



WHITE PAPER

Quantum Data Management for ADAS, Autonomous Driving, Robotics, and Industrial AI/ML-Based Solutions

Overview and Reference Architecture

Quantum[®]

CONTENTS

Introduction	3
ADAS/AD Development Workflow	4
ADAS/AD Platform High-Level Architecture	6
Platform locations	7
Test Vehicle	7
Remote Site (Garage)	7
Development Center	8
Public Cloud	8
Quantum ADAS/AD Reference Architecture	9
Quantum StorNext File System Overview	9
Quantum ADAS/AD Storage Technology Stack	12
Quantum In-Vehicle Storage Appliance	12
Remote Site Reference Architecture	13
Data Ingestion Tier	14
StorNext File System	14
Performance Tier	15
Capacity Tier	16
Object Storage	16
Archive Tier	16
Clients and Connectivity	17
Development Center	17
Object Storage	18
Archive Tier	19
Clients and Connectivity	19
Conclusion	19
Related Documents & Links	19

ABSTRACT

This document defines a reference architecture for a high-performance, scalable storage infrastructure deployed during the development and production of autonomous and connected automated systems. The storage infrastructure is implemented with Quantum software and appliances.

INTRODUCTION

Although most of the examples in this article are focused on Autonomous Driving, the described data management workflows and storage technology are directly applicable to many other industries and use cases. They enable the development and production systems for autonomous Smart Farming, Industrial Robots, and Drones to achieve the required scale and performance. These systems too, rely on collecting and processing extremely large amounts of unstructured data and sophisticated Machine Learning (ML) algorithms.

The automotive industry is going through a significant business and technology transformation driven by the desire to improve vehicle safety, comfort, and efficiency and to release fully autonomous vehicles as the ultimate goal.

The US Department of Transportation (DoT) uses J3016's six levels of automation for on-road motor vehicles in its "Federal Automated Vehicles Policy." This document became a de facto global standard adopted by most stakeholders in the automated vehicle industry.

The Society of Automotive Engineers (SAE) organization developed a visual chart describing the six driving automation levels.



SAE J3016™ LEVELS OF DRIVING AUTOMATION

		SAE LEVEL 0	SAE LEVEL 1	SAE LEVEL 2	SAE LEVEL 3	SAE LEVEL 4	SAE LEVEL 5
What does the human in the driver's seat have to do?		You <u>are</u> driving whenever these driver support features are engaged – even if your feet are off the pedals and you are not steering			You <u>are not</u> driving when these automated driving features are engaged – even if you are seated in “the driver’s seat”		
		You must constantly supervise these support features; you must steer, brake or accelerate as needed to maintain safety			When the feature requests, you must drive	These automated driving features will not require you to take over driving	
What do these features do?		These are driver support features			These are automated driving features		
		These features are limited to providing warnings and momentary assistance	These features provide steering OR brake/acceleration support to the driver	These features provide steering AND brake/acceleration support to the driver	These features can drive the vehicle under limited conditions and will not operate unless all required conditions are met	This feature can drive the vehicle under all conditions	
Example Features		<ul style="list-style-type: none">• automatic emergency braking• blind spot warning• lane departure warning	<ul style="list-style-type: none">• lane centering OR• adaptive cruise control	<ul style="list-style-type: none">• lane centering AND• adaptive cruise control at the same time	<ul style="list-style-type: none">• traffic jam chauffeur	<ul style="list-style-type: none">• local driverless taxi• pedals/steering wheel may or may not be installed	<ul style="list-style-type: none">• same as level 4, but feature can drive everywhere in all conditions

Figure 1 - SAE J3016 Levels of Driving Automation

Many car manufacturers and their suppliers are developing and deploying Advanced Driver-Assistance Systems (ADAS) and Autonomous Driving (AD) systems to achieve these automation levels. While a few manufacturers have achieved Level 3 & 4 in pilot deployments, industry analysts agree that Level 5 remains several years away as companies compete aggressively to develop the systems to ensure customer safety. These systems incorporate various complex Data Management, Analytics, ML, and AI technologies, which rely heavily on vast amounts of data mostly collected by test cars or environment simulators.

The ADAS/AD technology is also related to the Connected and Automated Vehicle (CAV) development. Connected vehicles use different communication technologies to interact with other cars on the road (vehicle-to-vehicle [V2V]), roadside infrastructure (vehicle-to-infrastructure [V2I]), cloud services [V2C], pedestrians [V2P]. There are a variety of sensors, including multiple cameras, radars, and LIDARs, deployed on the test vehicles, generating between 10 TB and over 400 TB daily. The ADAS/AD system has to process hundreds of petabytes of data during the development. This large amount of data presents a significant challenge for the development process. ADAS/AD system developers must deploy a fast, scalable, and reliable data platform to deal with this challenge effectively. The data platform often operates with data distributed between multiple test vehicles, car garages (Remote Sites), development centers, public clouds, or at least some of these locations. The data has to be available for efficient and secure access at every location and archived for extended periods to meet legal and regulatory requirements.

This document describes Quantum's advanced storage solutions for every stage of the ADAS/AD development platform. Starting at the edge, the Quantum R-Series ruggedized in-car storage devices support high-speed data capture in the vehicle and fast data upload at the data centers. Quantum data center solutions scale from hundreds of terabytes (TB) to multiple zettabytes (ZB) of managed storage capacity.

Quantum provides a flexible, multi-tier software-defined storage infrastructure controlled by the high-performance, parallel Quantum StorNext® File System. Following user-defined policies, data is moved seamlessly across ultra-fast Quantum F-Series NVMe appliances, cost-effective high capacity Quantum QXS™, on-prem Quantum ActiveScale™ Object Storage, Quantum Scalar® Tape Archives, and Public Cloud repositories while remaining under the same Quantum StorNext namespace. Quantum solutions deliver unprecedented scale and performance at the lowest Total Cost of Ownership (TCO) in the industry.

ADAS/AD DEVELOPMENT WORKFLOW

ADAS/AD system development cycle includes the following steps:

- Data Acquisition and Ingest
- Data Preparation
- ML Model Development
- Simulation and Validation
- ML Model Deployment

These steps are standard for most machine learning projects. They are continuously executed in parallel by the ADAS/AD development system while processing the stream of newly collected data and deploying the updated ML models to the fleet of test vehicles. The reference architecture described in this document focuses mainly on the acquisition, movement, storage, and curation of the data necessary to develop autonomous vehicle software.

