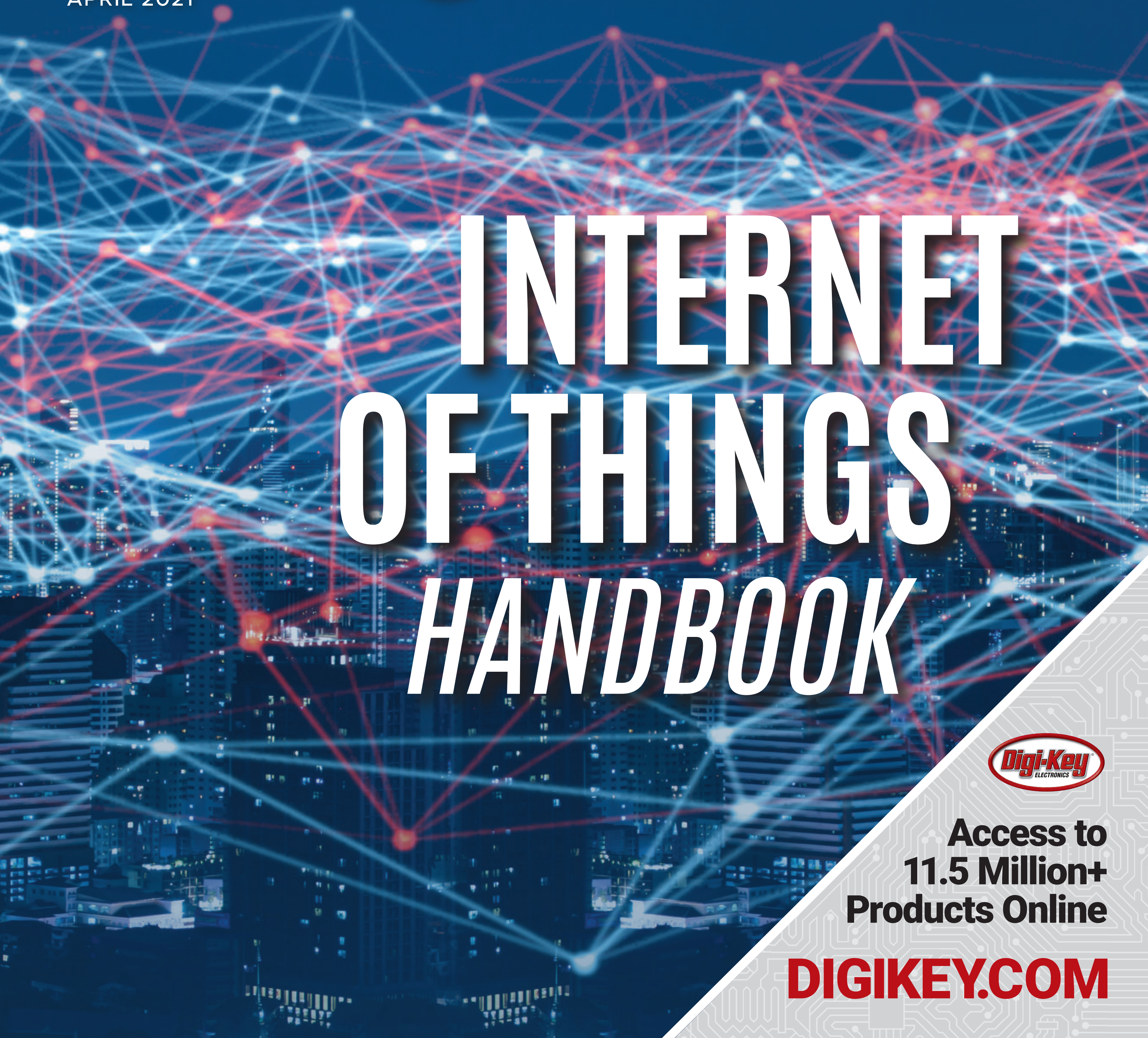
Edge computing applications for equipment designers **Page 8**
Tips on your network installation **Page 30**
Securing the IIoT **Page 40**

Design World

APRIL 2021



INTERNET OF THINGS *HANDBOOK*

Consolidating workloads at the rugged edge: **Enabling inference computing anywhere**

Severe settings are tapping into real-time processing, advancing a new level of machine learning and AI.

Dustin Seetoo,
Director of Product Marketing • Premio Inc.

Deep, real-time data is foundational to artificial intelligence, making it imperative for industrial environments to capitalize on the deluge of data generated by intelligent, connected devices. Yet all too often, these harsh and unstable environments have not been fully served by high-performance computing based on the hurdles to achieving reliable machine learning performance. This is the difference between the edge and the rugged edge, where more demanding, real-time industrial computing has the power to change the way industrial businesses operate and compete. The rugged edge is a unique place, where heavy industry relies on computing systems that blend the latest high-performance technologies to accelerate data processing fed by a variety of sensor input data.

