White Paper

The Automotive DataSphere: Increased Sensing Driving Growth

Sponsored by: Western Digital

Nina Turner Edward Burns
October 2021

IN THIS WHITE PAPER

Over the past few years, the automobile has been at the center of research and development on several different fronts. Advanced driver-assistance systems (ADAS) and autonomous vehicle (AV) functionality, infotainment and the digital cockpit, and the connected vehicle have all advanced significantly over the past few years. These advances have contributed to the continuing evolution of the automobile from a largely mechanical system to a modern data-driven, actuating vehicle. The increasing amount of sophistication in automobiles is demonstrated by the increased semiconductor content in the automotive space. According to IDC's Worldwide Semiconductor Applications Forecaster, automotive semiconductors are expected to grow from $39.4 billion in 2020 to $63.7 billion in 2025, at a five-year CAGR of 10.0%. As vehicles become more complex, the management of data from compute to storage becomes the most challenging aspect of vehicle development and operation.

The amount and complexity of data generated by each vehicle will vary depending on the penetration rates of these different technologies into the vehicle. The key areas of data consumption, creation, and communication include:

- ADAS functions – to autonomous vehicles
- Infotainment, navigation, and the digital cockpit
- Connected data – from the data sent to OEMs to communication with outside systems (infrastructure)

The advances in automotive technology are expanding beyond the traditional automobile and driving changes into commercial transportation and traffic infrastructure and even creating entirely new business models.

In this white paper, IDC examines the different systems that generate data, how the data is stored, and how the management of data within the vehicle is evolving in from low-end to high-end vehicles as well as some of the challenges facing OEMs and their suppliers and municipalities as vehicles transition from simple individual vehicles to complex, autonomous, and managed vehicles.

DRIVERS OF DATA GENERATION IN THE VEHICLE

From ADAS to Autonomous Vehicles

The sensor systems required for ADAS and eventually full autonomy will create the majority of data within the automobile, challenging the processors, software, and the network within the automobile. No one type of sensor is able to provide all the information necessary to handle the wide variety of driving conditions and situations that automobiles experience. The number of sensors, compute capabilities, and storage for ADAS vehicles varies by OEM, brand, and even model but will continue to increase as the complexity of functions grows.

Cameras

Cameras are seen as one of the primary sensors as we live in a visual world with signs and indicators directing how we drive. Initial implementations of ADAS functionality (emergency braking, lane-change warning, pedestrian warning, etc.) utilized a single forward-facing camera system. As ADAS functions began to require autonomous actuation, additional cameras were gradually added to the automobile to give the vehicle a better understanding of its environment. At the middle tier, vehicles can contain anywhere from 1 camera to 4 cameras, while higher-end vehicles can contain anywhere from 5 to 10 cameras and autonomous research vehicles can contain more than 20 cameras.

Automotive cameras range from 1.5Gbps to 1.0TBps (4K), each generating 672GB–5.4TB of raw data per hour depending on the frames per second and bits per pixel. Cameras can also address different parts of the light spectrum, providing additional information to the decision-making processors.

Radar

While cameras can act as the primary sensors, they cannot adequately handle situations where visibility is low such as nighttime conditions and poor weather. In those cases, radar can provide an additional sensing modality. Today, short- to medium-range radar utilizes 24GHz, and long-range radar utilizes 77GHz spectrum. Newer high-resolution radar with increased resolution adds the ability to detect objects in more than two dimensions and even shape. Data generated per sensor ranges from 0.1MBps to 24MBps depending on the number of receiving channels. In some ultra-high radar sensors, data generated can reach 375GBps.

Ultrasonic

Ultrasonic sensors provide short-range object detection with midtier automobiles having 0–4 sensors and high-end vehicles utilizing 8–16 sensors to surround the vehicle. Each ultrasonic sensor generates less than 0.01Mbps.

