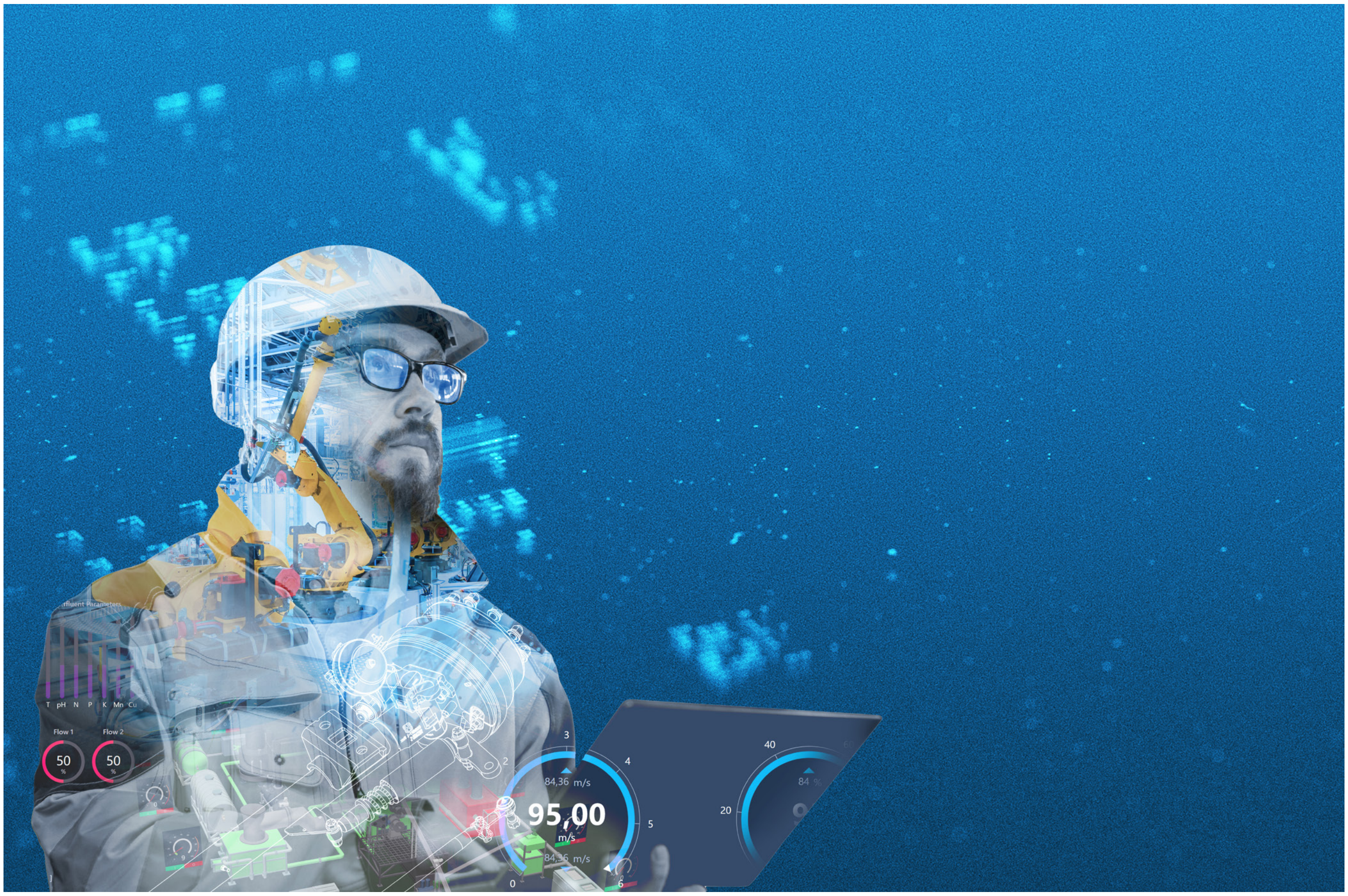
**White Paper**

00870-0100-0100
November 2021

# Liberating Stranded Data Via IIoT



# Liberating Stranded Data Via IIoT

Automation systems are in service almost everywhere, controlling and monitoring machines, industrial manufacturing, infrastructure, and other facilities. This activity generates volumes of data which could be used to improve operations and business outcomes—but only if the data can be accessed, managed and analyzed effectively. Unfortunately, this valuable stranded data often remains inaccessible for many technical and commercial reasons.