An ADAS/AD platform implements a multistage data pipeline as part of an ML model development, training, and verification process. The data pipeline contains the following primary stages:

- **Data Acquisition and Ingest:** Each test vehicle collects data captured by various sensors observing the surrounding environment and the driver's actions on an in-vehicle logger device. The sensors include multiple video cameras with 4K or higher resolution, radars, LIDARs, GPS, and others, generating 10 TB to over 400 TB data per car per day. This data is periodically ingested from the vehicles to a data center over either a direct network connection between the in-vehicle logger and an ingestion station or by moving a removable storage magazine to the ingestion station. The ingested data is mostly unstructured, and a fast ingestion process demands high raw throughput from the storage system.
- **Data Preparation:** At this stage, the data is prepared to facilitate the needs of downstream operations. The data preparation involves removing and transforming low-quality data, performing integrity checks, removing redundant information, data enrichment with location, weather, time, and other metadata. These functions help build a catalog with searchable metadata allowing quick search and identification of driving situations and scenarios. The platform has to analyze data coverage and identify needs for more data collection. Automated and manual data labeling for object detection and object tracking are performed at this stage to achieve high-quality training data sets. The data preparation tasks perform multiple ETL (Extract, Transform, Load) operations, which require high random I/O data access performance from the storage infrastructure. The Data Preparation process might consume a significant portion of the overall ADAS/AD development time.
- **ML Model Development:** The main goal of the ADAS/AD system is to perform accurate object detection, identification, classification, localization, and movement prediction. This is achieved by developing and deploying a variety of deep learning ML models. They rely on an extensive model training process that requires very large training data sets. The ML model training is a GPU-driven operation and requires storage infrastructure with very low latency and high read throughput.
- **Simulation and Validation:** After the ML model is developed its accuracy has to be validated. This is achieved with Hardware-in-the-Loop (HiL) / Software-in-the-Loop (SiL) or other similar approaches. Captured raw sensor data is replayed at test setups with thousands of iterations running in parallel. This process is continuously repeated as new ADAS/AD ML models are developed to verify the functionality against previously captured data. To improve the ML models even further, simulated driving scenarios data is generated and used for model training and testing. ML models are enhanced and tested with this data without the need to capture more live data with test vehicles. The Simulation and Validation stage requires storage infrastructure with high-bandwidth parallel data access.
- **Archiving:** Following the final validation, the data is moved to lower-cost archive storage for potential future use. Data archiving is required for meeting regulatory and contractual commitments. The storage system should be scalable to support large amounts of data and maintain high security and reliability.

ADAS/AD PLATFORM HIGH-LEVEL ARCHITECTURE

An advanced ADAS/AD development platform is a distributed machine learning (ML) system capable of processing vast amounts of data. Figure 2 exemplifies the high-level architecture of such a platform. The platform is globally distributed across test vehicles, remote sites, development centers, and public cloud infrastructure. At its core, it implements a globally distributed data lake containing all raw sensor data, simulated data, and metadata. Many of the data pipeline operations listed in the previous chapter are performed at multiple locations, including vehicles, remote sites, development centers, and the public cloud. For quick data search and access, a global data catalog and data management system are required. Placing the data across multiple storage layers guarantees data availability at the required speeds, scale, and cost at each location. Multiple types of distributed and federated machine learning frameworks are frequently used. They impose intense performance demands on the globally distributed data lake. A combination of high-speed parallel file systems, NVMe block and file storage, scalable object store, and large-scale data archival storage must be deployed to meet these demands.

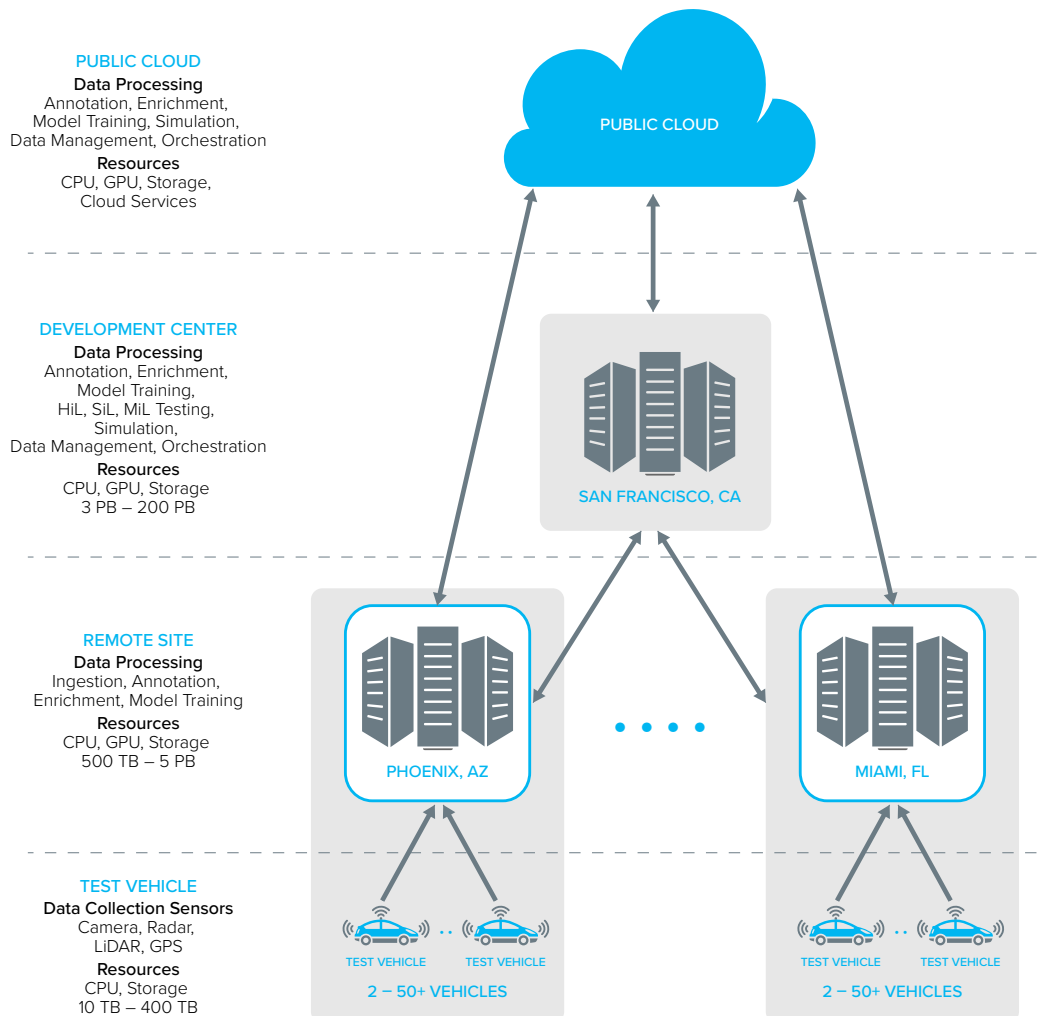


Figure 2 - ADAS/AD Platform Architecture

Platform Locations

The described platform architecture is modular, scalable, and capable of supporting the needs of companies of any size. The ADAS/AD platform contains four types of locations – Vehicle, Remote Site, Development Center, and Public Cloud (Figure 2). Companies developing and deploying ADAS/AD systems could maintain all or some of these locations. They could deploy very few or hundreds of test vehicles operating at one or multiple locations. All data could be processed at a single location, multiple locations, or pushed for processing at the company's public cloud-hosted infrastructure.

