By DR. DEBENDRA DAS SHARMA,
PCI-SIG Board Member and Intel Fellow, Intel
Corp., and ANIL KUMAR, Principal Engineer,
Internet of Things Group, Intel Corp.
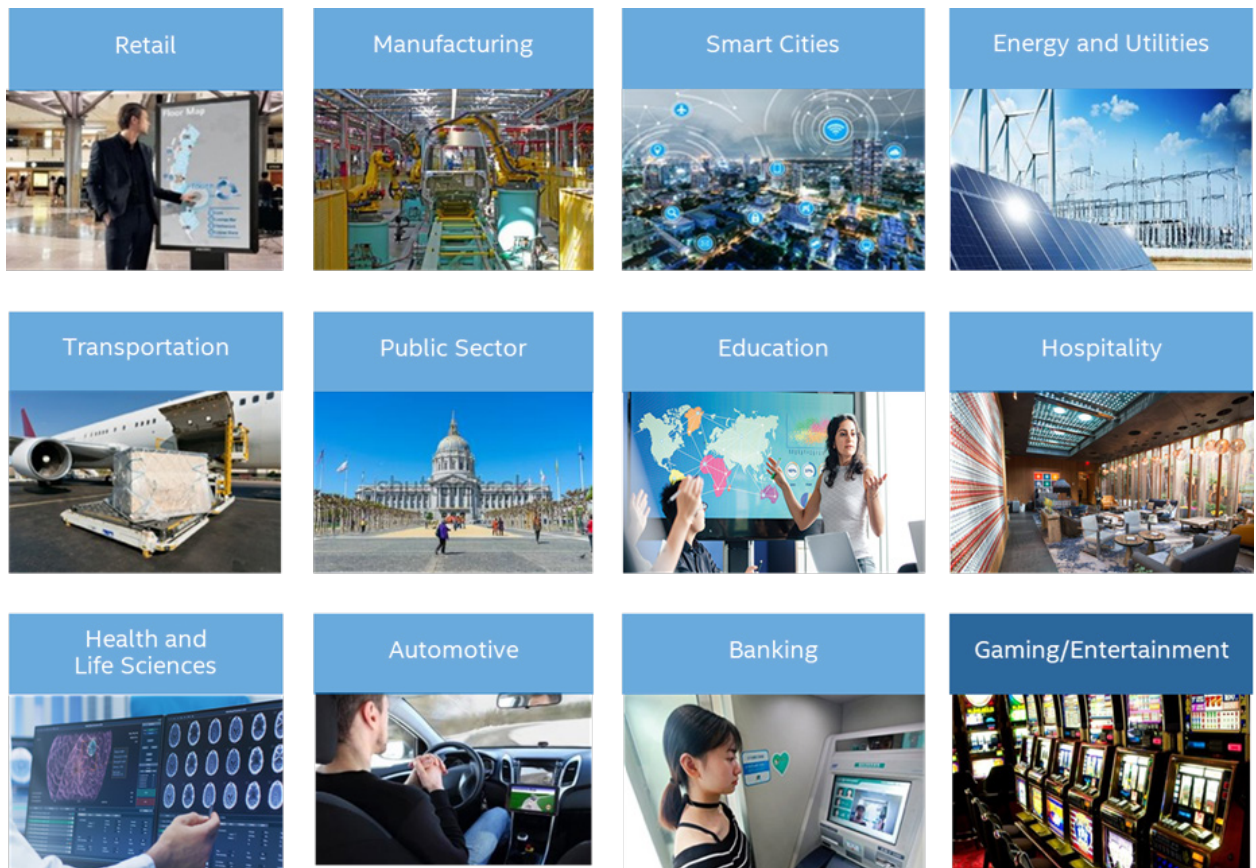
PCI Express Leads the Way in IoT Connectivity

High bandwidth and low latency are just two of the reasons why PCI technology has evolved to become the primary I/O interconnect in the IoT space.