Newer architectures are changing this predicament by combining flexible and capable edge computing with a cloud computing model, making it not only feasible, but actually practical, to analyze this data, gain new insights, and to make the results available to users.

Whether deployed at the edge or in the cloud—and more commonly both—any solution must be compatible with operations technology (OT) systems and legacy technologies, yet incorporate the latest information technology (IT) innovations for communications, accessibility and security. At the edge, in the cloud, and for the connections between them, openness, interoperability, cybersecurity and connectivity are vital to successful implementations.

This white paper identifies some of the most common reasons end users struggle with stranded data, and how a modern edge-to-cloud industrial internet of things (IIoT) solution makes it easy to access and use this data.

General limitations and ranges of the AMS Asset Monitor Tach CHARM will be provided in this document. This information is provided for the purpose of comparison and compatibility verification to select a compatible supported sensor.

## How Existing Infrastructures Strand Data

Until recent years, most manufacturing data was sourced from PLCs, HMIs, SCADA, and historian systems running in the OT domain. These systems are built using industrial automation products, which maintain a tight focus on achieving control, visibility, production and performance to maximize operational efficiency and uptime. As such, accessing and analyzing the associated data beyond immediate production goals is a secondary concern.

Designers architected and scaled the OT infrastructure correspondingly, leading to design choices like:

- Selecting proprietary protocols meeting performance requirements, but without supporting flexibility and cross-vendor interoperability.
- Minimizing control and sensor data collected to maximize system reliability and simplicity.
- Implementing localized on-premise architectures to minimize cybersecurity threats.
- Vendor lockout schemes to protect intellectual property and promote reliable machine operation, often at the expense of connectivity.

While within the OT environment these data sources appear to be open, they are still quite difficult to access for applications outside of the OT environment where the data could be more easily analyzed. In addition, many potentially valuable sources of data—such as environmental conditions, condition-monitoring information, and utility consumption—are simply not needed for production or equipment control and are therefore ignored or not collected by these automation systems. Even though the resulting systems performed admirably from a local standpoint in the context of their operational goals, they suffered from data “blind spots” and did not benefit from analysis of the full set of potentially accessible data. Big data analytics capabilities continue to expand, but the constrained access to this stranded data continues to limit their potential.

Traditionally, industrial automation systems were installed strictly on-premises and largely constrained by the specific hardware and software technologies available. OT devices were often installed with “just enough” processing power and therefore lack the necessary computing and storage elasticity to deliver on the full potential of the data often desired with the new class of analysis tools. Most of the familiar OT-centric software is not designed from the ground up for cloud connectivity, and certain features, and even basic security requirements, may be implemented inconsistently as add-ons/bolt-ons. In many cases, IT-centric software is not attuned to the always-on, low-latency, and high-volume data needs of an OT environment, and fails to offer a broad set of vendor connectivity controls. Security as an afterthought is not good enough; it must be built-in at all levels of the solution.

Pulling together ill-suited software to create an IIoT solution is problematic, can require kludgy custom scripts, and is hard to support. With these needs and challenges in mind, some users may question how to recognize stranded data so they can begin taking steps to liberate it.

## Types of Stranded Data

Stranded data exists in many forms. It originates at machines, the factory floor, and other systems throughout a facility, often as part of the OT but also associated with other support and utility systems. This data can be as granular as a single temperature reading, or as extensive as a historical data log identifying the number of times an operator acknowledged an alarm.

