of the underlying data server and storage infrastructure can cause significant limitations. Determining the types and amount of resources required prior to deploying a workload is a complex task and often leads to inefficiencies.

Many may start by using a public cloud. It's a quick way to start, allows for a decreased learning curve, and can keep costs down. However, as projects start to grow, costs can quickly escalate, and organizations soon look to move projects on-premise. The problem with bringing projects in-house is the lack of experience, which usually leads to underutilized GPUs and storage. Or worse, a lack of capacity to properly scale, which leads to wasted investment dollars.

To properly enable AI and ML, a move to an infrastructure that is programmable, adaptable, and API-driven is a must. The ability to deploy underlying resources on the fly allows for matching the demands of each varying stage of the data pipeline. And, the ability to compose resources to match the demands of a variety of workloads allows for cost-effective scaling and increased resource utilization across the deployment of AI, DL, and ML workloads.

A COMPOSABLE PLATFORM FOR DATA ANALYTICS

Together, DriveScale and Mellanox provide a disaggregated, high performance, scalable, end-to-end compute, GPU, storage, and data fabric solution that increases efficiency and utilization across data analytic workloads. Industry-standard compute nodes, and commodity storage, which helps reduce overall costs, can be configured as resource pools that can be provisioned as required and optimized for different data and workload requirements to enable higher utilization and performance.

As compute, accelerators and storage requirements change, hardware deployment can be orchestrated to adapt to exactly what is needed through

a single pane of glass (or RESTful API). Through the DriveScale Composable Platform building blocks – one SKU for server nodes, one for NVMe Storage, and one for HDD storage can be used to create a vast range of cluster configurations through the touch of a button. An orchestrator composes resources that can be expanded, reduced, replaced, and redeployed on-demand.

HIGH-PERFORMANCE FABRIC

Each varying stage of the data pipeline places a variety of strain on the underlying physical network. Data analytic workloads require fast access to data, immediate delivery of results, and demand intensive CPU processing. If the data can't be reached or CPU cycles are not available, applications can choke. The result is limited efficiency. Likewise, scalability is a significant concern as the network must accommodate rapid changes and the growth of data. Mellanox ConnectX™ adapters support NVMe-oF, which utilizes RDMA for more efficient and lower latency data transfers. NVMe-oF is a high-performance storage protocol designed to take advantage of faster flash storage over RoCE. RoCE is a network protocol that allows remote direct memory access over Ethernet — utilizing RDMA or NVMe-oF offloads the data transfer functions to the network adapter to bypass the CPU. These technologies provide for a more efficient and faster way to move data between networked computers or storage while lowering latencies and CPU utilization.

With more powerful CPUs and GPUs and faster storage used in data analytics, the switch is at the heart of the network, and it must be able to keep pace with much faster and intensive data movements between compute and storage servers. This is accomplished by moving data at low latency between CPUs, GPUs, memory, and storage. Using NVMe allows for accessing remote data at almost the same speeds as accessing local data. Mellanox Spectrum™ Ethernet switches provide 100/200GbE line-rate performance and consistent low latency with zero packet loss. Spectrum is also the only RoCE-ready switch that can deploy RoCE effortlessly, offer end-to-end automatic RoCE acceleration and real-time RoCE visibility for easy troubleshooting.
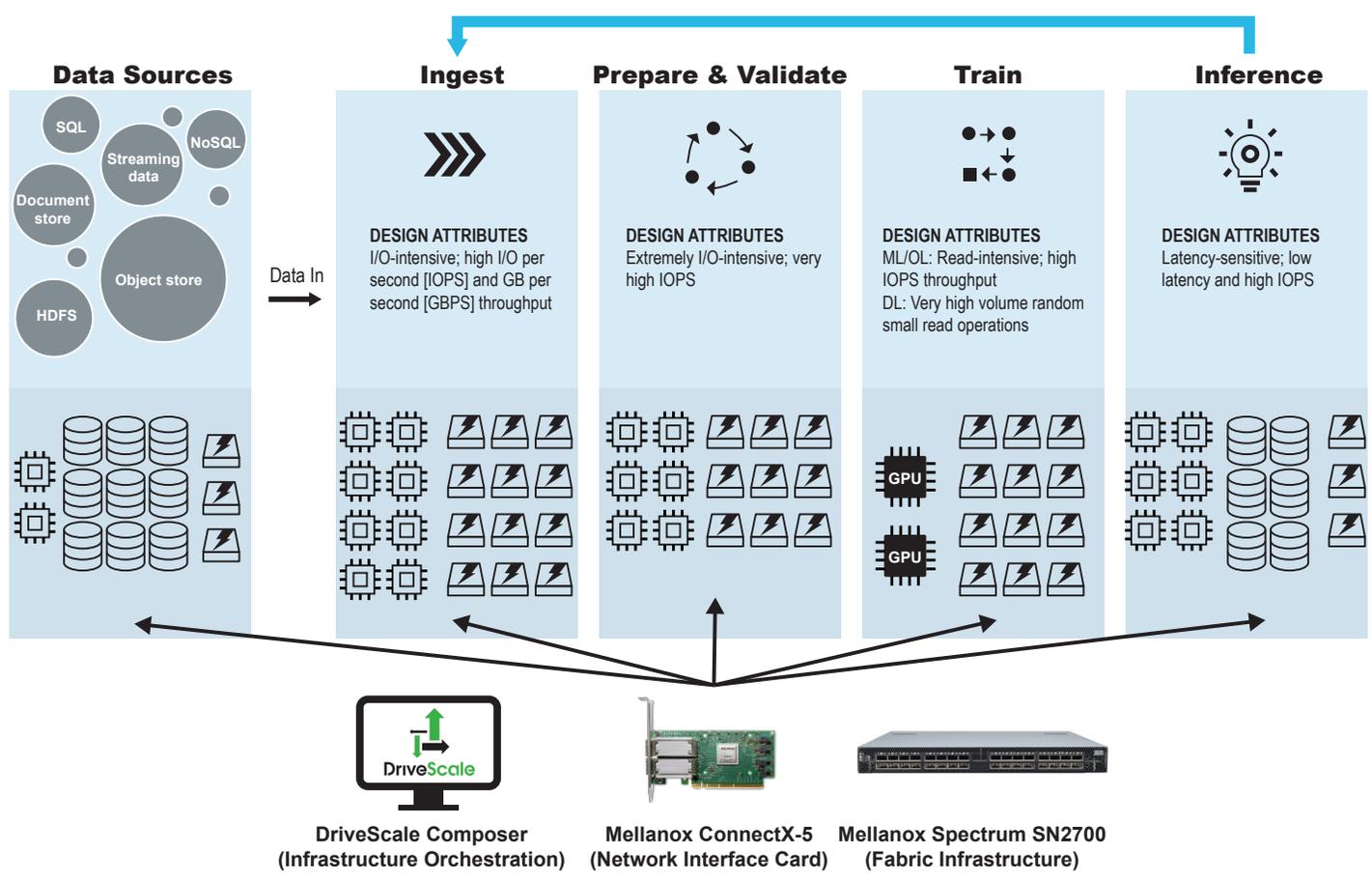
Figure 2. DriveScale Composable Platform and Mellanox high-performance infrastructure

