PURESTORAGE®
Uncomplicate Data Storage, Forever

Tech Mahindra

TECHNICAL WHITE PAPER

Winning the ADAS Race with Tech Mahindra and Pure Storage®

Why innovation in ADAS and AD programs is fueled by the power of data.

Contents

Introduction 3

Development Programs Typically Follow Five Key Stages: 3

Challenges in Accelerating ADAS & AD Programs 4

Key Challenges Faced by R&D and Infrastructure Teams in ADAS Development 4

Partners in Automotive Innovation 5

The Technology Driving Transformation 6

High-Level ADAS Framework 6

ADAS Data Center Architecture Framework 6

Autonomous Drive Phases 7

Unified Fast File and Object Storage 7

Data Collection Center 8

Eliminating Bottlenecks 9

Workload Consolidation for a Future Proof Architecture 10

Future Proof ADAS Solutions for Modern Containerized Applications 10

Driving Outcomes for OEMs and Tier 1s 11

Use Case: Shaping the Future for a European Luxury Car Manufacturer 11

Challenges 11

Objective 11

Outcomes 11

Conclusion 12

12

Introduction

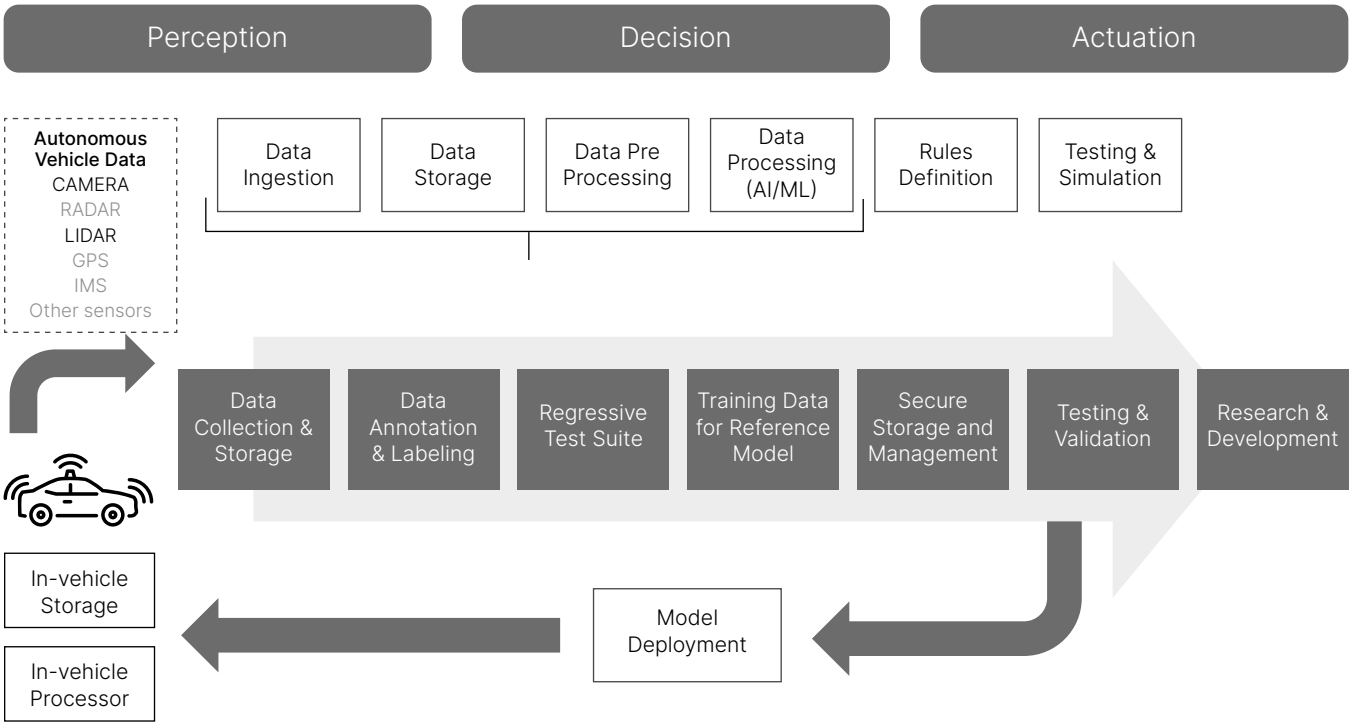
The Global ADAS Market size is projected to grow from USD 27.2 billion in 2021 to USD 74.9 billion by 2030, at a CAGR of 11.9%[1].