As the data accumulates, the design of the data lake becomes critical for preventing data access bottlenecks. Having the flexibility to automatically move data between different locations and perform data preparation and model training on-premises or on the public cloud brings significant performance gains and cost reductions to the whole development process. For example, outsourcing the data labeling process is very common and requires moving data to the cloud in many cases. The global data lake needs to be capable of storing at least 2 PB of data per test car per year. Support for cold or archival storage is also necessary to keep a large amount of infrequently accessed data.

Test Vehicle

ADAS/AD test vehicles are equipped with multiple sensors to capture a wide variety of sensor data. The sensors include multiple video cameras, radars, LIDARs, GPS, and others, generating 20 TB to over 400 TB data each day. This data is captured by an in-vehicle data logger and storage appliance and stored as files with ADTF v2/3, RTMaps, MDF4, and ROS/rosbag formats. For data uploading at the data center, the appliance supports a high-speed network interface and, in some cases, a removable solid-state storage cartridge. The data logger appliance may perform initial data preparation, such as data trimming, automatic labeling, and other types of data enrichment if hardware resources are available.

Remote Site (Garage)

The data collected by the vehicles is periodically offloaded at regional data centers, also called Remote Sites or sometimes Garages. The data ingestion typically occurs at device speeds as removable drives are cycled between the vehicle and the garage equipment. A direct network connection between the vehicle logger and the garage equipment could also be used for data transfer, but this approach is slower and requires the vehicle to be out of services for the duration of the data transfer.

In a typical ADAS/AD development cycle, after the raw data is ingested from the cars, it goes through the Data Preparation stage and later is used for ML Model Development. Some of the operations related to these stages are performed locally at the Remote Site, and the data is transferred to a central Development Center for further processing.

Multiple Data Preparation functions can be performed at Remote Sites, such as data classification, deduplication, compression, and enrichment. The reason to perform data cleaning and lossless data compression at this point is to reduce the amount of required local data storage as well as network bandwidth to move the data to a Development Center or Public Cloud infrastructure. Identifying the necessary car driving scenarios at the Remote Site could determine which data needs to be uploaded to the Development Center and which data could be stored at hot or cold storage locally. In many situations, less than 5% of the captured data is necessary for further processing.

Some ML Model Development functions could also be performed at the Remote Site. These operations require not only the availability of GPU resources but also storage infrastructure with high sequential read/write performance and low latency. To minimize the overall data storage cost but still deliver the required performance and scale, multi-tier data infrastructure could be deployed where only the necessary data is stored at the highest performance and most expensive storage tier. In most cases, the high-performance storage tier relies on NVMe technology, and HDD/SDD devices support the file or object store large capacity tier. For long-term archiving and cold storage, magnetic tape storage is the best solution.

During the ADAS/AD platform design, it should be considered that transferring large amounts of data to the Development Center or the Public Cloud infrastructure is costly, introduces significant development delays, and might introduce security, privacy, and regulatory concerns.

In most cases, the Remote Sites maintain up to 5 PB of data which is sufficient to store and process the locally collected data. High-throughput, wide-area network connectivity to the Development Center and the Public Cloud is deployed to enable the data aggregation from all Remote Sites when necessary.

Development Center

The main part of the ADAS/AD development is performed at an ADAS/AD Development Center. All necessary data from all Remote Sites is collected, and a full ADAS/AD data pipeline is implemented. As it was described earlier in this document, the data pipeline includes Data Ingestion, Data Preparation, ML Model Development, Model Validation, and Data Archiving stages. The Data Ingestion stage involves the collection of all data from all Remote Sites either over the network or through transferring of storage devices offline. The aggregated data size at the Development center could exceed hundreds of petabytes, which creates a significant challenge for the storage infrastructure. The data storage multi-tier architecture implemented at the Remote Sites is implemented at the Development Center but with a much larger storage capacity at each tier.

While most of the Data Preparation functions can be executed at Remote Sites, the Development Center is the best location for ML Model Development. The data collected at all Remote Sites by all test vehicles is aggregated locally and made available for ML model training and validation. Extensive HiL or SiL testing could be performed as well.

Public Cloud

During the ADAS/AD development, Public Cloud infrastructure and services can be used in multiple situations. All or part of the data can be copied to the cloud for long-term storage by a cloud-based object storage service. Public cloud providers offer hundreds of services that can be used to implement parts of the ADAS/AD data pipeline. Data is shared with other service providers for services such as manual or automatic data labeling. The Public Cloud is best positioned for delivering centralized infrastructure and data management services. It may host a global data catalog as well as a global infrastructure monitoring dashboard.

QUANTUM ADAS/AD REFERENCE ARCHITECTURE

At the core of the Quantum ADAS/AD Reference Architecture sits the Quantum StorNext File System.

Quantum StorNext File System Overview

The Quantum StorNext File System is a state-of-the-art, high-performance, scalable, cost-effective storage solution that satisfies the demanding ADAS/AD storage requirements. It consists of a few different component types, as illustrated in Figure 3. At the core of the system are nodes that run StorNext data services. Although a single node installation is possible, a cluster of nodes is most commonly used for providing redundancy and scale. The StorNext policy engine data services manage the data placement across flash, HDD/SSD disk, tape, object storage, and cloud storage devices. Regardless of where the data resides in the system, it is always accessible in the namespace where it was written. The support of multiple, user-defined storage policies enables optimization of storage performance, cost, and protection.