To break new ground in these rigorous settings, rugged, high-reliability systems must efficiently pass data back and forth to the core – capitalizing on reduced latency, accelerated processing, and increased data storage capacity right where it is needed. It's a convergence of the latest IoT technologies from compute, storage, connectivity, and ruggedization, driving a new level of reliability in harsh, mobile, and remote environments.

What are the design principles that enable systems to survive and thrive at the rugged edge? These can be best demonstrated by looking closely at some of the emerging IoT applications in play, for example, telematics and autonomous fleet routing, where hardware supports inference analysis and machine learning for in-vehicle computers.

INFERENCE COMPUTING IN ACTION

Successful autonomous fleet routing relies on a wealth of data provided by various sensors that indicate what is happening in real time. In addition to increasing

An autonomously driven truck is an example of a rugged edge system. It captures data on its route with a range of sensors in the vehicle often sending that data over CAN bus. Critical data, such as, vehicle telematics, GPS, and speed, are transmitted consistently while scheduling information and safety data, including road conditions or obstructions, are simultaneously received.

vehicle efficiency and safety, these data sources are instrumental in vehicle tracking, routing, reporting, and even monitoring the need for vehicle service. They also simultaneously process data concerning weather and traffic conditions to optimize fuel consumption.

A proof of concept demonstrates the challenge, outlined by this list of requirements established by a trucking fleet manufacturer developing an autonomous routing application:

- Commercial off-the-shelf (COTS) industrial-grade computers for fast customization and deployment
- High reliability, supported by validation to heavy shock and vibration in a small system footprint
- Fanless computing design able to withstand wide operating temperatures and input voltage from vehicle batteries
- Embedded architectural structure, purpose-built to aggregate and monitor high-value data and transmit it back into a real-time neural network
- PCIe GPU support for real-time analytics and compute processing for real-time detection
- Storage capacities up to 32 terabytes to log large amounts of available data
- Compact IoT gateway equipped with diverse I/O and connectivity options for the most convenient integration with various vehicles for real-time processing and data telemetry

In a system developed based on these parameters, an autonomously driven truck captures data on its route from point A to point B with a range of sensors deployed in the vehicle, such as cameras, LiDAR, radar, and CAN bus. Critical data, for example, vehicle telematics, GPS, and speed, are transmitted consistently while scheduling information and safety data, including road conditions or obstructions, are simultaneously received. Data are collected, aggregated, and then offloaded from the system for further analytics. Smart algorithms use these data to train a car to become more intelligent and safer, leveraging the rugged edge system itself as a data aggregation device to enable the vehicle's neural network training.

In this type of inference computing application, the system recognizes that a four-legged object is crossing the road, likely a dog, by relying on powerful processing near the edge to inform critical decisions in real time through hardware acceleration and machine intelligence. In addition to dedicated hardware to process and run algorithms effectively, the application demands a passive cooling design, validated to operate in extreme temperatures, has resistance to shock and vibration, and supports wide-voltage power protection for vehicle batteries. Hardware must be explicitly designed for performance in the context of mission-critical mobile computing rigors, featuring powerful multicore CPU/GPU processing, more bandwidth, low latency, and seamless connectivity – principles outlined in the following sections.

BLEND PERFORMANCE ACCELERATORS FOR REAL-TIME PROCESSING AT THE EDGE

Long gone are the days when edge computers only collected information from IoT devices. Advances in edge computing hardware enable performance computing at the edge. Careful consideration of processor configuration is required, determining the number of workloads the system can handle and the speed at which it completes tasks. Systems performing tasks such as data telemetry and remote monitoring of vehicles may be able to tap a low-powered yet efficient processor. In contrast, performing complex and demanding workloads, like inference analysis and AI edge computing, might require a more powerful and robust option, for instance, the Intel Core i3, i5, and i7 processors. When low power is a requirement, consider SoC (system-on-chip) options, which combine all the components of a computer on a single substrate such as CPU core, integrated graphics GPU, and memory storage on a single chip. This ensures a balance of power consumption and performance.

For complex industrial workloads that demand more power, a socket CPU system might be the optimal option. Socket CPU systems generally offer more performance than SoCs – with more cores, higher clocked cores, and a higher TDP, they operate at a much higher temperature than SoC PCs. However, socket PCs consist of a motherboard with a CPU that is mechanically installed into the socket. Industrial-grade edge computers that use socket type CPUs can also be engineered in fanless designs that use passive cooling for better reliability in the harshest environments. These types of computing solutions strike a solid balance between multicore performance and rugged reliability in hardened designs.

Other components, such as GPU and memory, must also be manually inserted into the system through other slots on the motherboard.