Automotive Original Equipment Manufacturers (OEM) and Tier 1 suppliers are in a race to deliver ADAS and AD solutions earlier than their competitors. At least 30% of their investment is being made in data management and hybrid data centers[2].

In this white paper, we explore how suppliers can optimize their infrastructure and the way they ingest and manage growing volumes of data that ADAS and AD solutions rely on, while also under pressure to assign higher CAPEX budgets.

Development Programs Typically Follow Five Key Stages:

1. Build and run test vehicles using test path and data collection.
2. Pre-process and manage the data.
3. Undertake vehicle training and inference development.
4. Test the software and hardware modules.
5. Complete vehicle integration and validation.



1 Source: https://www.marketsandmarkets.com/Market-Reports/driver-assistance-systems-market-1201.html
2 Source: Tech Mahindra internal research

Challenges in Accelerating ADAS & AD Programs

In OEMs and Tier 1s, ADAS development programs are managed by the engineering teams, and the data management and infrastructure is managed by the CIO. This creates a number of challenges for R&D and infrastructure teams in ADAS development:

Key Challenges Faced by R&D and Infrastructure Teams in ADAS Development

Challenges	Needs
Higher CAPEX investment	CAPEX to OPEX
Multi-site & multi-party	Web-frontends, DMS, Data Moving instances (SiL, HiL, HPC)
Country specific regulations	Domain expertise on country specific regulations and data management
Optimize On-Prem and Cloud environments	Data Sizing, Data Architecture expertise
Iterative Process (train, test, re-train, improve)	Integration along development chain
Data Growth	Powerful DB, Metadata management, storage systems scalability
Complex T&V, workflow and CI/CD	Fully automated toolchain

Taking Camera-Based Adas Features as an Example, Here Are Some of the Challenges:

Object detection and classification is a computer vision technique that recognizes images, and it's the most valuable technology in enabling such systems. In modern cars, AI systems read data from cameras and other sensors, and then use pre-trained models to take action. These systems use complex neural networks to read the incoming video and sensor feed and generate inference data.

Developing these pre-trained models creates a challenge for compute and storage, as training a system that is accurate 99.99999% of the time requires a significant amount of video data. Multiple video feeds may be required from different angles and locations to compensate for a moving environment, weather factors can also alter the appearance—the training algorithm must account for such issues. In addition, there could be supplemental data in the form of radar, infrared images or light detection and ranging (LIDAR) to improve the accuracy of the model. So the compute engine needs to be able to handle high volumes of data from different sensors.

The training stage is currently performed on a server in a data center. The data is collected from a car in motion and the neural network model is trained for days or months. It's then downloaded to the vehicle periodically. Storage is a critical element of the infrastructure for both training and inference because the sheer volume of data in a vehicle could overwhelm a storage system. For instance 360° cameras could require up to 17 video streams to be processed, stored, and then fed into the training system. And that volume increases further when considering additional data from elsewhere, such as radar sensors.

In vehicles, storage requirements are constrained by real-time decision-making needs. Automatic emergency braking (AEB), for example, must calculate the results within a permissible time, depending on the speed of vehicle, to avoid a collision. The storage system onboard has to cope with multitudes of cameras and sensor data processing in real-time.

Manufacturers are therefore seeking ways to enable training in real-time, so that vehicles can be updated on any issues immediately. For instance, a pothole could be identified quickly by a training system and the information passed on to all vehicles coming up to it to avoid an accident.

Partners in Automotive Innovation

In partnership, Tech Mahindra and Pure Storage can help address these challenges and power the development of ADAS and AD programs.

With its unique end-to-end engineering capabilities, Tech Mahindra has more than 18,000 associates in the automotive sector and partners within 80% of the top ten automotive OEMs globally. It delivers continuous innovation to OEM and Tier 1 customers in the automotive sector, with cutting-edge solutions and turnkey product design and development programs. As well as smart solutions for infotainment, customer experience (CX), factory and aftermarket domains.

Pure Storage is supporting the next era of human advancement with leading-edge AI and data analytics solutions. Providing the world's first full-stack, AI-ready infrastructure—powered by NVIDIA and FlashBlade®, it helps customers reduce time to insight and drive AI at scale.