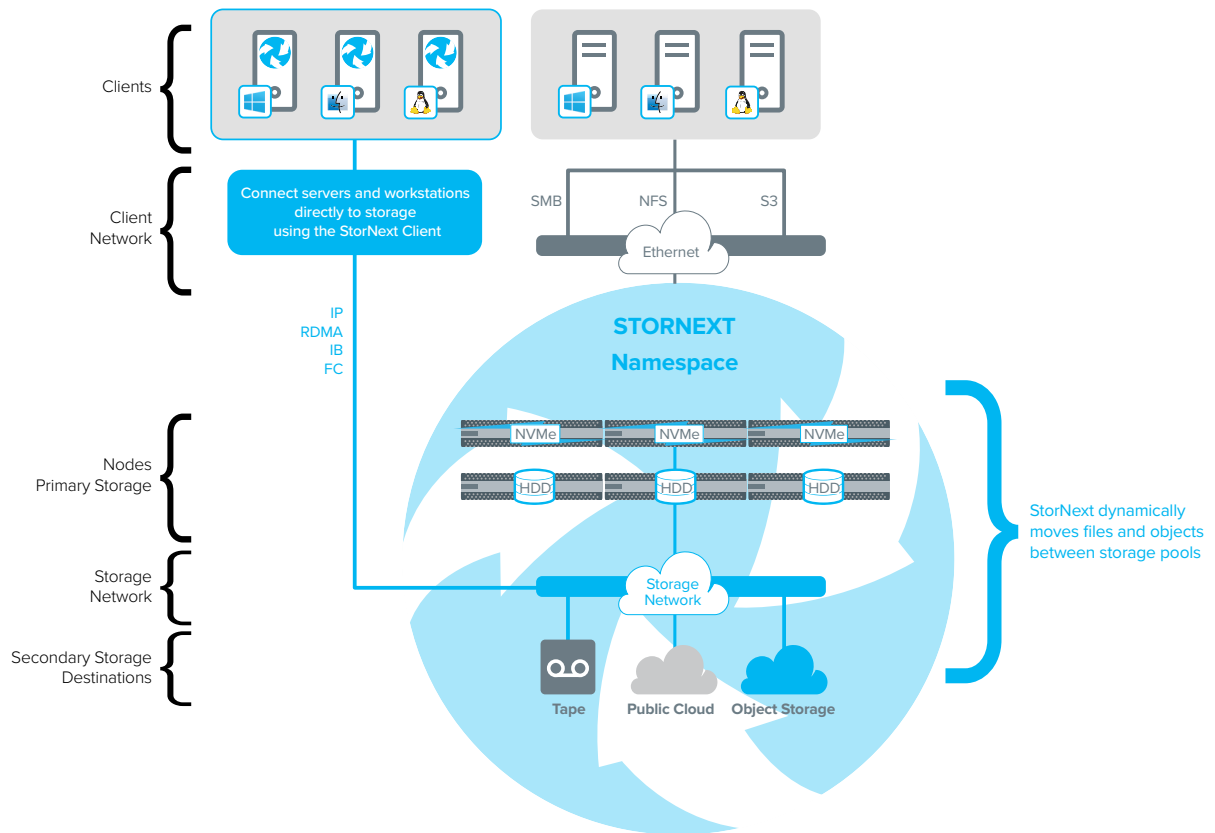


Figure 3 - High-Level StorNext File System Architecture

StorNext supports standard protocols such as NFS and SMB, but StorNext client software needs to be deployed on client machines in order to achieve the best performance. S3 protocol access is also supported where portions of the StorNext system could be set aside for use as on-premise or cloud-based object storage.

The Quantum StorNext File System architecture includes two types of networks. The Client Networks provide access for the user client machines to the file system. The Storage Networks connect the storage nodes and allow access to external storage destinations. Clients running the StorNext client software also connect directly to the Storage Network, which provides very high-performance and low latency access.

At the heart of the StorNext architecture is the StorNext File System, aka SNFS. The StorNext File System is an enterprise-grade parallel file system that provides superior resiliency, scalability, and control. Much of the power, scalability, and flexibility in a StorNext cluster is due to the capabilities of SNFS. A StorNext cluster may host one or multiple StorNext File Systems, depending on customer needs and objectives. It is well equipped to handle the most demanding ADAS/AD workloads. SNFS is device-agnostic and supports HDD, SSD, NVMe, or any other block storage types in the same file system while maintaining their performance characteristics.

SNFS brings an extra level of automation through the implementation of storage pools abstraction. Storage pools enable transparently moving data between different classes of primary storage such as NVMe flash and HDD. This movement is triggered manually by an administrator or automatically by a pre-defined policy. From the clients' perspective, the files always remain in their original locations in the StorNext namespace, regardless of which pool they physically reside in. The storage pool functionality brings high efficiency in cases where data sets have different storage performance needs at different points in their lifecycle. It simplifies the data moving to the storage type that best matches the current requirements. For example, an ML model training process requires very high bandwidth and low latency data access provided by NVMe storage, while HDD storage offers a cost-effective way to store the data while waiting to be processed.

On a capacity basis (\$/TB), NVMe is significantly more expensive than HDD storage and not applicable for extensive data sets. Storage pools enable leveraging a small amount of fast storage like NVMe when and where needed, bolstering it with more affordable bulk storage for parts of the workflow or data lifecycle that do not require extreme performance. Implementing storage pools also enables content separation between projects so that heavy use of files on one project does not negatively affect other projects' performance. Storage pools are also suitable for high-performance data ingest scenarios, where data lands on a pool of fast storage at capture time and is automatically aged off to less expensive, slower storage to make room for new content. Providing strong support for these types of scenarios is critical for deploying a high-performance yet cost-efficient ADAS/AD development platform.

StorNext also delivers the sustained performance required for Hardware-in-the-Loop (HiL) and Software-in-the-Loop (SiL) testing and removes the data access-related performance bottlenecks. The provided parallel data access shortens the time required for testing and maximizes equipment utilization. Utilizing StorNext client software achieves the highest performance for SiL and Model-in-the-Loop (MiL) test systems.

StorNext removes the biggest data access bottlenecks during the ML model training phase, where massive I/O bandwidth with extreme I/O parallelism is needed to feed data to the deep learning training cluster for processing.

For the ML model training, the ability to deal with randomized I/O is more critical than high IOPS performance for the storage system. Deep learning is random-read intensive, while the output is negligible, and therefore, a random-read optimized system is required.

By taking advantage of the SNFS advanced capabilities, the vast amount of ADAS/AD sensor data could be efficiently stored and managed in the multi-tier architecture shown in Figure 4.