The Internet of Things (IoT) represents a wide range of vertical markets spanning retail, manufacturing, smart cities, energy and utilities, transportation, public sector, education, hospitality, health and life sciences, automotive, banking, gaming, and entertainment—touching practically every aspect of our human lives *(Fig. 1)*.
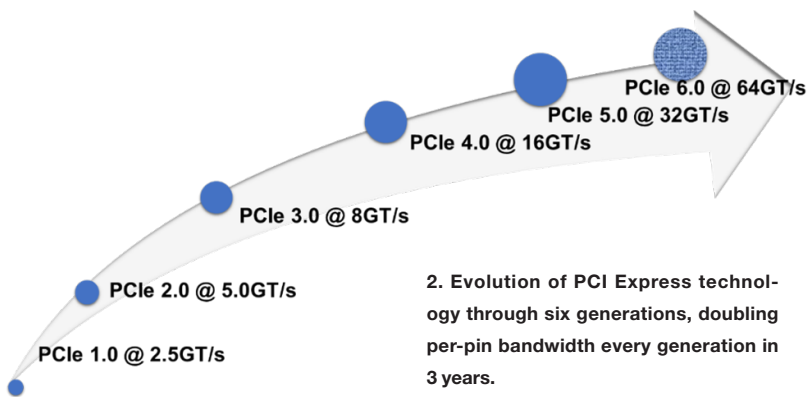
IoT represents a wide range of diverse systems spanning edge/gateway, data-acquisition devices, automotive, modu-lar computers, industrial PC, etc. One of the characteristics of the emerging edge and IoT market is the exponential growth in data as illustrated by a few statistics below:

- 80 zettabytes (ZB) of data by 40+ billion IoT devices in 2025, accounting for more than half the data.[1]
- 50% of data processed outside the data center or cloud by 2022.[2]
- 43% of AI tasks taking place on edge devices (vs.

1. IoT is ubiquitous and at the center of our rapidly changing world.

2. Evolution of PCI Express technology through six generations, doubling per-pin bandwidth every generation in 3 years.

PCIe 6.0 @ 64GT/s
PCIe 5.0 @ 32GT/s
PCIe 4.0 @ 16GT/s
PCIe 3.0 @ 8GT/s
PCIe 2.0 @ 5.0GT/s
PCIe 1.0 @ 2.5GT/s

cloud) in 2023.[3]

• Annual business-to-business (B2B) IoT revenue more than US$300B.[4]

Most of that data will be processed, stored, and analyzed at the edge to help deliver better latency, conserve network bandwidth, and improve reliability, security, and privacy. As data growth explodes, high-bandwidth I/O capabilities along with AI will become necessary to transport and transform this data into actionable information and valuable insights. In this article, we will delve into the role played by PCI Express (PCIe) technology as the primary I/O intercon-nect in the IoT space.

PCIe architecture has provided I/O connectivity for computing platforms for more than three decades. This is due to the ability of PCIe technology to seamlessly deliver a cost-effective, HVM (high volume manufacturing) friendly, power-efficient, high-bandwidth, and low-latency solution through six generations of technology evolution, doubling the data rate every generation and having full backward compatibility with prior generations (Fig. 2).

Additional protocol enhancements such as precision time measurement (PTM), I/O virtualization, and security features like link encryption are essential in the IoT space. Even though a multitude of form factors (e.g., PXI, M.2, U.2, CEM, various flavors of small form factors) have evolved to meet the needs of diverse systems across the IoT compute continuum, they all use the same silicon ingredients based on a common PCIe base specification.

The success of PCIe architecture as a ubiquitous I/O interconnect in the IoT space is due to it being an open industry standard. And it's backed by a robust compliance program to ensure seamless interoperability between devices made by different companies.

PCI-SIG, a consortium of more than 800 member companies spread across the globe, owns and manages PCI specifications and runs the compliance program. We expect PCIe technology to continue to evolve to meet the I/O needs of the IoT segments for many more years to come. Let's take a closer look at how PCIe architecture is used in the following IoT segments: edge computing, test equipment, embedded/industrial PCs, and automotive.