LiDAR

Most OEMs believe that light-detection and ranging (LiDAR) sensors will be necessary for automobiles to reach full autonomy and LiDAR suppliers have been able to reduce prices enough to be used in high-end luxury vehicles. For full autonomy, OEMs are considering anywhere from one to five sensors per vehicle, with each sensor generating from 58,000 points per second to over 37 million points per second. This large range is due to the rapid advancement of LiDAR technology. The few production vehicles with LiDAR feature technology from four to six years ago, and newer LiDAR technology has much higher resolution. Data generation depends on the field of view (FOV), angular resolution, and refresh rate. At 16 bits per point, data generated ranges from less than 1Mbps to 600Mbps.

Other Sensors (IR, HD Mapping, and V2X)

Other sensors and technologies such as vehicle to X (V2X), where X could be another vehicle or infrastructure, are being added as another safety system. Other sensors that are under consideration for addition in AVs include infrared (IR) sensors and high-definition (HD) maps. An IR sensor could generate up to 41Mbps, while high-definition maps would likely be a hybrid of onboard data with data being transmitted regularly for updates.

Example Data Generation for Audi A8 L2+ System

Consider the 2021 Audi A8 that operates at the SAE 2+ Level of autonomy. Announced in 2017, the 2021 Audi A8 was to be a L3 vehicle, but the software was not felt to be ready, so it was kept to a L2+ feature set. The Audi A8 sensors include five cameras (front camera and four 360-degree environment cameras), five radars (front long-range radar and four midrange side- and rear-assist radars),12 ultrasonic sensors, and a forward-facing LiDAR. Table 1 details an estimate of the amount of data generated by the A8 if all sensors are running and utilized.

This estimate shows that a Level 2+ vehicle can generate 3.9TB of data per hour. If 30 seconds of data is stored, this would require 32GB of solid state storage. Note that some of the sensors in this Audi configuration were developed three to five years ago, and resolution has increased significantly for both cameras and LiDAR today.

TABLE 1

Data Estimated for Audi A8 L2+ Sensor Configuration

	Megabits per Second	Gigabytes Generated per Hour	Number of Sensors	Total Gigabytes Generated per Hour
Front lane–assist camera (30fps, 2.1MP [1,920 x 1,080], 24 bits)	1,493	672	1	672
4 360-degree cameras (30fps, 2.1MP [1,920 x 1,080], 24 bits)	1,493	672	4	2,688
IR night vision camera (336 x 256 pixels, 30Hz)	41	19	1	19
Front long-range radar (77Ghz)	192	86	1	86
4 midrange side- and rear-assist radars	192	86	4	346
LiDAR (0.25 x 0.8 horizontal x vertical [H x V] degree resolution, 145 x 3.2 [H x V] FOV, 16 bits, 25Hz)	0.928	0.417	1	0.417
12 ultrasonic sensors (8 bits, 70kHz)	0.010	0.005	12	0.054
Total	3,412	1,536		3,811

Source: IDC, 2021

Example Data Generation for Fifth-Generation Waymo Driver

Waymo, a self-driving technology company spun out from Google, is on its fifth generation of autonomous vehicle technology, and its current AV sensor configuration includes 5 LiDAR sensors, 29 cameras, and 6 short-range and long-range radars. The number of sensors in a production automobile could decrease as Waymo optimizes cost versus performance. Table 2 details an estimate of the amount of data generated by the current published configuration if all sensors are running and utilized. This estimate shows that a Level 4 AV could generate over 22TB of data per hour. Utilizing the estimate from Table 2, storing just 30 seconds of raw data would require 184GB of storage.

TABLE 2

Data Estimated for Published Waymo Fifth-Generation Sensor Configuration

	Megabits per Second	Gigabytes Generated per Hour	Number of Sensors	Total Gigabytes Generated per Hour
1 front camera (HD 8.3MP, 30fps, 24 bits)	6,006	2,703	1	2,703
26 surround cameras (2.1MP [1,920 x 1,080], 30fps, 24 bits)	1,492	672	26	17,472
5 LiDAR (0.05 x 0.05 horizontal x vertical [H x V] degree resolution, 120 x 26 [H x V] FOV, 30Hz, 16 bits)	599	270	5	1,348
6 radars	192	86	6	518
Total	8,289	3,731	38	22,041

Source: IDC, 2021

Infotainment and the Digital Cockpit

Vehicle manufacturers use the infotainment system as a means of differentiation, allowing consumers to extend their digital lifestyle to their automobiles. Today's smartphone connected consumers expect their vehicle infotainment and control systems to be responsive and easy to use. While automotive systems used to take OEMs four to six years to develop and implement, the consumers' expectations have driven OEMs to improve their product development processes and incorporate more advanced interfaces within one to two years. The in-vehicle infotainment (IVI) system is still undergoing rapid changes as it evolves with OEMs still developing their applications and cloud services. More screens in the vehicle means more compute and more data to manage and store.