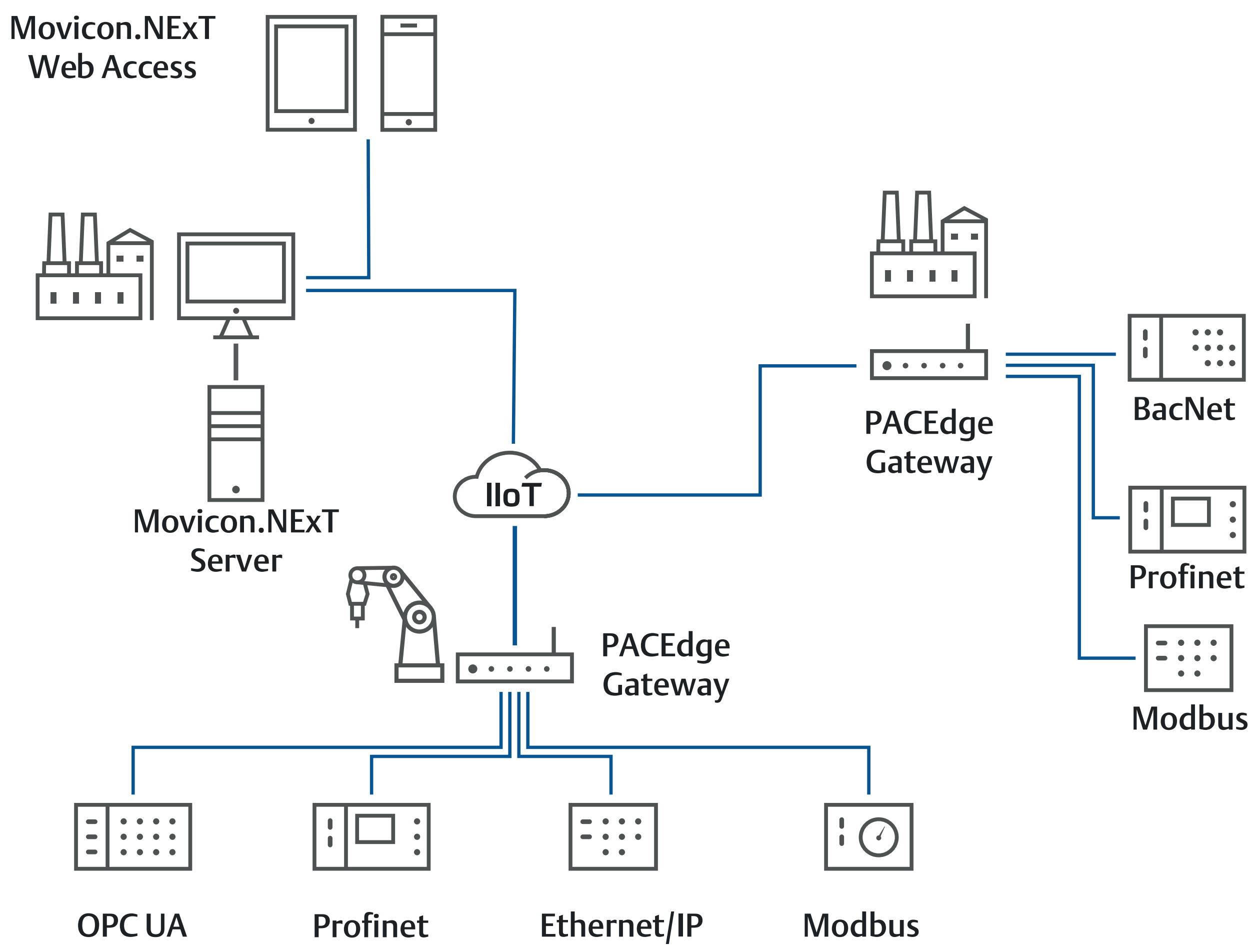
One goal of IIoT initiatives is to transmit this on-site data to cloud-based IT systems for deeper and longer-term analysis than is typically done for production-oriented goals. The variety of data types, source devices, and communications protocols adds to the difficulty of building a comprehensive solution. Following are five categories of stranded data sources:

- Isolated:
  Assets within a facility with no network access to any OT or IT system.
  This is the most straightforward case, but not necessarily the easiest to solve. Consider a standalone temperature transmitter with 4-20mA connectivity or even Modbus capability. It needs to connect with some type of edge device—PLC, edge controller, gateway or other—to make this data stream accessible. In many cases, such data is not critical to machine control, so it is not available through traditional legacy PLC/SCADA data sources. Pulling the data in through the nearest machine PLC risks voiding OEM warranties due to required changes in the programming logic.
- Ignored:
  Assets connected with OT systems and generating data, but that data is not being consumed.
  Many intelligent edge devices provide basic and extended data. A smart power monitor can easily provide basic information like volts, amps, kilowatts, kilowatt-hours and more using hardwired or industrial communication protocols. But deeper data sets, such as total harmonic distortion (THD) may not be transmitted due to lack of application requirements, low bandwidth communications or limited system data storage capacity. The data is there, just never accessed.
- Under-Sampled:
  Assets generating data but sampled at an insufficient data rate.
  Even when a smart device is supplying data to supervisory systems via some type of communication bus, the sampling rate

may be too low, or the latency too great, or the data set is so large so that the results are not obtained in a usable fashion. Sometimes, the data may be summarized before it gets published resulting in a loss of fidelity of the data.