Edge Computing

Edge computing allows for the placement of resources needed to move, store, and process data closer to the source of the data or the point of service delivery (Fig. 3), catering to different levels of quality-of-service (QoS) expectations.
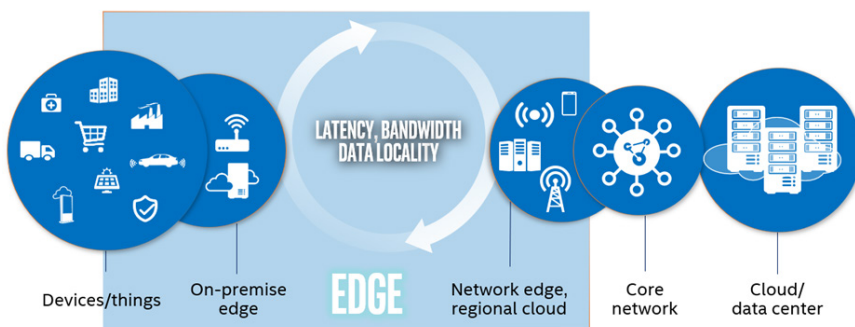
Based on optimized location, edge computing will be central to developing a new class of applications that require faster decision-making. Edge computing is an evolution and extension of cloud computing to transform the underlying architecture and create an environment ripe for application, service, and business model innovation.

The PCIe specification, which is widely deployed in the data-center market, is thus uniquely positioned to accelerate the development of edge solutions. The three key performance indicators driving edge value proposition and how PCIe technology capabilities fit into this big picture include:
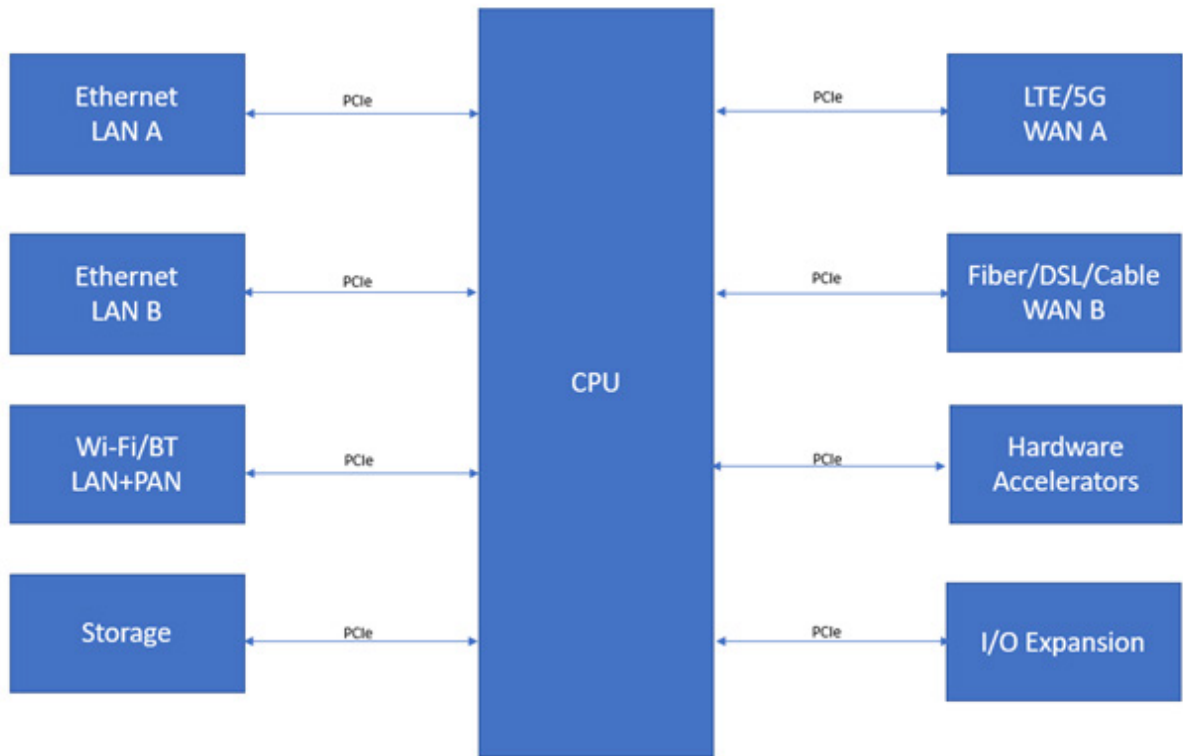
• *Latency and determinism:* Advanced applications such as factory automation, robotics, industrial process control systems, video surveillance and security, immersive media applications, autonomous vehicles, and content delivery demand sub-milliseconds to tens of milliseconds of end-to-end latency and deterministic deadline-based response. A traditional cloud-based solution can't meet this requirement.