Note that adding GPUs for hardware acceleration may be ideal for AI and machine learning applications, allowing edge computers to store, process, and analyze large amounts of data without having to move data to the cloud – these access an abundance of cores for parallelism versus a sequential CPU, for real-time inference. As a result, their deployment could save a considerable amount of money and internet bandwidth, especially if a metered data plan is in place.

ADDRESS THE REALITIES OF THE PHYSICAL COMPUTING ENVIRONMENT

Edge computing hardware must be rugged enough to withstand operation in volatile environments where systems are exposed to frequent shocks, vibrations, dust, debris, and even extreme temperatures. Fanless design is a defining feature for this type of deployment. No fans or ventilation holes are necessary, preventing the possibility of dust and debris entering the system and damaging sensitive internal components. The elimination of fans themselves adds reliability and removes a common point of failure in air-cooled systems.

Systems should also be cableless, ensuring optimal handling of shock and vibration. Cable-free systems are not subject to the possibility of a loose cable that could render the system inoperable. Enclosures of aluminum and steel complete the environmental protection strategy, making cleaning the computer's outer portion easy while preventing the system from corroding and deteriorating if exposed to water and dirt. Protection levels vary, so evaluate the need for ingress rating protections as a design feature that adds reliability and longevity to system performance.

> Edge computing hardware must be rugged enough to withstand operation in volatile environments where systems are exposed to frequent shocks, vibrations, dust, debris, and even extreme temperatures. Fanless designs are often preferred here. Designs without fans or ventilation holes prevent the possibility of dust and debris entering the system and damaging sensitive internal components.

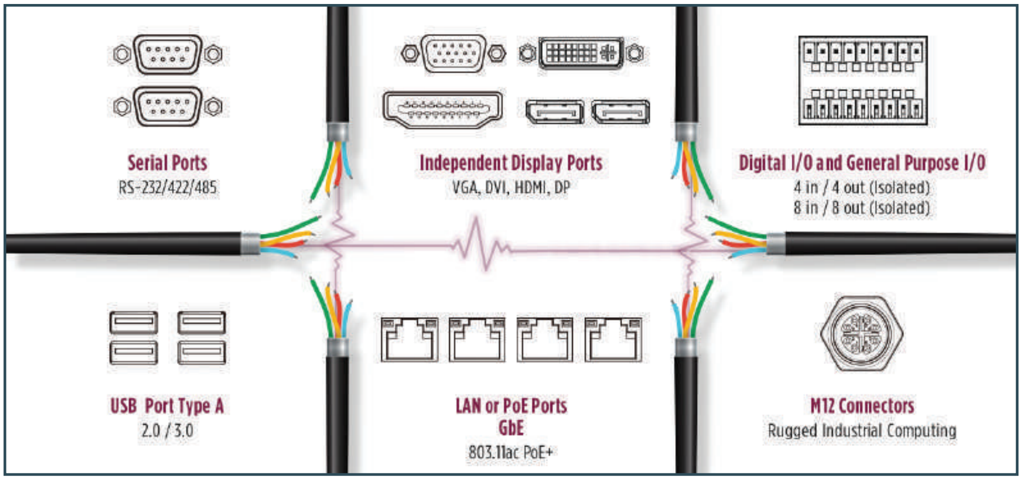
INCREASE RAM FOR RESPONSIVENESS, SSDS FOR RELIABILITY

Edge PCs often collect, process, and analyze large amounts of data gathered from machinery, equipment, and industrial IoT devices – sufficient storage and quick data access are critical features. The more memory added to a system, the greater its responsiveness to accessible workloads for immediate data cache.

For long-term storage, small form factor rugged PCs can be equipped with HDDs (hard drives) and/or SSDs (solid-state drives). A single enterprise-grade SSD might hold terabytes of data and transfer data at faster speeds than hard drives. Additionally, rugged PCs can be configured with NVMe SSDs, extremely fast storage devices capable of read speeds up to 3,500 MB/s and write speeds up to 2,500 MB/s. In configuring an ultra-rugged small form factor PC, designers should opt for SSDs because they store data on silicon chips, which offer magnitudes of reliability over hard drives reliant on spinning metal platters for data storage. SSD cost is often prohibitive in terms of price per capacity, but it is potentially shortsighted to defer reliability and ruggedness in system storage that may experience shock and vibration. That said, for organizations requiring large amounts of data storage, hard drives can be added for additional volume. For example, large storage capacity works well for long-haul driving data, including a 10 GbE connection to move data bi-directionally – effectively stored and accessible to continuously train machine learning algorithms for ongoing improvement.