Audio

Audio systems are ubiquitous in the automobile and have undergone a steady evolution to digital radio, satellite radio, and smartphone integration and streaming services. Audio systems, either through a user's smartphone or through a separate connection in the automobile, consume different average rates of data:

- Low-quality audio streams at 96kbps, using 0.72MB/minute, or 43.2MB/hour.
- Normal-quality audio streams at 160kbps, using 1.20MB/minute, or 72MB/hour.
- High-quality music audio streams at 320kbps, using 2.40MB/minute, or 115.2MB/hour.

Data rates often auto-adjust based on network conditions. Many OEMs have converted to connecting with consumer devices to provide music and entertainment rather than add unnecessary cost to the vehicle.

Video

While consumption of audio is standard functionality, video in the driver area is primarily used for rear backup, surround view cameras, or mirror replacement rather than entertainment. Data rates for cameras are covered in the From ADAS to Autonomous Vehicles section. Rear-seat video entertainment is sometimes a premium option for some vehicles catering to families or is installed in the aftermarket. Some systems utilize DVD/Blu-ray players to provide the video feed, while connectivity options for consumers include Bluetooth, USB ports, SD cards, HDMI, and Wi-Fi.

Average streaming data rates can depend on a number of factors including compression, network conditions, and frame rates:

- Low-quality video (240p or 320p) streams from 0.18GB/hour to 0.45GB/hour.
- Standard-definition (SD) video (480p) streams from 0.48GB/hour to 0.66GB/hour.
- High-definition video (between 720p and 2K) streams at about 0.9GB/hour (720p), 1.5GB/hour (1,080p), and 3GB/hour (2K).

Digital Cockpit

The digital cockpit is the merging of the compute and management of the different visual displays in the front of the car. It combines the infotainment and navigation with the instrument cluster and potentially other screens in the driver cockpit area. Because data flowing through this system has to be managed for different functions, the processors and operating systems must be able to operate independently through hypervisors and/or independent compute cores. The system must be able to process the audio, video, and navigation as well as the human-machine interface (HMI) functions involved in driving while interfacing with ADAS control and management. Data management would also be a shared resource for functionality, with priority given to the instrument cluster screens and functions.

Gaming

Recently, AMD and Tesla announced that AMD processors would be used to power the newest Tesla S and X model infotainment systems. AMD processors are used in today's leading gaming consoles; Tesla also is enabling gaming for passengers. This unique function requires significant compute and onboard data storage. Today's console games have 16GB of DRAM and over 800GB of storage. The Tesla Arcade has 8GB of DRAM with 8–64GB of embedded MultiMediaCard (eMMC) flash memory.

Navigation

Consumers have become used to immediately being able to utilize maps to determine directions, understand latest traffic conditions, and locate the nearest points of interest such as restaurants or other places utilizing their smartphones. This expectation has moved to vehicles. Most OEMs have increasingly added cellular connectivity with 4G and 5G connectivity. There are multiple ways OEMs are implementing navigation, from smartphone connectivity through Apple CarPlay or Android interfaces to vehicles with cellular connectivity, GPS, and over-the-air (OTA) map updates.

Navigation today utilizes limited amount of data, but as OEMs partner with companies like HERE or implement their own datacenters to collect sensor data from vehicles for services like crowdsourced real-time traffic and driving condition updates, data is likely to increase. The use of high-definition maps for autonomous driving will increase data requirements, but the use of metadata layers will keep the amount of data minimal when compared with the amount of data generated by sensor systems used within the automobile for ADAS or autonomous systems.