• *Bandwidth:* By 2022, 82% of IP traffic will be video content. This traffic is expensive to transport to the cloud and process at a cloud data center. End users and service providers are motivated to process video data at the

3. The different edge locations in a things-to-cloud IoT system.

Devices/things On-premise edge EDGE Network edge, regional cloud Core network Cloud/data center

LATENCY, BANDWIDTH DATA LOCALITY

4. Shown is a simplified edge gateway block diagram: PCIe technology is the preferred choice within the system connectivity.

edge to lower latency, reduce jitter, improve video quality, and create new revenue with value-added services such as content delivery networks (CDNs), cloud gaming, or video analytics for enhanced retail experiences and smart-city applications.

- *Data locality and regulatory compliance:* Data privacy and security are driving the need for broad edge deployments as governments implement policies to protect customer data and maintain data within their geographies, such as the General Data Protection Regulation (GDPR) in the EU.

PCIe architecture, being a high-bandwidth, low-latency and secure data-connectivity standard, is ideally suited for these applications. *Figure 4* shows a simplified block diagram of an on-premises edge gateway and how PCIe technology provides the connectivity for the various subsystems within it. Let's look at this in detail:

- Every business needs to proactively plan for network failover for seamless operation. Wide-area network connectivity can be implemented using cellular 4G LTE/5G or wired solutions such as DSL, cable, or fiber-optic connections, each connected through PCIe. In this example, if a wired WAN connection goes down, one can failover to LTE/5G.

- The router can provide multiple LAN connections using wired Ethernet or wireless LAN based on Wi-Fi modules. For example, Ethernet-based cameras can be connected to the CPU via an Ethernet switch for a digital surveillance and security (DSS) system. Wireless peripherals such as a point-of-sale (PoS) system in a retail store can connect

via the Wi-Fi module.

- Low data-rate wireless connections, such as a handheld bar-code scanner or a temperature and humidity sensor on the shelf in a retail store, can connect to the edge system via Bluetooth. To provide office automation, light bulbs, window shades, remote controls, etc. can communicate via Bluetooth, Zigbee, or Z-Wave wireless connectivity solutions.

- DSS and networked video recorders, medical imaging systems, etc. need to store the captured data for further processing. Storage that's fast, reliable, and secure is critical for IoT applications. A high ratio of capacity to physical footprint, along with the ability to withstand harsh environmental conditions, also is an important consideration as edge devices are installed in outdoor harsh environments such as a petroleum refinery.

- The NVM Express (NVMe) specification is an interconnect protocol used to access high-speed storage media. It's designed to connect high-performance NAND flash memory to compute resources over native PCIe links. It's a streamlined memory interface with 64K I/O queues, each supporting up to 64K I/O operations, eliminating the legacy SCSI command stack and direct-attached-storage (DAS) bottlenecks associated with traditional hard-drive interfaces. This results in a uniquely tuned I/O architecture optimized for solid-state media.

Storage at the edge needs the ability to support 24x7x365 ingestion of new data and must have enough extra bandwidth and I/O capacity to handle batch and ad hoc queries against this data. A PCIe-based host interface also provides

5. An overview of workload consolidation based on industrial PCs.

NVMe SSDs with the high bandwidth needed to process this aggregated data at the edge into knowledge and convert it into actionable intelligence.

- Depending on the vertical market, IoT gateways support a host of legacy and I/O standards as well as custom I/O interfaces. Since the same CPU design is typically used to develop edge solutions servicing multiple markets to achieve economies of scale, PCIe technology provides an efficient way to connect a variety of I/O controllers to expand and aggregate these interfaces. For example, in the industrial-automation market, industrial Ethernet solutions such as EtherCAT, EtherNet/IP, PROFINET, CC-Link-IE, and serial protocols such as CIP (Common Industrial Protocol), PROFIBUS, Modbus, etc. are commonly implemented using PCIe interface aggregator chips. Similarly, BACnet for building automation and control networks, KNX, etc. are used in the building-automation industry.