DRIVESCALE, MELLANOX AND NVME

DriveScale Composable Platform software, when used with Mellanox ConnectX intelligent Ethernet adapters and Spectrum Switches (Figure 2.), is fully automated through the discovery of servers, storage, and data fabric capabilities. DriveScale works with Mellanox to detect data fabric connectivity options, including NVMe over Fabrics (NVMe-oF) using RoCEv2 for RDMA or NVMe-TCP, as well as performance optimized, multi-path iSCSI, offering choice and reduced complexity in preparing the data fabric. Users choose the instances of compute, and the instances of storage needed for a workload and DriveScale's software platform automatically sets up the highest performing fabric, mounts the devices, and creates a cluster ready for workload deployment.

CONCLUSION

DriveScale and Mellanox transform industry-standard compute nodes, GPU nodes, and storage systems into elastic resource instances enabling users to deploy IT resources on-demand, expand GPU compute capacity or storage as workloads require. Redeploy GPU compute nodes to additional datasets after a workload completes, and replace failed components instantly via DriveScale software. Mellanox high-performance, low-latency fabric combined with accelerators and offloads ensure the proper fabric to provide efficient data analytics and ensure future scalability. By deploying DriveScale on a Mellanox infrastructure, IT organizations can reduce cost and complexity while driving up CPU efficiency and GPU utilization. Enterprises can achieve hyperscale performance at a fraction of the cost of public cloud solutions and without the need for a highly trained staff.

About DriveScale

DriveScale is a team of entrepreneurs and experts in diverse disciplines who are united by a common quest: to re-imagine static compute infrastructure as adaptable and programmable, making every data center a composable, agile, scale-out cloud. This gives any company the ability to enjoy the same speed, flexibility, and cost-efficiency that the Cloud Giants create for themselves. With the growth of data-driven and distributed applications, enterprise and cloud companies will benefit from the rapid deployment, flexible operations and high availability previously only found in the largest data centers in the world. This is our vision for your infrastructure.

For more information, visit www.drivescale.com.

About Mellanox

Mellanox Technologies is a leading supplier of end-to-end InfiniBand and Ethernet interconnect solutions and services for servers and storage. Mellanox interconnect solutions increase data center efficiency by providing the highest throughput and lowest latency, delivering data faster to applications and unlocking system performance capability. Mellanox offers a choice of fast interconnect products: adapters, switches, software, cables and silicon that accelerate application runtime and maximize business results for a wide range of markets including high-performance computing, enterprise data centers, Web 2.0, cloud, storage and financial services.

To find out more, visit our website: www.mellanox.com

WANT TO LEARN MORE?

DriveScale Compostable Platform:

https://drivescale.com/composable-platform/

Mellanox ConnectX Ethernet Adapters:

https://www.mellanox.com/products/connectx-smartnic/

Mellanox Spectrum Ethernet Switches:

https://www.mellanox.com/page/products_dyn?product_family=218&mtag=spectrum_ic

350 Oakmead Parkway, Suite 100
Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com