Streaming data usage through smartphones is less than 1MB per 20 minutes of driving, or less than 5MB/hour. Similarly, today's onboard maps are updated one or two times per year with sizes of 8–10GB per download, depending on the size of the maps. Standalone navigation systems use more than 32GB of solid state NAND memory, but manufacturers may combine the storage needed for all IVI functions (including navigation) into a single storage device, driving storage capacity for IVI to 256GB and beyond.

The Connected Automobile

Connectivity is increasingly important for the vehicle as both a source of data and a means of vehicle management. As automobiles shift from being primarily mechanical to software defined, OEMs have begun to define how they will connect to the automobiles they manufacture, what data they will transmit, and how they will provide data and software security. Data transmitted between automobile and OEMs could include:

- Diagnostics
- Over-the-air software updates
- V2X connectivity
- Vehicle sensor data for autonomous vehicle development
- Feature activation
- Artificial intelligence–powered voice recognition
- Services that utilize cloud data
- Navigation data

Automobiles will have multiple ways of connecting including Bluetooth, Wi-Fi, and cellular.

Diagnostics

Consumers no longer have to take their vehicles for service to determine what is wrong with their vehicles. The availability of diagnostic data and vehicle status can be available to both the consumer and the OEM through the vehicle's cellular connection to the cloud and into the consumer's app. This data is small but provides the OEM and dealerships with potential service opportunities and a closer relationship with the consumer. The OEM cloud could also aggregate the data and determine potential manufacturing issues.

Over-the-Air Software Updates

With all the advanced features being added to vehicles, they are becoming software-defined vehicles. Today, vehicles can have over 100 million lines of code. It is estimated that a Level 5 autonomous driving could have up to a billion lines of code. In comparison, Windows 10 had around 50 million lines of code. Software security is also of utmost importance as vehicles become more connected. Security patches will likely need to occur over time. With this much complex software, the need for security, and the potential for feature enhancements, OEMs must be able to provide OTA software updates. Today, most OTA updates are limited to infotainment systems or telematics systems rather than with electronic control systems (ECUs) that are involved in safety systems and vehicle operation. This will evolve as OEMs gain experience managing their customers' vehicles. The ability to provide OTA would reduce the need for consumers to bring their cars in for software upgrades and ensure vehicles have the latest security updates.

Vehicle to X (Infrastructure, Pedestrian, and Other Vehicles)

Connectivity can also provide vehicles with additional awareness of their environment. V2X is in its infancy as the number of vehicles on the road with the capability is still low and standards are still being fought over. Vehicles can have an average age of 7–15 years depending on the region. This means that V2X implementation will not reach critical mass for a number of years. During this time, smartphones indicating pedestrian location, intelligent transportation systems (ITSs) on infrastructure, and commercial and safety vehicles will grow to include the ability to communicate with the vehicle. Connectivity technologies include dedicated short-range communications (DSRC) and 5G cellular (C-V2X), and with 5G's reduced latency, C-V2X has gained some traction. Some chipsets can provide both connectivity technologies in one package to hedge between which standard will win. With vehicles knowing where another vehicle is, accidents could be prevented.

CHANGES IN THE VEHICLE, TRANSPORTATION, AND INFRASTRUCTURE

Domain Architecture

With all the additional sensors around the vehicle, the network within the vehicle has become more complex. The disparate ECU architectures allow the ECUs to operate independently and require individual processors in each control unit. With the previously discussed trend toward OTA software updates, vehicle architecture should be designed to better enable this functionality. Centralizing compute power with a more powerful processor in a functional domain controller keeps sensors and actuators smaller and lighter. While some software would reside with the sensor, this would be less likely to require software changes compared with the processing done in the domain controller. While the domain controller architecture consolidates processing by function, reducing the number of ECUs, network complexity remains high as data must travel from all areas of the vehicle back to the domain controller. OEMs are gradually incorporating this architecture in their new model vehicles, but it will take a number of years before this architecture becomes pervasive. The centralization of processing power also means the ability to centralize some of the storage into fewer devices, both for the domain architecture and, to an even greater extent, for the zonal architecture.