- Emerging applications like augmented reality, depth perception, and gesture recognition require multiple image-sensor interfaces to connect to the application processor with minimal latency between frames. A MIPI-to-PCIe bridge can interface multiple MIPI CSI image sensors and aggregate data to a PCIe output. In the industrial sector, intelligent automation system captures and processes video for building security and quality control. A PCIe bridge captures videos from HDMI, DVI, or SDI interfaces and can transfer full uncompressed video and audio to a host processor for further processing. Similar bridges with LVDS as the video input also are available to support cameras that have LVDS as the video-output interface.

Edge computing is seeing increased deployment of heterogeneous computing with the PCIe interconnect due to its low-latency, high-bandwidth load-store semantics. For example, the growth in data processing at the edge has led to the use of dedicated accelerators for artificial intelligence (AI), machine learning (ML) and deep-learning workloads. These PCIe-based accelerators perform parallel computation, leading to faster execution of AI workloads compared to traditional CPUs.

Similarly, vision processing units (VPUs) are helping to offload vision workload involving multiple camera streams with high frame rates that have latency-critical processing requirements. A graphics processing unit (GPU) with its massive potential of parallel computing can be used in a vision workload to accelerate the processing of pixel data, in data mining, as well as other AI and ML workloads.

In applications where the algorithms and deep-learning neural networks are still evolving, reconfigurable field-programmable gate arrays (FPGAs) play a major role. Besides the acceleration functions, FPGAs provide many different connectivity options to connect to cameras, sensors, and other devices.

Test Equipment

Data acquisition is the process of measuring an electrical or physical phenomenon, such as voltage, current, temperature, pressure, or sound. From semiconductor device validation to automated production tests, PCI eXtensions for Instrumentation (PXI) systems deliver high-performance modular instruments and other I/O modules that feature specialized timing and synchronization, as well as key software features for test and measurement applications.

PXI systems use commercial PC-based PCI and PCI Express bus technology while combining rugged modular packaging. With this approach, one can meet requirements for timing, synchronization, and throughput across high-channel-count test applications.

Measurement accuracy is the most important consideration in designing any data-acquisition system. The overall performance of the system, including I/O sampling rates, throughput, and latency, are other top considerations. PXI can meet the ever-more-demanding application needs by incorporating the new PCIe standards into future products.[5]

Embedded/Industrial PCs

Industrial and embedded PCs (IPCs) are rugged PCs designed to operate continuously (24x7) in a harsh environment, including outdoor applications such as mining, agri-
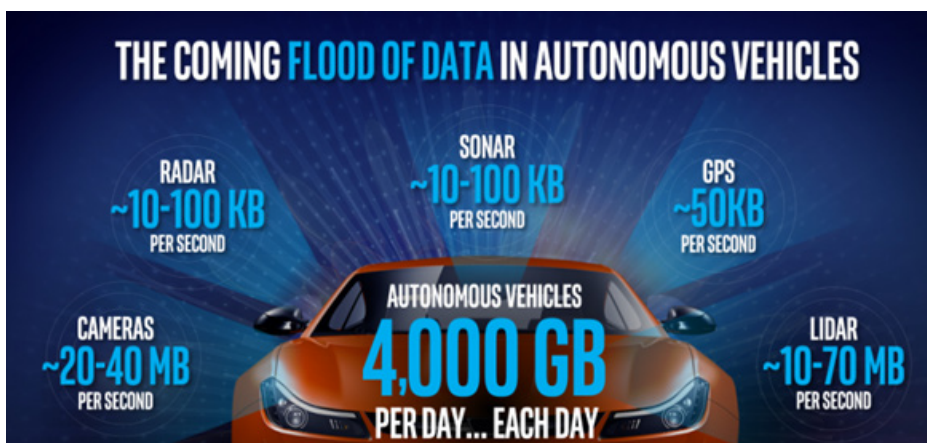
6. The connected car is akin to a data center on wheels.

culture, and energy distribution, and industrial applications like plant automation and process automation. They take advantage of the open PC architecture and deploy advanced video and AI technologies.

To realize the promise of the fourth industrial revolution, known as Industry 4.0, modern automation systems have moved away from purpose-built hardware to user-friendly networked PCs with high reliability, more data storage, and greater computing power. Supervisory control and data-acquisition (SCADA) systems and manufacturing execution systems (MES) are examples of such systems.