By combining innovative technology and industry knowledge, Tech Mahindra and Pure Storage can help customers drive key outcomes, including workload consolidation, improved TCO, reduced emissions, simplified architecture, shorter time to market, enhanced user experience, improved quality and global presence. Let's explore how.

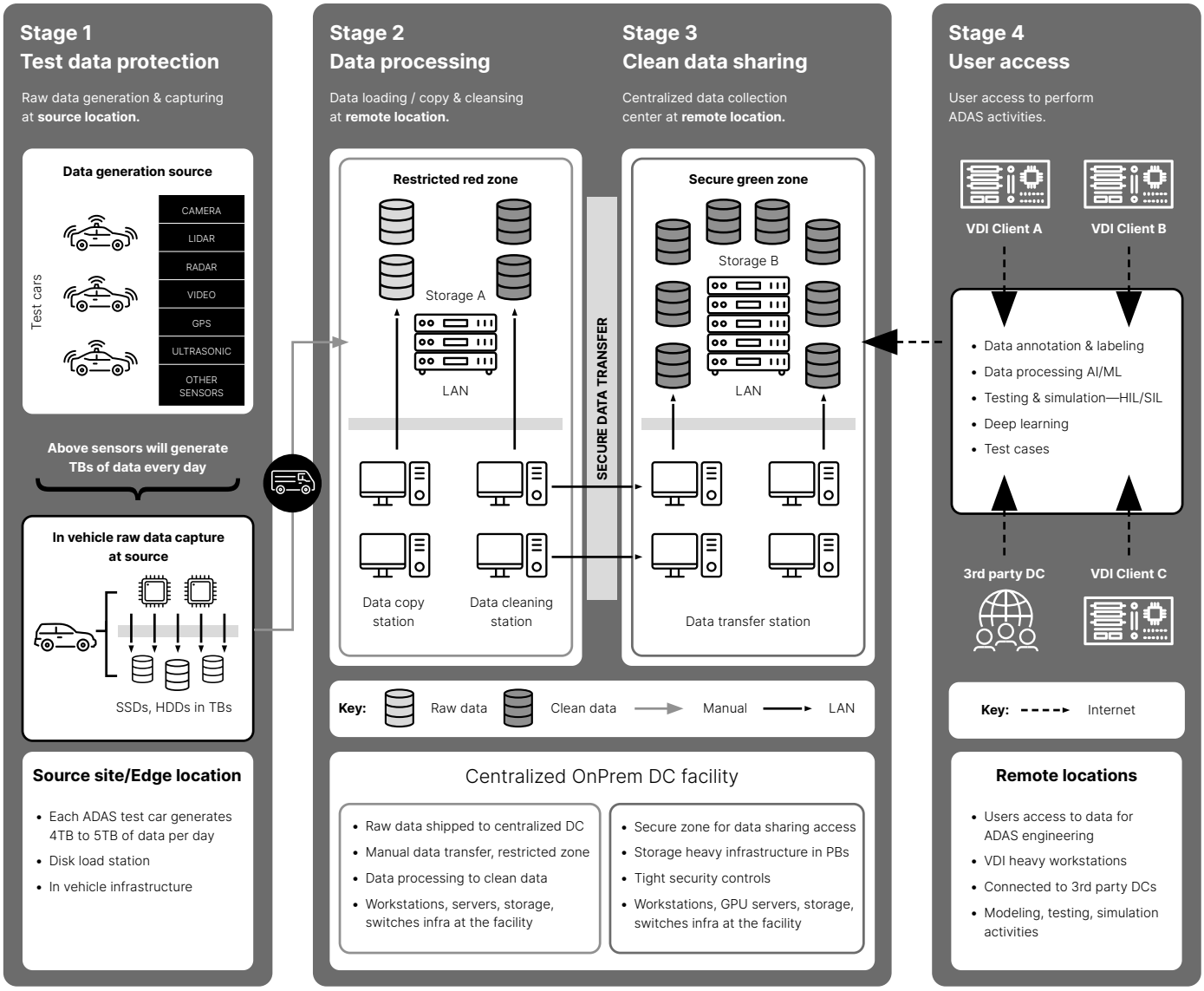
- Edge data centers and ingestion stations need to be established and managed globally. We have a global presence, so can manage local regulatory requirements in different countries and acquire any necessary licenses.
- Our knowledge and capabilities bridge the gap between disconnected IT and engineering teams.
- We help customers reduce time to market by providing the latest technology.
- Building their own infrastructure means customers face the risk of incorrect sizing and incompatible product selection. We deliver it as a service, so these risks are eliminated allowing the customer to focus on vehicle development not IT infrastructure.
- We implement multiple edge platforms, which enable engineers to remotely access and analyse sensor and video data to develop their autonomous driving systems.
- The process is simplified and disruption is reduced, so developers and analysts can work more productively.