Zonal Architecture

The increased number of actuators and sensors required for ADAS and AVs places significant demands on the wiring harness, which is one of the heaviest, most costly, and complex parts of a vehicle. The zonal architecture consolidates compute power further and utilizes a time-sensitive networking backbone. This architecture is still in development as it increases software complexity because a single zone processing unit would be responsible for controlling multiple sensor types and actuators in concert with other zones. A central compute unit is linked to the sensors and devices through a network of zonal processing units. This architecture provides the benefit of a more simplistic network, cabling harness, consolidated ECUs, and consolidated storage but requires more complex software and powerful processors with virtual machines performing different functionalities simultaneously.

Commercial Transportation

Many of the advances in the automobile are applicable to commercial vehicles including medium-sized and large trucks, buses, and other commercial vehicles. Fleet management driver monitoring is emerging as an application that businesses are using to improve driver performance and safety. Today, some services only report driver monitoring by how the vehicle reacts, but an interior-based camera system can detect driver distraction and signs of drowsiness and even verify driver identity. In some regions, driver monitoring is expected to be mandated. The driver-facing camera data will be locally processed for alerts and driver assistance, but the video data will also be stored locally when any incidents occur or to evaluate driver performance and then uploaded utilizing the vehicle's connectivity when cost is optimal. Commercial driver monitoring systems today utilize standalone data storage but could be integrated into the vehicle as advanced driver features become more pervasive.

Mobility as a Service

The intersection of connectivity and transportation has created the new business model of mobility as a service (MaaS). MaaS enables users to plan, schedule, and pay for services that can help them travel from one location to another utilizing shared resources rather than personally owned modes of transportation. Companies utilize the cloud to manage the transportation resources and ensure the user and mode of transportation connect. Companies are piloting and rolling out autonomous vehicle MaaS programs before L4/L5 autonomous vehicles reach the consumer market. These pilot programs can act as real-world tests and enable data collection utilizing a tightly monitored and managed fleet of vehicles. These MaaS fleets will have larger data storage onboard than consumer vehicles, as price sensitivity will be lower, utilization will be higher, and the data generated will be important for autonomous driving development.

Intelligent Transportation Systems

As connectivity, processing power, and features in the vehicle have grown, some countries and local municipalities are developing methods to improve safety, increase mobility, and provide benefits to the environment. Technologies that work between local infrastructure and the vehicle can improve traffic flow, identify and fine traffic violators, prioritize signals at intersections to improve traffic and transit performance, and quickly notify emergency services about incidents and reroute traffic.

New York's Connected Vehicle Project is implementing an ITS with the goal to eliminate traffic-related deaths and reduce crash-related injuries. The pilot program deployment is looking to utilize vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and vehicle-to-pedestrian (V2P) communications technologies. As part of the project, connected vehicle telematics gateways will be installed on over 3,000 city vehicles that travel, in aggregate, around 1 million miles per day.

Approximately 450 roadside units (RSUs) were installed in Manhattan, New York. In addition to performing V2I operations, select RSUs were used to efficiently update up to 9,000 vehicles over the air for software and firmware updates. The pilot program firmware files could exceed 100MB+ in size, requiring storage at each RSU and individual vehicle telematics unit. Vehicles passing near an RSU download packets over time, until the full update is downloaded to update the vehicle. The pilot programs utilized DSRC, which has lower data transmission rates, so a 5G C-V2X system could potentially support larger OTA updates, which would increase storage per RSU and vehicle telematics gateway.

ITSs can also include camera sensors that send data to a central traffic management center for object and vehicle identification, pattern recognition, and traffic flow analysis. The most widely used legacy sensor is the inductive-loop sensor, but other sensors could include acoustic, ultrasonic, and infrared sensors to account for different visibility conditions. These infrastructure systems could also provide other public safety functions beyond transportation. Connectivity to the ITS RSU can come through power over Ethernet or through a cellular connection. Management of data can emphasize compute at the edge or send compressed video to a central control center either at an edge location or in the cloud. Locating high-performance compute near the camera and sensors reduces latency, can improve real-time responses, and reduces network bandwidth usage. While keeping the majority of processing at a central datacenter could reduce individual RSU costs, network bandwidth can limit the number of sensors on the network. Processing of sensor data at the edge could require local storage, which retains a set amount of data that could be sent to the datacenter upon request or identification of a type of incident.