PCIe technology plays a key role in these PCs for storage, networking, accelerator attach, and even to connect to traditional automation systems using purpose-built hardware. It's resulted in reduced capital costs (CapEx), increased efficiencies, and simplification of operations to lower operational costs (OpEx), leading to seamless convergence of operational technology (OT) and information technology (IT) infrastructure.[6]

Systems utilized for workload consolidation *(Fig. 5)* are leveraging virtualization. Virtualization partitions the host platform into multiple software-defined, isolated environments with shared resources managed by a hypervisor.

Such a software-defined infrastructure has been made possible by advances in multicore CPUs, real-time operating systems (RTOS), lightweight hypervisors, and hardware-assisted virtualization technologies. I/O virtualization support in the PCIe specification has facilitated offloading of packet processing to network adapters, as well as direct assignment of virtual machines (VM) to virtual functions, including disk I/O.

In control and automation systems, applications often are required to be synchronized across the network or devices within the subsystem to a common clock, with different levels of accuracy depending on the end usage. Ethernet time-sensitive-network (TSN)-based connectivity ensures

synchronization between systems using the IEEE1588 and 802.1AS standards.

The PCIe specification's Precision Time Measurement (PTM) protocol enables TSN network interface cards (NICs) or other peripherals to synchronize the network time to the CPU timestamp and ultimately to the software application. PTM minimizes the translation errors and thus increases the coordination of events across multiple components with very fine precision. Data from different sources, such as cameras, transducers, etc., can be reassembled in the host processor or accelerators through timestamps enabled by PTM to perform accurate analysis and take actions, even in safety-critical applications.

IPCs offer greater customization, reliability, and scalability, plus a longer product lifecycle. The wide range of systems include low-power and fanless designs in space-constrained environments, industrial panel-mounted PCs installed on surfaces, industrial rackmount PCs fastened to walls or in cabinets, industrial workstation PCs in closets, and embedded PCs in kiosks or medical equipment.

A popular form-factor is COM Express developed by the PCI Industrial Computer Manufacturers Group (PICMG).[7] The ability to plug a COM Express module or standard SFF wireless modules like M.2 onto a carrier board reduces the time and cost to develop a product. The user doesn't need to understand the typically complex details associated with high-speed signaling, RF design and tuning, etc. for the latest SoCs.

This approach also streamlines the effort needed to get regulatory approvals such as FCC certification for the final product. Moreover, a modular design approach allows us to address the "mile wide and inch deep" IoT market by customizing the solutions based on common solutions driving economies of scale and wider adoption.