WIDE POWER OPTIONS PROTECT MISSION-CRITICAL PERFORMANCE

Rugged small form factor PCs equipped with a wide input voltage allow rugged computer systems to run on power that ranges from 9 to 50 VDC, adding compatibility with different power input scenarios. Rugged small form factor PCs must also be equipped with overvoltage protection, power surge protection, and reverse polarity protection. Overvoltage protection cuts off the system's power whenever the system senses that the voltage exceeds a predetermined level to protect the sensitive internal components. Power surge protection is required to protect the computer system from power surges. Whenever the system detects a power surge, it diverts electricity into the ground to avoid damaging the sensitive electronics onboard the device. Finally, reverse polarity protection ensures the system is not damaged if power supply polarity is reversed. When such a system detects a reverse in polarity, power to the system is cut to prevent sensitive components from sustaining damage. These features are unique requirements in industrial-grade computing solutions, designed to assure reliability and performance in mission-critical enterprise deployments.



- IoT integrators can best address I/O requirements with a plug and play option. Popular I/O options include USB 3.1 Gen 2 ports capable of data transfer speeds up to 10Gbs, COM ports for legacy devices, and RJ45/M12 Ethernet ports for LAN/ PoE+ devices, and General Purpose I/O ports (GPIO) to program on/off triggers in automation.

RICH I/O PROTECTS FLEXIBILITY FOR INTEGRATION WITH NEW AND LEGACY EQUIPMENT

Rugged edge PCs must offer a diverse slate of I/O ports and flexible PCIe expansion because they often connect to both new and legacy factory machinery, devices, and equipment. Rugged edge PCs also have to offer integrations for modular daughterboards to accommodate configurations that need additional I/O. This key design feature allows IoT integrators to address I/O requirements with a plug and play option, often not available in fixed I/O configurations. Popular I/O options include USB 3.1 Gen 2 ports capable of data transfer speeds up to 10Gbs, COM ports for legacy devices, and RJ45/M12 Ethernet ports for LAN/ PoE+ devices, and General Purpose I/O ports (GPIO) to program on/off triggers in automation. GPIO ports can accommodate peripherals, sensors, and devices that do not have a common interface, for instance, a USB port or legacy serial port. Ultimately, GPIO ports allow edge computing hardware to connect to digital devices regardless of age; if a device (or its sensors) is functional, it can be connected to an edge computing solution.

UNDERSTANDING THE RUGGED EDGE

Just like factory equipment continually learns to recognize objects on the processing line, a telematics solution in autonomous vehicles enables operability and responsiveness improvement via machine learning. These systems perform parallel computations to process data regarding road conditions, hazards, driver maneuvers, and vehicle control systems – with minimal latency. Working with massive amounts of data in real-time, autonomous routing solutions must be rugged enough to withstand the range of temperatures, shock, and vibration that can cause electronics reliability issues. Also critical are pinpoint accuracy to the vehicle's GPS location and the ability to accurately push all the data to a central database enabling the telematics manager to monitor what is happening in real-time.

These challenges represent a number of emerging IoT applications such as data center performance enabled on the factory floor, intelligent vending systems using modular designs for smart workload consolidation, processing plants deploying inference computing to improve output and safety, and much more. Understanding the rugged edge – what it needs and what it enables – will help designers hit the mark with a full spectrum of new and exciting inference computing applications in the harshest remote and mobile environments.

WIRED, WIRELESS, OR ALL THE ABOVE

Autonomous routing solutions must be able to accommodate a variety of wired, wireless, and cellular connectivity options for seamless connectivity from tower to entire fleet network. Optimized systems should be equipped with two RJ45 LAN ports for very fast wired data transfer, ranging from 1 GbE to even 10GbE along with support for Wake-on-LAN and PXE. If the end-use application requires additional RJ45 LAN ports or M12 LAN ports capable of PoE+ (IEEE 802.3at) for data and power through a single cable, easy-to-install expansion daughterboards can be deployed.

In the event that wired connectivity is not available, edge computers should still be able to connect to the internet via options for Wi-Fi6 (IEEE 802.11ax), used for ultra-reliable, low latency communications through high-speed wireless LAN connectivity. WiFi6 also offers configuration flexibility for range and power consumption as well as performance that is closest to wired LAN in terms of latency and bandwidth. In the absence of Wi-Fi, edge computing solutions should be able to offload critical data via cellular 4G, LTE, and 5G connectivity. Multiple SIM module sockets should be integrated as well, allowing organizations to add up to two data carriers for redundancy. 4G tops out at a theoretical speed of 100 Mbps, whereas 5G tops out at a theoretical speed of 10 Gbps – the kind of fast cellular connectivity is extremely important for remote deployments where stable internet connectivity is not always available. Bluetooth connectivity is an additional option as a lightweight technology that offers quick and straightforward connectivity for low-powered devices. Although Bluetooth does not provide the speed and range that Wi-Fi offers, it does provide reliable one-to-one and many-to-many connectivity.