The Technology Driving Transformation

High-Level ADAS Framework

Tech Mahindra brings together end to end services from AI and ML simulation, app development and testing and comprehensive simulation all integrated within a vehicle.

ADAS Data Center Architecture Framework



To enable feature development and training, data is ingested from a vehicle, stored for labeling and then processed through a test suite. The defined rules are applied, and testing and validation is carried out before deploying the model back to the vehicle. And that cycle is repeated.

Each of these steps require different software stacks, and traditionally separate silos of storage, with data being copied and transformed between each silo.

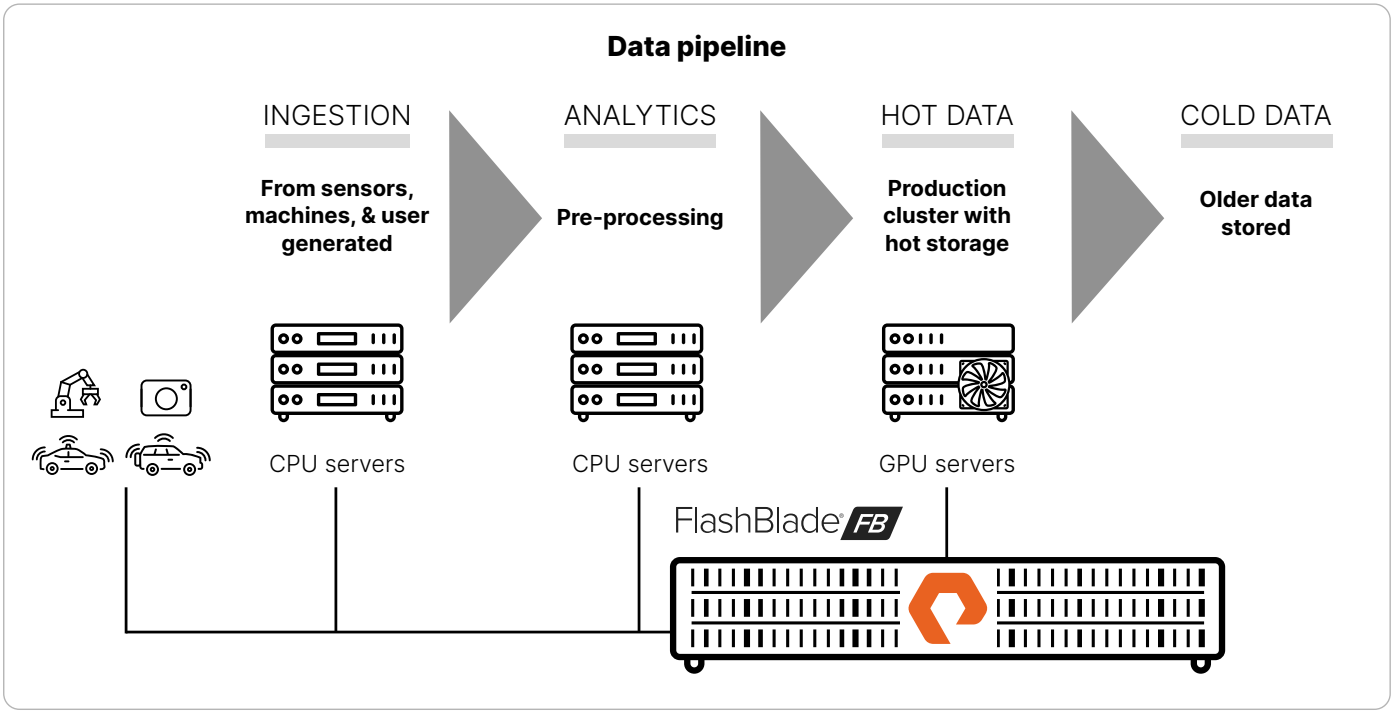
Autonomous Drive Phases

Data storage characteristics

	GET DATA Ingest	**GET DATA** Refine	**TRAIN & ADJUST** Training	**DEPLOY & TEST** Testing
	RAW	LABELING	GPU DNN TRAINING	SIMULATION
Media type	Mechanical/Hybrid	Mechanical/Hybrid	NAND/Performance Optimized	Flash/SSD
Processing mode	Batch	Batch	Real-time	Batch
I/O type	Sequential	Sequential	Random & Sequential	Sequential
Architecture	Scale-out	Scale-out	Scale-out/ Massively parallel	Parallel