Automotive

Automotive electronics has been growing in sophistication and complexity for over a decade.[8] Modern cars with their in-vehicle infotainment (IVI) systems, advanced driver-assistance systems (ADAS), and vehicle-to-vehicle (V2V) and vehicle-to-everything (V2X) connectivity can be considered as data centers on wheels *(Fig. 6)*. To satisfy the consumer expectations for comfort and meet regulatory requirements for passenger and pedestrian safety, these sys-
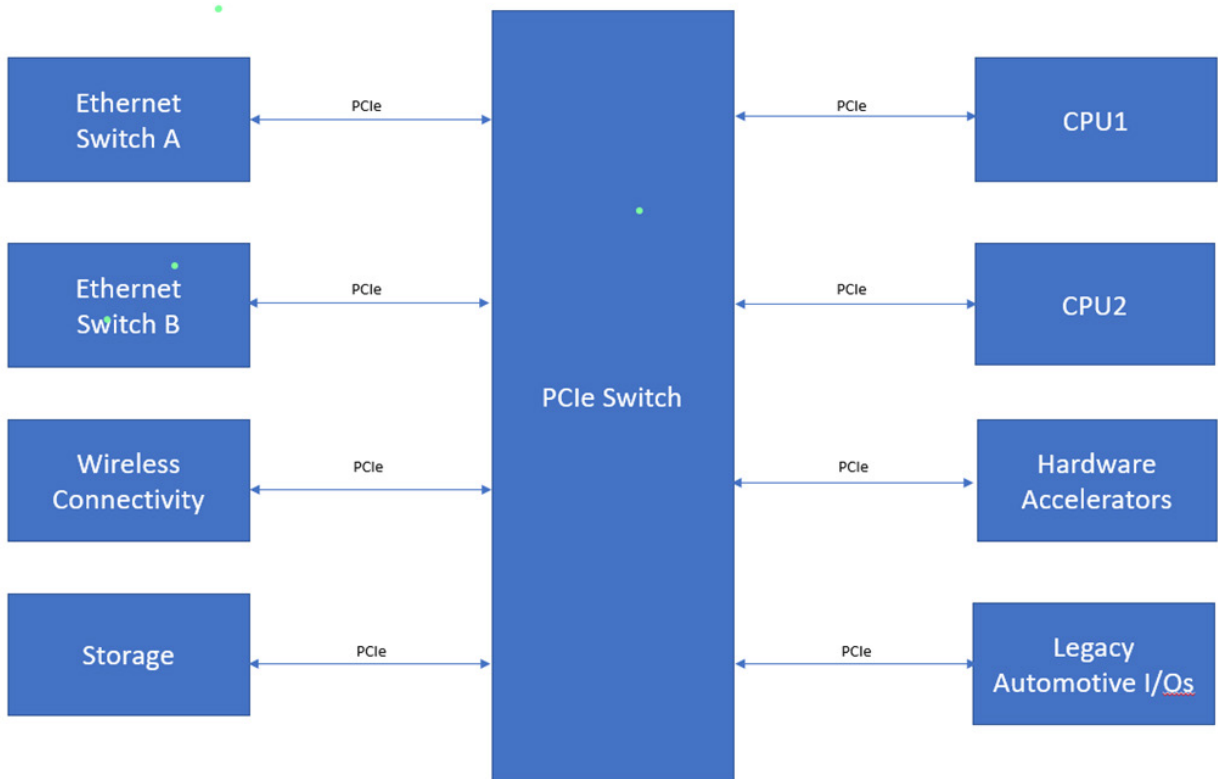
7. The diagram illustrates an overview of high-performance automotive compute architecture.

tems are becoming prevalent in mass-market models and not just limited to luxury automobiles.

With the increased amount of data to be transported, stored, and processed, proven data-center technologies such as PCIe architecture are becoming very common in automotive electronic systems to connect powerful multi-socketed CPU systems. Beyond PCIe-connected GPUs, Wi-Fi, V2X, and LTE/5G cellular modules, specialized vision and AI accelerator ASICs, Ethernet NICs, and FPGAs are common in automotive electronic control units (ECUs).

As expected, this segment has very stringent regulatory, safety, and reliability requirements while operating under extreme conditions, with an expectation of availability without replacement for the entire useful life of about 15 years of a car. PCIe technology, with its high reliability and availability features, is the interconnect of choice to interface to multiple data sources comprising cameras, LiDAR, radar, and ultrasound and infrared sensors. It provides wireless connectivity through Wi-Fi, BT, LTE, 5G, and satellite for the car. It's also used for connection to storage devices and accelerators *(Fig. 7)*.

Conclusion

The PCIe interface is natively supported by most high-performance CPUs. In addition, the ecosystem includes a large range of available devices for AI/ML acceleration, networking, storage, and wireless connectivity. Availability of IP blocks, off-the-shelf standard form factor/modular com-

ponents, debug tools, and a large ecosystem of developers familiar with the technology from the IT world is accelerating the penetration of PCIe architecture into IoT applications. The *table* summarizes the key attributes of PCIe that drive optimal solutions in the IoT space.

With a rich and successful history of navigating several technology transitions in a backward-compatible manner spanning three decades, PCI-SIG is well-positioned to continue leading the evolving IoT landscape going forward. The power and promise of this open standards organization, backed by the combined innovation capability of 800+ member companies, makes the technology nimble, scalable, cost-effective, power-efficient, and multi-generational with relevance across all IoT segments and usage models for the foreseeable future.

References
1. DataAge 2025, November 2018
2. Gartner, December 2018
3. ABI Research, May 2018
4. Forbes, December 10, 2017
5. PXI System Alliance
6. Workload Consolidation in Industrial IoT
7. COM Express Overview
8. Data is the New Oil in the Future of Automated Driving