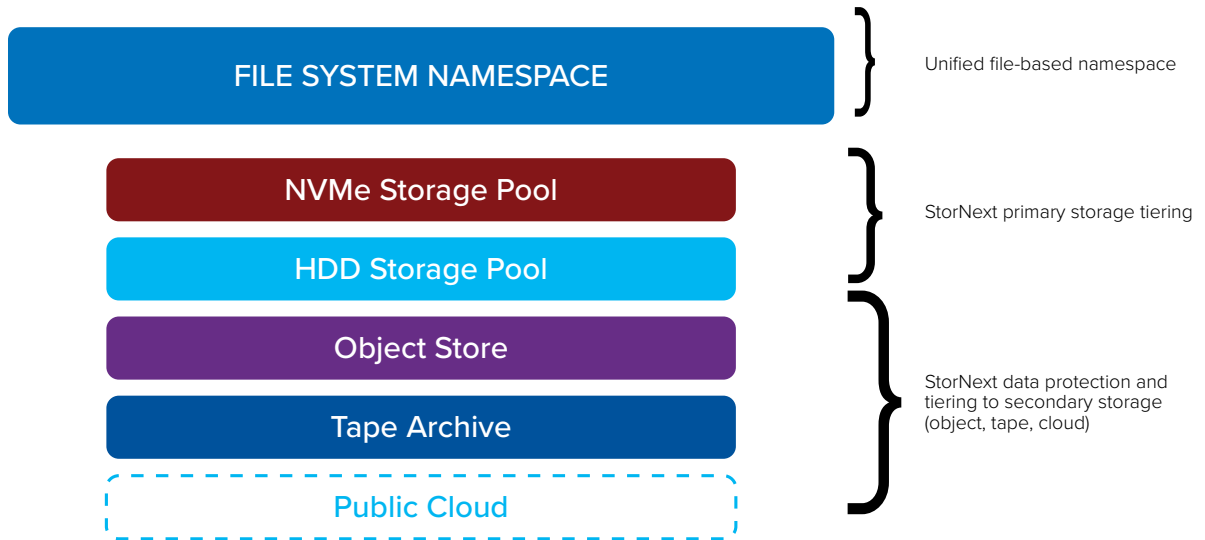


Figure 4 - StorNext Data Tiering

In addition to the file system, StorNext contains a robust selection of integrated data services.

Data Service	Description	Purpose
Data Migration	Move files between stripe groups	Resource Management
Quotas	Control sharing of file system capacity	Resource Management
QoS/Bandwidth Management	Control sharing of file system performance	Resource Management
FlexTier™	Transparently extend StorNext File Systems to secondary storage	Resource Management Data Protection Data Distribution
FlexSync™	Directory replication / synchronization	Data Protection Data Distribution
Import/Export	Import files from or export files to an external system such as Tape or Object Storage	Data Interchange Data Migration
Web Services API	Enables programmatic integration with outside applications such as asset managers, schedulers, etc.	Application Integration

Table 1 - Summary of StorNext Data Services

This rich selection of data services makes the implementation and the operation of a distributed ADAS/AD data lake possible and efficient.

Quantum ADAS/AD Storage Technology Stack

The Quantum portfolio of data management software and storage hardware appliances provides all storage components required to deploy an end-to-end, high-performance, flexible, cost-effective ADAS/AD development and production environment. The huge amount of data required for the ADAS/AD system development moves through multiple stages, each with very specific storage requirements. The Quantum technology stack meets these requirements at every stage, starting with the in-vehicle Data Acquisition and going through the Data Ingest, Data Preparation, ML Model Development, Simulation, Validation, and Archiving processes. The Quantum ADAS/AD storage reference architecture is implemented using the appliances listed in Table 2.

Data Function	Physical Location	Quantum Product
Data Capture	Test Vehicle	R-Series
Data Ingestion	Remote Site	R-Series
File System	Remote Site, Development Center	StorNext 7.0, StorNext Appliances
Performance Tier	Remote Site, Development Center	F-Series (NVMe)
Capacity Tier	Remote Site, Development Center	QXS (HDD/SDD)
Object Storage	Remote Site, Development Center	ActiveScale
Archive Tier	Remote Site, Development Center	Scalar (Tape Archive)

Table 2 - Quantum ADAS/AD Storage Technology Stack

Quantum In-Vehicle Storage Appliance

Quantum developed the R6000 appliance specifically for high-speed data capture in a moving vehicle. It provides a large data storage capacity necessary for the in-vehicle logger to store the collected sensor data for an extended period of time. It has a small form factor well suited for the trunk of a self-driving car, and it is designed to withstand the demands of a rugged environment.

R6000 provides two 100 GbE interfaces for high-performance data access. The data is stored on NVMe SSD drives packaged in a removable magazine. There are multiple types of storage magazines with up to 122 TB storage capacity per magazine. The magazines are field-replaceable and easy to transport for the data to be offloaded at the data center. Multiple magazines could be used if necessary during data collection to extend test data capacity and minimize car downtime.

R6000 is purpose-built for high availability and reliability. RAID support is implemented to ensure high performance and reliability.



Figure 5 - Quantum R6000 Ruggedized Storage Appliance

Remote Site Reference Architecture

Figure 6 illustrates the Remote Site storage infrastructure architecture built with Quantum storage appliances and software components. The architecture includes data ingestion, performance, capacity, and archival storage tiers.