- Inaccessible:
  Assets generating data (often non-process, yet still important for things like diagnostics), but in a generally inaccessible format or not available via traditional industrial systems.
  Some smart devices have on-board data like error logs which may not be communicated via standard bus protocols but nonetheless would be very useful when analyzing events that resulted in downtime.
- Non-Digitalized:
  Personnel generating data manually on paper, clipboards, whiteboards, and the like, which misses the opportunity to capture this information digitally.
  For many operating companies, workers complete test and inspection forms and other similar quality documents in a physical paper format, without any provisions for integrating this information with digital records. A more modern approach would use digital methods to gather this data, leading to a "Paperless Plant."

Operational personnel may have some awareness of stranded data sources. But the reality is that this data is not required as part of their day-to-day operational work of keeping production running. As a result, they really don't proactively try to access this data during their ongoing process improvement initiatives and thus don't recognize its potential.



## Whose Job is the IIoT?

Industrial operations—especially remote locations—typically have limited personnel available for administering and managing specialized OT/IT systems. Even if IIoT projects were attempted, some organizations would store data far longer than OT systems could handle, resulting in a form of data bloating that would negatively impact the primary production-oriented functions of OT computing assets.

Another challenge related to integration between OT and IT revolves around which parties "own" and "need" the data. OT personnel administrate the PLC/HMI/SCADA systems, while enterprise and IT personnel have a greater interest in carrying out IIoT initiatives and have the knowledge to manage large data warehouses or data lake projects.

Companies need a way to coordinate OT and IT efforts. Normally, the experts who understand the IT cloud infrastructure and have good ideas on what to do with available data lack the expertise to locate and interpret OT-sourced data, connect to it, and bring it to the cloud computing platforms. On the other hand, the people that understand the OT data and connectivity challenges well generally don't understand the cloud infrastructure or the potential this infrastructure can have. Systems making both the data access and cloud transfer simple and easy to set up are a critical missing link. While traditional automation platforms with the right features certainly play an important role in IIoT data projects, it is equally important to consider including technologies that can form an IIoT data gathering path parallel to the existing automation.

## Gaining Value from Edge-Sourced Data in the Cloud

Stranded data is of significant interest to data engineers, data scientists, and other subject matter experts looking to analyze operational performance across one or more entire production facilities. This group of people is very interested in finding solutions to transmit this stranded data from the field to the cloud, for logging, visualization, processing, and deeper analysis.

This connectivity, especially to high-level on-site and cloud-based enterprise information technology (IT) systems, is needed so that the many types of edge data can be historized and analyzed to achieve deeper and longer-term analytical results, far beyond what is typically performed for near-term production-oriented goals. When an organization—whether an end user or an OEM—can liberate stranded data from traditional data sources, and transmit this data to cloud-hosted applications and services, many possibilities are enabled:

- Remote monitoring
- AI/ML model training for predictive diagnostics
- Pro-active vs reactive maintenance
- Planning across machines, plants and facilities
- Root-cause analysis
- Long-term data analytics
- Like-for-like asset analysis within and across multiple plants
- Fleet management
- Cross-domain data analysis and analytics (deep learning)

- Insights into production bottlenecks
- Crew schedule efficiency studies
- Performance comparisons across shifts
- Identifying where process defects are originating, even if they are not detected until further in the production process

Solving the challenges of stranded data, and effectively connecting edge data to the cloud where it can be analyzed, is ultimately the purpose and strength of IIoT initiatives. IIoT solutions incorporate hardware technologies in the field, software running at both the edge and the cloud, and communications protocols, all effectively integrated and architected together to securely and efficiently transmit data for analysis and other uses.

## Creating an Edge Solution


Digital transformation is necessary at the edge to liberate all forms of stranded data, and to transport this data to the cloud for analysis. Edge solutions can be an integral part of automation systems, or they can be installed in parallel to monitor data not needed by the automation systems. Many users prefer the latter approach because they can obtain the necessary data without impacting systems that are already operating in production. However, the key is that these new digital capabilities can connect with all previously identified forms of stranded data.

Edge connectivity solutions take many forms; following are a few popular examples:

- Compact or large PLCs ready to connect with industrial PCs (IPCs) running SCADA or edge software suites.
- Edge controllers that are “edge-enabled” running SCADA or edge software suites.
- IPCs running SCADA or edge software suites.