The data hub philosophy of Pure Storage eliminates data hub philosophy eliminates silos and drastically reduces data movement from the ingest stage through to training. It can handle mixed access patterns, sequential or random, and work with any file size. This parallel architecture ensures workloads are always serviced with low latency.

Unified Fast File and Object Storage

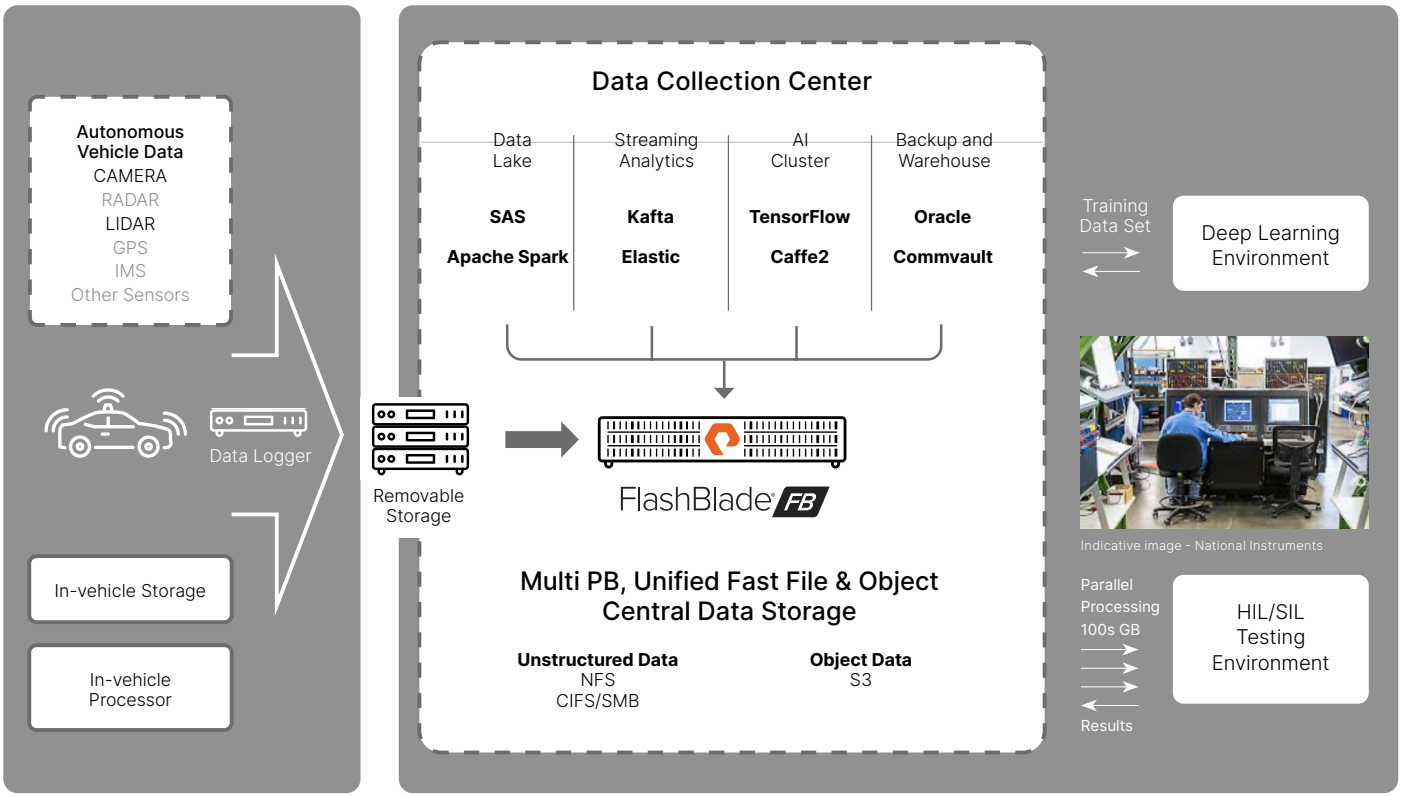


A traditional ADAS infrastructure is often hand-built with multiple components, delivered by different vendors. Not only does this bring a number of challenges, it is also not sustainable and hinders innovations as needs constantly change and environments grow.