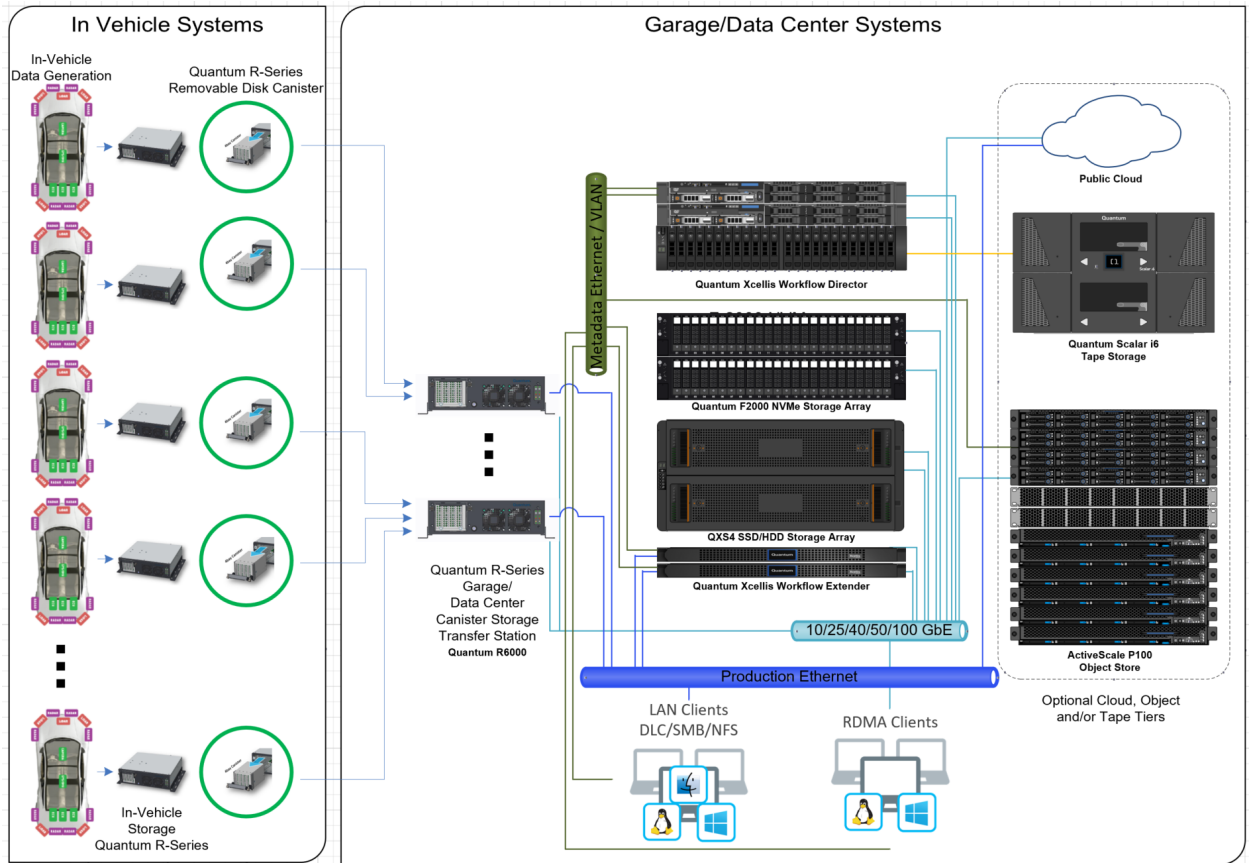


Figure 6 - Remote Site Reference Architecture

Data Function	Quantum Product	Number	Max Capacity
Data Capture	R6000	1 per car	120 TB
Data Ingestion	R6000	2	120 TB
File System	StorNext Metadata Controller	1	Metadata
	StorNext Gateway Nodes	2	Gateway
Performance Tier	F2000	2	368.6 TB
Capacity Tier	QXS-4	1	5.4 PB
Object Storage	ActiveScale P100	1	27.2 PB
Archive Tier	Scalar i6	2	9.6 PB (native)
			24 (compressed)

Table 3 - Quantum Storage Products (by Tier)

Data Ingestion Tier

R6000 appliance is installed as a data ingestion station at the data center. Once a storage magazine is plugged into the ingestion station, the data is automatically uploaded into the Quantum StorNext File System and ready for further processing.

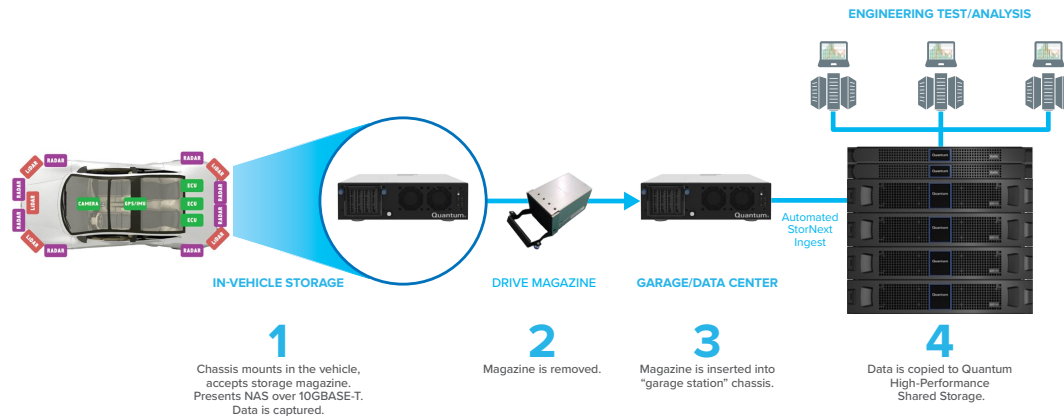


Figure 7 - Data Ingestion Process

R6000 has two 100 GbE interface for high-performance data ingest. It is a standard NAS device. It supports data access speeds of up to 10 GB/sec.

The data ingestion is performed on a separate network, and it is not interfering with the rest of the data traffic. Multiple data ingestion stations operating in parallel could be configured at each Remote Site.

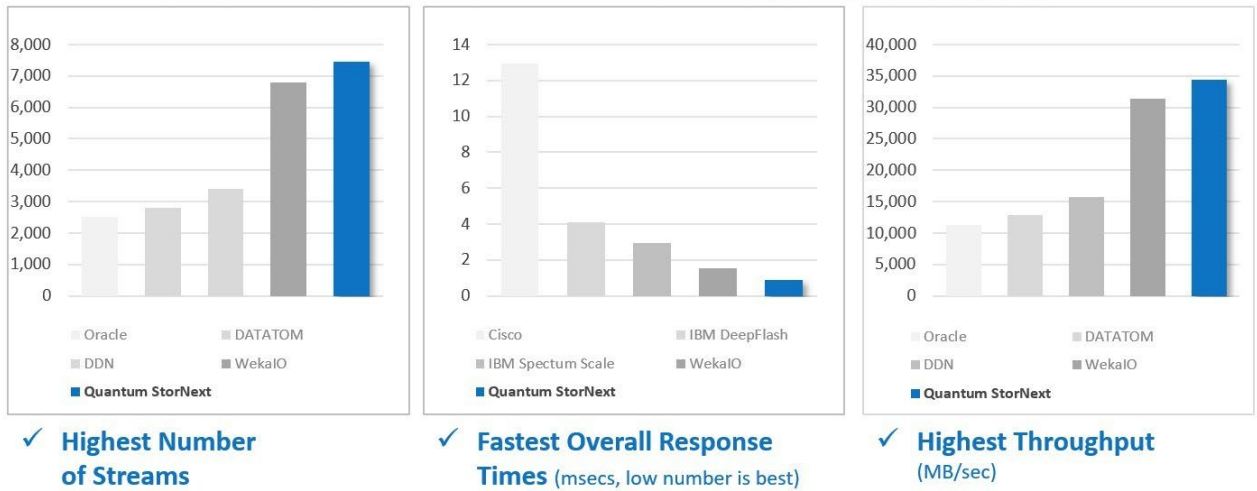
StorNext File System

At the core of the Remote Site storage platform sits the StorNext File System. StorNext supports complete data lifecycle management from data ingestion to long-term data archival. It provides continuous data access as data moves across storage media: NVMe, SSD, HDD, tape, cloud, and performance tiers (hot, warm, cold) and storage formats (file, object). For performance-sensitive applications, StorNext leverages iSER/RDMA or IB direct to storage connections. Thousands of parallel data streams are supported to meet the requirements of the specific applications. Once data is no longer required for high-speed access, it is automatically placed on lower-cost storage tiers, following software-defined policies. The data could be moved or copied to on-prem highly scalable, reliable, and cost-efficient Quantum ActiveScale Object Storage, Quantum LTO tape libraries, or public cloud S3 object storage.

In the reference architecture described in Figure 6, the StorNext File System is configured and deployed on the Quantum Xcellis® storage appliances. The Xcellis Workflow Director appliance is hosting the StorNext Metadata Controller ("MDC") software. The Xcellis Workflow Extender appliance is deployed to expand the number of supported client connections and bandwidth. These appliances are custom built to maximize StorNext's performance and connectivity capabilities.