Challenges for this market include budget prioritization, uncoordinated development of infrastructure by local government, interdepartment conflicts, concerns about privacy, and vandalism.

IMPLICATIONS FOR DATA STORAGE AND THE NETWORK

ADAS and Autonomy Data Management and Storage

OEMs have not united on how data is handled in the vehicle. Currently, most of the data is sent to a central driver assistance computer as raw data or with some limited signal processing. This raw data stream from many sensors requires a more powerful central driver assistance computer. Eventually, some of the sensors may process data like image recognition at the sensor, decreasing the amount of data that must be delivered through the automobile's network. The key issue will be real-time performance and processing latency, as well as the capability of the network and central processors to handle the amount of data generated.

Retention of data will be important as vehicles become more autonomous. The recreation of accidents involving autonomous vehicles will be important in determining if the vehicle decision making is at fault. While government regulations have not kept up with the changes in technology, regulations could be put in place that required a certain amount of sensor data before an incident. 30 seconds of data before an incident with a full suite of data in the Audi A8 would be 32GB, while a L4 vehicle like the Waymo fifth generation would require almost 200GB. Data storage in vehicles uses automotive-grade solid state memory, while in-process compute memory is automotive-grade DRAM. The automotive industry is seeing a change from discrete systems, each with their own storage, to a shared storage system with one controller, resulting in the need for larger storage in the vehicle. Factors driving the increase in NAND per ECU includes the increased number of sensors to support ADAS and autonomous operation, the increasing complexity in the software in vehicles, and the implementation of OTA, which effectively doubles memory per ECU. Embedded NAND flash performs much better than removable cards due to shock, vibration, and other environmental factors in vehicles.

As more vehicles with full sensor suites hit the road, the data generated by these vehicles will be valuable for OEMs as AV training scenarios and can be transmitted back to the OEM either through Wi-Fi at the vehicle home/business or compressed and sent through cellular networks. This data collection requires additional storage to allow for delayed transmission to the OEM.

Infotainment and Digital Cockpit Data Management and Storage

OEMs have not enabled extensive storage of video media, but automotive-grade flash is available from 8GB to 256GB in embedded MultiMediaCard and up to 512GB as automotive-grade universal flash storage (UFS), which can further enhance the vehicle experience. The infotainment systems in a vehicle are prime opportunities to add leading third-party entertainment software; however, automakers are concerned about security. In addition, if the applications for infotainment share storage with other systems in the vehicle, this could limit the amount of storage available for infotainment. This likely means that the storage for the infotainment systems must remain separate from the storage for other systems in the vehicle. A company like Tesla may allow the consumer to plug in a solid state drive (SSD) for content or additional storage, but other automaker OEMs may limit the ability to add additional storage.

Connected Vehicle Management and Storage

OTA requires that vehicles have enough storage to communicate their system configuration, retain a copy of the current software image, download the new software image, and implement it across the vehicle. Some techniques to reduce the size of the image are determining a delta software package that can reduce the size of the download by over 90%. This storage memory would hold the software images for the OTA management system. There are two methods of upgrades in the vehicle. One requires that individual ECUs could need additional flash to retain the "current" firmware and have space for the new version as it waits to be switched. If anything goes wrong, the firmware can be rolled back to the previous version. In another approach, only one version of the firmware exists on the ECU and the delta is applied. This does not allow the update to occur while the ECU is in operation. The more conservative approach is the method that retains both the current and new versions of firmware, as it ensures that the updated component is rebootable at all times.

LOCALITY OF STORAGE

Cloud

There are many examples of data that will be transmitted to the cloud for processing and storage including less latency-sensitive data such as navigation, vehicle algorithm improvement opportunities, maintenance, performance, and traffic experience. Most of the storage in the cloud will be on HDDs due to less stringent latency requirements and the environmental controls in place for cloud datacenters. Data ownership philosophy and the experience operating a datacenter vary by OEM. Initially many OEMs will rely on outside cloud services, working with telecommunication companies and cloud datacenter companies to manage their data management needs. However, as the vehicle becomes more autonomous, software defined, and capable of OTA updates, OEMs may shift operations in-house.