With Tech Mahindra and Pure Storage, the data hub seamlessly integrates with the ADAS framework, removing silos of storage and the need to copy data between different zones. A single FlashBlade data hub can simultaneously handle data ingestion, performing cleansing and training without the need to copy data between various silos.

Data Collection Center

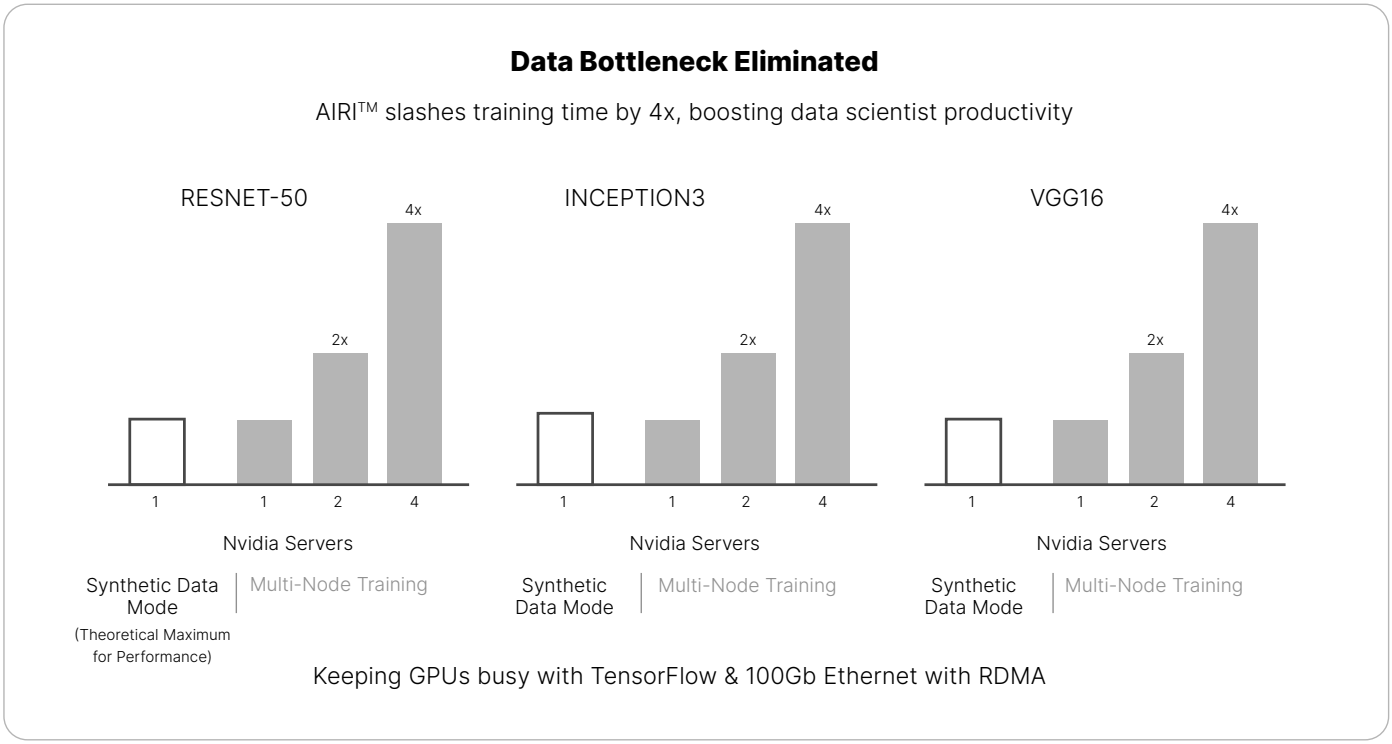
The overview below, demonstrates our perspective on vehicle training, simulation and testing:



Multiple application types can be run, from data lakes to data warehouses, on a single unified fast file and object platform, which allows everything from ingest to training simulation and testing to be completed on a centralized platform. It simplifies deployment, adapts to change, increases quality, and reduces time to market.