StorNext 7 outperforms all other file systems across video SPEC-SFS 2014 metrics using 57% fewer storage nodes than the next nearest competitor.

StorNext demonstrates best value in terms of cost per concurrent streams



It set new performance records for:

- The highest aggregate throughput (34,391 MB/sec)
- The lowest latency (0.9 milliseconds overall response time)
- The highest number of concurrent streams (7,450 streams)

Performance Tier

The ADAS/AD reference architecture performance storage tier is implemented with Quantum F2000 NVMe appliances. F2000 is an ultra-high-performance storage system based on NVMe-attached Flash modules. Designed specifically for use with Quantum Xcellis appliances and StorNext software, F2000 excels at handling a large number of simultaneous streams of high-resolution video or similar high-bandwidth unstructured data. The F2000 is available in three capacity configurations (46 TB, 92 TB, 184 TB) with Dual-port 100 Gb Ethernet interfaces in a 2U rack-mounted chassis. Multiple units are deployed for a larger file system capacity, with the aggregation of capacity and performance provided by StorNext. The StorNext architecture delivers linear horizontal scaling as more F2000 appliances are deployed. A single appliance provides 23 GB/sec, with a second appliance, the performance increases to 46 GB/sec, and so on. Quantum Client software needs to be deployed on the client computers to achieve the best data access speeds. More details about the client software are provided below in the Clients and Connectivity section.

ML model training, HiL and SiL testing, and some of the ADAS/AD development system's data preparation functionality rely heavily on the Performance Tier capabilities. The reliable support of thousands of parallel data streams is critical for the development process. NVMe appliances play an essential role in training a machine learning model by providing high bandwidth and low latency data access. Even if the training itself runs from RAM, the memory is fed from non-volatile storage. Paging out old training data and bringing in the new data is done as fast as possible to keep the GPUs loaded. Checkpoint setting also benefits from NVMe technology. When the ML training process is long, the system can choose to save a snapshot of the memory into non-volatile storage to allow a restart from that snapshot in case of a crash.

Capacity Tier

The capacity storage tier is designed to support petabytes of data in a cost-efficient way. It is implemented by utilizing a Quantum QXS-4 series appliance. Designed and engineered from the ground up to support demanding video and other unstructured data processing workflows, QXS is tuned to deliver high levels of sequential I/O performance, with over 7 GB/s and 320K IOPs per chassis. It offers up to 5.4 PB of capacity per system implemented with SSD and 10K HDD drives. The QXS-4 appliance is connected to the storage network over 10G Ethernet (iSCSI) interface.

With this implementation, the Capacity Tier provides a cost-effective way for storing the extremely large ADAS/AD data sets. It also provides sufficient performance for data preparation operations.

Object Storage

An object storage system is the most efficient way to store and process the vast data sets used by an ADAS/AD platform. One of the main components of the Quantum ADAS/AD reference architecture for the Remote Sites is an enterprise object storage tier implemented with the Quantum ActiveScale P100 appliance. ActiveScale P100 supports configurations from 864 TB to 27.2 PB (raw) capacity and 600 million to 16.2 billion objects. It supports a standard S3-compliant object interface and an NFSv3 interface, and it is easily expandable when extra storage capacity is required. The S3 interface enables workloads and applications to move seamlessly between public cloud and on-prem infrastructure. ActiveScale provides strong data protection by implementing Dynamic Data Placement and Repair functionality by placing objects across separate storage units.

The object storage technology offers the greatest level of scalability and a simplified access protocol via HTTP(S). Object stores are capable of supporting multiple concurrent I/O requests, but they generally can't offer the throughput provided by file or block storage devices.

Archive Tier

The Quantum ADAS/AD platform provides long-term storage facilities for digital media and image archiving implemented with Quantum Scalar i6 tape appliances. These appliances offer the best storage density at the lowest cost per TB of storage on the market. LTO tape is well suited to store the petabyte-scale ADAS/AD data sets at a fraction of the cost of public cloud and other cold storage solutions.

Quantum Scalar i6 appliance offers a unique solution to data center backup and long-term storage. Users can start with a single, 6U library unit and grow to fill an entire 48U rack with a single, integrated library that can store up to 24 petabytes.

For the ADAS/AD Remote Site storage needs, two Scalar i6 modules provide up to 6.0 PB of storage in a 12U form factor. They are connected with an 8 Gbps FC interface for data access and a 1 GbE management interface.

Clients and Connectivity

StorNext File System is accessible by client computers running Windows, macOS, and Linux operating systems. The clients connect to the StorNext cluster by using either OS native or separately installed client software. Each client connection method has distinct characteristics, as outlined in Table 4. StorNext simultaneously serves a mix of clients using different connectivity methods based on their specific needs and capabilities.

Client Connection via	Performance	Path to Storage	Client Software	Data Network
NAS (NFS/SMB)	Good	NAS Gateway	Built into OS	Standard Ethernet
StorNext LAN Client	High	StorNext Proxy Gateway	Installed Separately*	Standard Ethernet
StorNext SAN Client	Highest	Direct	Installed Separately*	Fibre Channel, IB, iSCSI, RDMA Ethernet
S3 Client Software	Good	S3 Gateway	Installed Separately	Standard Ethernet

**macOS includes Xsan, a version of StorNext client software. Windows & Linux require StorNext client software installation*

Table 4 - StorNext Client Summary

To achieve the best possible storage performance, clients need to use the StorNext client software package. Apple macOS computers may take advantage of the Xsan client built into macOS for this purpose. Windows and Linux require the installation of StorNext client software.

StorNext client software mounted StorNext File Systems appear to the client machine as local file systems, not a remote NAS mount. It provides much higher performance than NAS mounts, especially for large files and streams.

Development Center

The storage infrastructure deployed at the Development Center provides sufficient scale and performance to enable the effecting execution of all ADAS/AD development activities. It is capable of storing hundreds of petabytes of data in the multi-tier architecture described previously in this document.

The Development Center storage infrastructure is a scale-up and scale-out implementation of the Remote Site storage infrastructure. Each storage tier and the network infrastructure are scaled appropriately to meet the significantly increased performance demands.

The ADAS/AD Development Center storage reference architecture is displayed in Figure 9.