To enable semiautonomous or autonomous vehicle development, OEMs are utilizing the vehicles on the road to collect data for algorithm improvement training. The data management methodology depends on what data the OEM believes it needs for R&D. Some OEMs may record sensor data, storing the data locally on the vehicle and selecting a subset of the data to upload when the vehicle has access to Wi-Fi. OEMs may also utilize vehicles to collect data for HD map development in a crowdsourcing model. Rather than send raw data, which would be costly and inefficient, the vehicle would collect metadata that gathers geometric and object identification data. This data summary can be condensed to tens of kilobytes per kilometer of driving, saving on storage requirements and enabling the data to also be transmitted over cellular networks. As more sensors are added to vehicles to support autonomy, storage will also have to increase.

For the infotainment system, noncritical data such as queries for information, navigation, or other data such as real-time map data will be regularly transmitted from the cloud. Navigation data updates to ensure a map with the most up-to-date information is maintained and will become more critical as vehicles begin to use HD maps for more autonomous operation.

Additional data from the vehicle performance such as engine output, emissions, mechanical issues, or driving performance is sent back to the edge or cloud for analysis. The frequency and amount of data tracked could increase as OEMs further develop their datacenter capabilities. This data will help improve car design, determine if software updates are needed, identify maintenance needs, and address performance.

While many of the regulatory requirements for storing vehicle data do not yet exist, there are likely scenarios that could take shape in the coming years. Recently, the U.S. National Highway Traffic Safety Administration issued an order requiring manufacturers and operators of vehicles equipped with SAE Level 2 ADAS or SAE Levels 3–5 automated driving systems to report crashes within one day of learning of certain crashes and an updated report within 10 days indicating whether ADAS or other autonomous function was in use within 30 seconds of the incident. While data reporting is not yet required, it is very possible that regulations could be put in place to require certain amount of data is stored in a vehicle's "black box" for incident investigation. Alternatively, certain data could be sent to the cloud and stored for a period of time after any incident.

Over time, more OEMs are also expected to realize the value of the data generated by vehicles and may look for opportunities to monetize it now or in the future, which in turn is expected to accelerate the demand for storage at the edge and in the cloud, generating incremental demand for SSDs and HDDs.

Edge

Semiautonomous or autonomous driving and the stringent latency requirements by many applications will require a great deal of storage and process to occur on the vehicle and in edge locations. V2X applications, for example, have response time requirements that cannot be met by the cloud and therefore will have data sent to edge infrastructure locations to be communicated to vehicles, requiring infrastructure data storage.

Another area of growth for edge applications is roadside units, which have not been implemented at scale yet mainly due to the limited implementation in vehicles today and a lack of public or governmental support. The edge concept is that each roadside unit would be tied into a centralized system (e.g., Caltrans traffic management system). The system would provide real-time traffic conditions or road conditions to the vehicle. These roadside units in theory would utilize cameras or other sensors for traffic, multifunction, safety (blind corner), and surveillance. All of these roadside units would utilize embedded or removable industrial-grade SSDs. The size of the capacity needed for these is still being determined as the technology implementations are largely in pilot tests today.

Most of the edge storage discussions have been focused on utilizing SSDs to meet the latency, environmental, and ruggedization requirements of many edge locations. If however the application requires higher capacities and is focused on warm or cold storage with less latency requirements and contained in an environmentally controlled environment such as a local telco hub, then HDDs would be the obvious choice.

Vehicles and transportation are the key use cases for telecommunication companies developing multi-access edge computing (MEC) due to the demand for lower-latency data access. The MEC enables data to be stored and processed closer to the endpoint application. Today's MEC devices utilize server systems and storage devices similar to those used in datacenters but may require industrial operating specification requirements depending on the location. As the technology and operational parameters evolve, more specialized systems could emerge.