Efficiency is improved by eliminating multiple copies of data, distributed across separate DAS cluster nodes. By consolidating them into one data hub, the deployment of new clusters and future applications is simplified and can be achieved on-demand in minutes.

Eliminating Bottlenecks



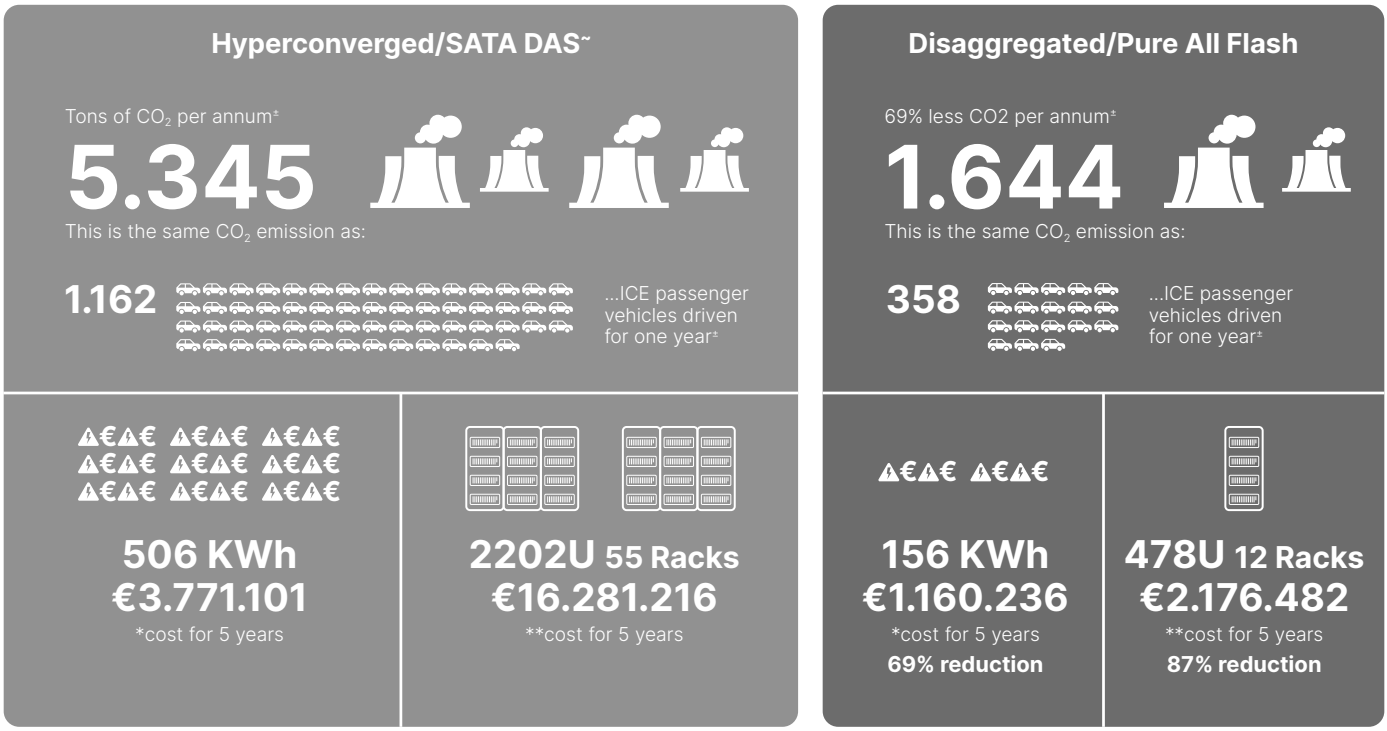
Data scientists generally run AI workloads from a single server, with multi-node scaling only used by AI experts presenting data, typically leaders in their field. Tech Mahindra's and Pure Storage's ADAS framework removes the barrier, enabling any data scientist to complete multi-node training.

In the chart above, three popular neural networks can be seen. It shows two patterns:

- That GPUs are kept at their maximum throughput.
- Pure Storage's unified fast file and object platform allows jobs to scale linearly as more systems are added to a single job.

Data needs to be processed from storage, over networks, through CPUs, across peripheral buses, and into the GPU. In DIY systems, the input pipeline can easily stall the GPU, slowing down the entire training process. With Pure Storage's data hub and Tech Mahindra's ADAS framework, the GPUs are continuously fed with data.