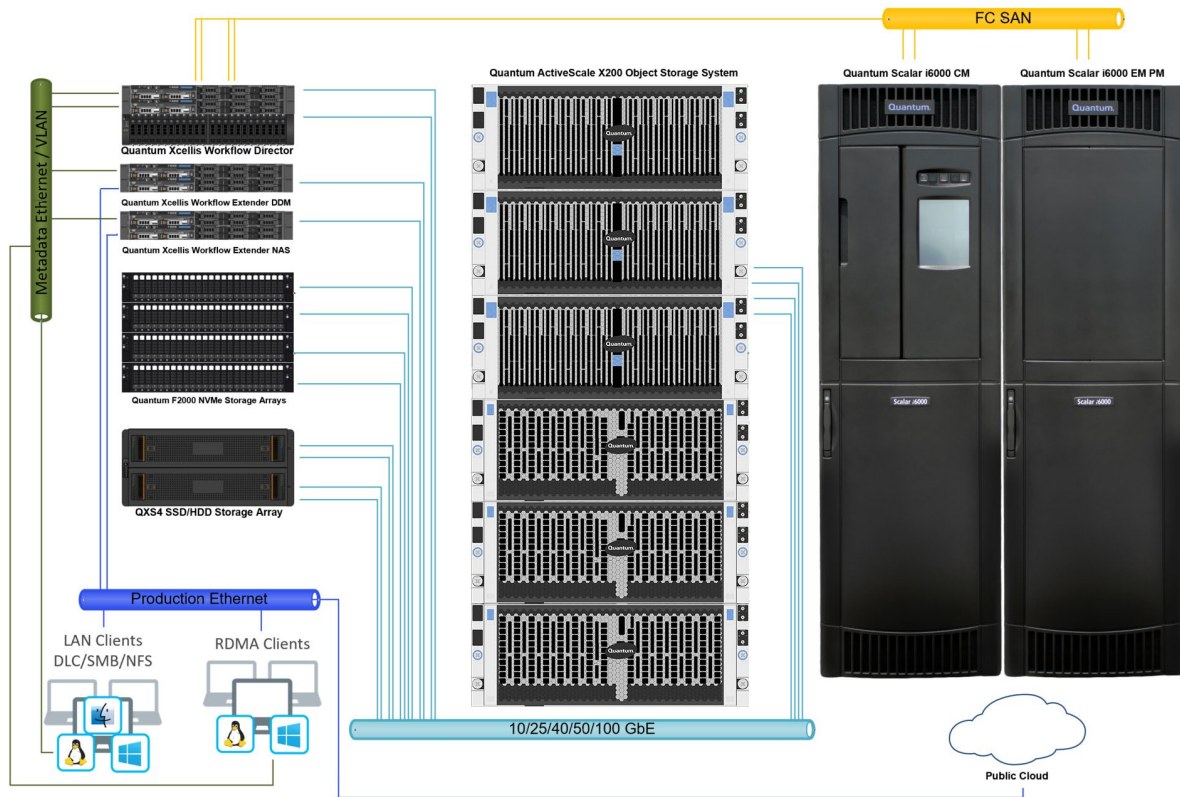


Figure 9 - Development Center

Data Function	Quantum Product	Number	Max Capacity (raw)
File System	StorNext Metadata Controller	1	Metadata
	StorNext Gateway Nodes	2	Gateway
Performance Tier	F2000	4	737.2 TB
Capacity Tier	QXS-4	2	10.8 PB
Object Storage	ActiveScale X200	1	Unlimited
Archive Tier	Scalar i6000	1	144 PB (native)
			360 PB (compressed)

Table 5 - Quantum Storage Products (by Tier)

Object Storage

The object storage infrastructure at the ADAS/AD Development Center is implemented with Quantum ActiveScale X200 appliances. ActiveScale object storage provides a new, innovative approach to creating a simple, 'always-on' data repository that scales with the extreme data durability, accessibility, and security required of petabyte-scale growth. With ActiveScale, performance, capacity, and scaling are seamless. As new nodes are added, network and computing resources come online immediately to load balance data requests and system tasks. New storage is added to the capacity pool and write activity is optimally distributed across storage resources using Dynamic Data Placement heuristics. ActiveScale's unique architecture achieves high performance with an extremely cost-effective approach. X200 offers base unit capacity at 1.62 PB, 3.24 PB, and 4.86 PB, and scales infinitely by deploying multiple base units.

Archive Tier

The Development Center data archival storage tier is implemented by using Quantum Scalar i6000 enterprise tape library. Scalar i6000 is the Enterprise library with the industry's best combination of capacity, security, and economy. A single unit can store up to 360 PB of data, and it scales using capacity-on-demand growth and slot-based pricing, achieving the densest and most economical storage. Its high-density design does not compromise robotic performance or reliability while providing up to 23.4 PB in a single, standard 19-inch rack footprint.

Clients and Connectivity

Client workstations and servers can connect to the Development Center storage infrastructure using all mechanisms and protocols described above in the Remote Site section. The Development Center hosts deep learning training clusters that could incorporate hundreds of servers with 4 to 8 GPUs per server. Quantum tiered storage infrastructure eliminates I/O bottlenecks and supports 100% GPU utilization.

CONCLUSION

ADAS/AD developers are faced with the immense challenge of dealing with hundreds and even thousands of petabytes of unstructured data collected and processed in multiple locations today. With the evolution of sensor technology, the amount of data quickly multiplies.

Applying decades of unstructured data management experience, Quantum helps customers to streamline their data management, development, and validation processes. By taking a holistic approach and creating an end-to-end architecture that addresses the specific needs of every stage of the ADAS/AD development, Quantum brings a strongly differentiated solution set to this market.

Industry-proven high-performance, cloud-scale, reliable Quantum technology enables ADAS/AD developers to unleash the power of their data. Quantum strongly cares about having the data well protected against ransom attacks which becomes more critical every day. Quantum recognizes that ADAS/AD collected data becomes one of the most valuable company assets and focuses on unlocking the data business value.

Quantum is constantly expanding the ecosystem of technology alliance partners. Below is a sample list of partners Quantum is working closely with to bring comprehensive customer solutions.



Related Documents & Links

More information on Quantum storage, StorNext software, and Xcellis storage appliances, including compatibility and detailed information on features, is available via the following links.

F-Series: www.quantum.com/f-series

Scalar: www.quantum.com/tape

StorNext 7: www.quantum.com/stornext

Documentation Portal: www.quantum.com/documentation

Xcellis: www.quantum.com/xcellis

Quantum StorNext Sets New Performance Records for Video Workloads: www.prnewswire.com/news-releases/quantum-stornext-sets-new-performance-records-for-video-workloads-301235534.html

R-Series: www.quantum.com/r-series

ActiveScale: www.quantum.com/object-storage



Quantum®

Quantum technology, software, and services provide the solutions that today's organizations need to make video and other unstructured data smarter – so their data works for them and not the other way around. With over 40 years of innovation, Quantum's end-to-end platform is uniquely equipped to orchestrate, protect, and enrich data across its lifecycle, providing enhanced intelligence and actionable insights. Leading organizations in cloud services, entertainment, government, research, education, transportation, and enterprise IT trust Quantum to bring their data to life, because data makes life better, safer, and smarter. Quantum is listed on Nasdaq (QMCO) and the Russell 2000® Index. For more information visit www.quantum.com.

www.quantum.com | 800-677-6268