In-Vehicle

The latency requirements in the vehicle demand onboard storage for data storage and management. The vehicle can utilize additional functions and provide data by being connected to edge or cloud datacenters to augment decision making and information. Companies must carefully profile vehicle data storage workloads. Read/write frequency can significantly impact the life span of solid state storage, so storage endurance must be taken into consideration when sizing storage devices against required system life span. Demanding applications such as vehicle black boxes that buffer data continuously during vehicle operation could require overprovisioning of storage to reduce the read/write frequency and meet system endurance requirements.

In-vehicle data storage needs are as follows:

- Retention of last 30 seconds of data for accident forensic examination will be a minimum requirement. This can range from 32GB minimum today to more than 200GB depending on sensor configuration. As more sensors are added to vehicles, the amount of storage will grow with data generation.
- For OEMs to collect data for their own autonomous vehicle development, crowdsourcing of data can provide real-world operational data. For this purpose, storage in some of today's production vehicles can vary from 32GB up to 1TB depending on the make, model, and trim.
- Storage for the infotainment system today ranges from 8GB to 256GB of embedded NAND. With the introduction of new infotainment options such as in-vehicle gaming or movies, storage requirements will likely see significant growth over time. Some manufacturers may allow USBs or other interface to increase storage, but this also can add intrusion security risks if the subsystem is not isolated from the vehicle's operational ECUs.
- OTA-capable ECUs will likely require double the existing embedded storage memory to mitigate upgrade issues. While the shift toward different automotive architectures will consolidate functionality and reduce the number of ECUs, the amount of software will increase, growing the demand for storage on a per-vehicle basis. Centralization of storage will occur as OEMs shift toward domain or zonal architectures.
- The harsh physical environment demands memory capable of surviving harsh conditions over a long lifetime. With the growing need for speed, the memory interface is a critical decision point, with UFS and PCIe interface providing faster performance than some of the older standards used in automotive.

SUMMARY

The changes in the vehicle and the advancements in technology brought about by sensors and IoT are all driving an exponential increase in data. Storage methodology varies depending on the data use, from requiring immediate action for operation to transmission to the cloud to longer-term retention. Data management in the vehicle is also significantly impacted by the location of the data computation – from processing at the edge sensor systems to the domain controller to the central driver-assistance controller to the cloud.

OEMs must:

- **Improve the ability to manage automobile data.** While most of the data will not be retained by the automobile and/or not sent to OEMs, OEMs will likely want to retain a percentage of that data to improve the performance of AVs. OEMs must strengthen their data management capabilities.
- **Enable a flexible, secure OTA methodology.** An OTA methodology that will work as vehicle models evolve will be essential as vehicles become increasingly software defined. It will be important to develop best practices for secure upgrades within the vehicle.
- **Secure the automobile.** As the automobile becomes a software-defined system with connectivity, it is essential that technology suppliers build in security into their automotive products. Processors, networking products, sensors, application processors, and microcontrollers all must include native security to ensure that there are multiple levels of security.
- **Establish standards for internal retention.** As more sophisticated ADAS and AVs emerge, the importance of what data gets retained will become a critical factor to determine the chain of responsibility. Developing these standards with various governments, insurance companies, and other suppliers will reduce confusion and eliminate a potential roadblock to the introduction of AVs.
- **Adopt the right memory interface.** With automotive technology rapidly becoming more sophisticated and performing at the leading edge of AI inference and autonomous operation, critical real-time decision making demands fast memory access. Choosing the right interface will be important to improving real-time performance.
- **Carefully profile data storage workloads when sizing storage.** Storage workloads must be considered when sizing storage, as read/write frequency can significantly impact storage life span. Overprovisioning should be considered for systems with more demanding storage functionality to meet system endurance requirements.

Beyond the vehicle, cities and countries are looking to utilize technology that will interact with the vehicle to improve safety and traffic. This is in its early stages but will only grow as vehicles become more capable. Growth in the use of sensors and compute in vehicles and transportation infrastructure will drive an increase in the amount of data used, stored, and communicated by municipalities.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

140 Kendrick Street
Building B
Needham, MA 02494
USA
508.872.8200
Twitter: @IDC
blogs.idc.com
www.idc.com