Workload Consolidation for a Future Proof Architecture



ILLUSTRATIVE CALCULATION BASED UPON CURRENT PROPOSED CONFIGURATION AND IS SUBJECT TO CHANGE AND CONFIRMATION OF COMPETITIVE CONFIGURATIONS

* Assumes DC PUE (Power Usage Efficiency) of 1.7 European average, source: https://www.opencompute.org/documents/the-current-state-of-data-center-energy-efficiency-in-europe-ocp-white-paper

**KWh uses market standard pricing - €0.10/kWh.

±CO_2 and vehicle Calculation Source: https://www.epa.gov/energy/greenhouse-gas-equivalencies-calculator.

~Assumes 4TB SATA disks, replication factor 3.

This approach also enables automotive manufacturers meet their net zero goals with significant environmental benefits. This example shows:

- A reduction of 69% in power and cooling.
- 87% savings in rack space vs using traditional DAS clusters with open-source software.

Future Proof ADAS Solutions for Modern Containerized Applications

Disaggregation with fast S3 has driven an architectural shift. Traditionally, data pipelines and analytics applications were built in a way that leverages direct attached storage, originating from HDFS and GFS. Advancements in technology mean this model no longer makes the same compelling case it did two decades ago.

These disaggregated models continue to gain momentum with key benefits such as operational simplicity as the data pipelines get larger and larger. And the separate stateless compute and storage offers a denser, more environmentally friendly architecture.

As needs evolve and complexity increases, the demand for agile containerized software environments grows. Pure Storage meets this demand with a Kubernetes data management layer that delivers persistent storage to modern containerized applications. Portworx® provides enterprise data services alongside migration and DR, allowing applications to utilize the most appropriate storage, whether it's DAS, block or cloud.

Driving Outcomes for OEMs and Tier 1s

Partnering with Tech Mahindra and Pure Storage to develop ADAS and AD programs, enables customers to deliver key business outcomes:

- **Simplified architecture**—Minimizes data movement, improving the management process and scaling a single namespace from 100TB to multi-petabyte scale, with consistent performance and enhanced user experience.
- **Workload consolidation**—Eliminates silos by consolidating multiple workloads into a unified data storage layer.
- **Boosts innovation with better insights**—Removes storage bottlenecks and allows faster analytics, which improves quality through more training iterations and reduces time to market.
- **Improved TCO**—Reduces management overheads through simplification and elimination of up-front investment.
- **Global presence**—Tech Mahindra and Pure Storage's global presence means all local regulatory requirements are met.
- **Environmental sustainability**—Supports customers net zero goals through reduction of power, cooling and data center footprint.

Use Case: Shaping the Future for a European Luxury Car Manufacturer

Challenges

- Needed a platform that enables engineers to remotely access sensor and video data to develop autonomous driving systems.
- Required hybrid cloud mobility and reduced complexity.
- Scalability and agility were critical.
- User experience needed to be a key consideration.

Objective

- To deliver a new scalable and agile subscription-based storage service.

Outcomes

- Offered balance sheet benefits from day one.
- Simplifies the supplier ecosystem and automated processes.
- Provides a fully managed, Evergreen-enabled consumption-based hybrid cloud.
- Accelerates and de-risks workload migration, and CMO decommissioning.
- Guarantees 99.999% data availability.
- Improves protection and performance.
- Flexibility of usage and forecasting.
- Platform for continuous innovation that evolves in line with changing usage requirements and demand.

Conclusion

Together, Tech Mahindra and Pure Storage enable OEMs and Tier 1s to innovate in ADAS and AD with a high performing, highly scalable future proof infrastructure. Through a strategic partnership with two leading innovators, customers benefit from:

- Core competencies to focus on development and innovation, and faster time to market, while their IT is in the capable hands of Tech Mahindra.
- Redeployed infrastructure budget for better product innovation, and the opportunity to consume infrastructure in a pay-as-you-grow model.
- Centralized management of all data across multiple locations.

Together, we provide the vision, best-in-class solutions and capabilities to unlock the power of data and fuel successful ADAS and AD journeys.

Contact your local Tech Mahindra or Pure Storage representative

©2023 Pure Storage, Inc. All rights reserved. Pure Storage, the P logo mark, FlashArray, FlashBlade, and FlashBlade//S are trademarks or registered trademarks of Pure Storage, Inc. All other names may be trademarks of their respective owners.

PS2437-01-en 05/23