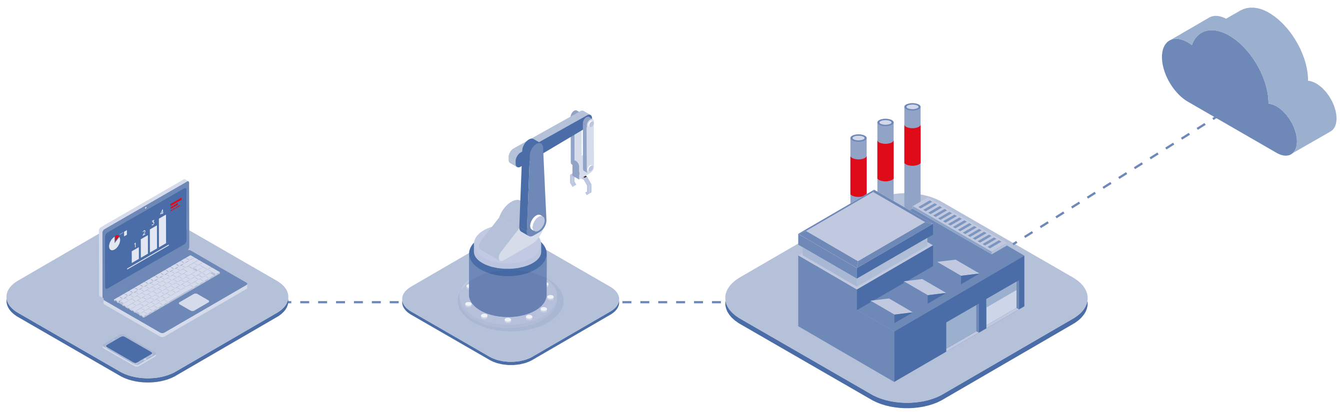
Hardware deployed at the edge may need wired I/O, or industrial communication protocol capabilities, or both, to interact with all sources of edge data. Once the data is obtained, it may need to be pre-processed or at least organized by adding context. Finally, the data must be transmitted to higher-level systems using protocols like MQTT or OPC UA.

For best usability and to minimize any subsequent processing efforts (and errors), OT data must be cleaned, transformed, structured, and conveyed with its context. This means including naming conventions, engineering units, and scaling values so the data stream self-describes the content. A seamless solution is needed to transport data from edge sources to cloud computing resources so the enterprise can fully take advantage by transforming the data into actionable information. Sending individual sensor information directly to the cloud is a first step but is not in itself an optimal method, compared with routing data through a proper OT automation system or a parallel installation of an edge-capable solution that will help provide semantic meaning.

Maintaining context is particularly important in manufacturing environments where there are hundreds or thousands of discrete sensors monitoring and driving mechanical and physical machinery actions. Modern automation software systems help preserve the relative relationships and context. Today’s OT/IT standards are developing in a way that ensures the consistency and future flexibility of data and communications. It is important for any solution to be flexible yet standard-compliant—as opposed to custom setups that will be impossible to maintain long-term.

Once an edge solution is in place and can get the data, the next step is to make it accessible to higher-level IT systems, with seamless communications to cloud-hosted software.

## Connecting Edge to Cloud



Once the right edge hardware and software is in place to connect with all types of data sources, pre-process this information to a degree, and then transfer it in context to the cloud, it is time to examine why the cloud is so compelling for creating value from this data. With clean, structured data in proper context and readily accessible in the cloud, data scientists and other analytics experts can apply big data principles to gain new insights from the data.

The term "cloud computing" refers to the distribution of computer services and IT resources on the internet, such as data storage, processing or transmission. A key point of this technology is the ability to use each of these aspects on demand, individually or in conjunction with one another, all via web access without having to install programs or occupy memory on local devices. Cloud computing is often referred to as an "elastic computing" environment because if more computing or data resources are needed, they can be added in real time as required.

The cloud is a container of data, information and services, accessible at any time from any device in any location connected to an internet infrastructure. The operating logic of the cloud is based on the concept of "access" rather than that of acquisition. In fact, the user pays only for what they use, allowing them to avoid investing in purchasing and managing IT infrastructures, and to focus their professionalism and resources on pursuing their core business-related objectives. The cloud also relieves the user from the problems of configuring IT hardware and software systems, along with deployment, management, performance, security, and updates of the same.

The cloud computing model has radically changed the world of IT and has become a key support factor for both big and small companies. It offers people and businesses the opportunity to take advantage of cutting-edge and constant up-to-date services at a low cost. These services were once unattainable because they were too expensive to purchase and run and were thus limited to large companies and organizations.

Thanks to the cloud, users can now store data on what is now called a virtual hard disk, without having to worry about the threat of losing important information due to local PC or Server problems, or the possibility of running out of storage. Users also gain the ability to share data with authorized people.

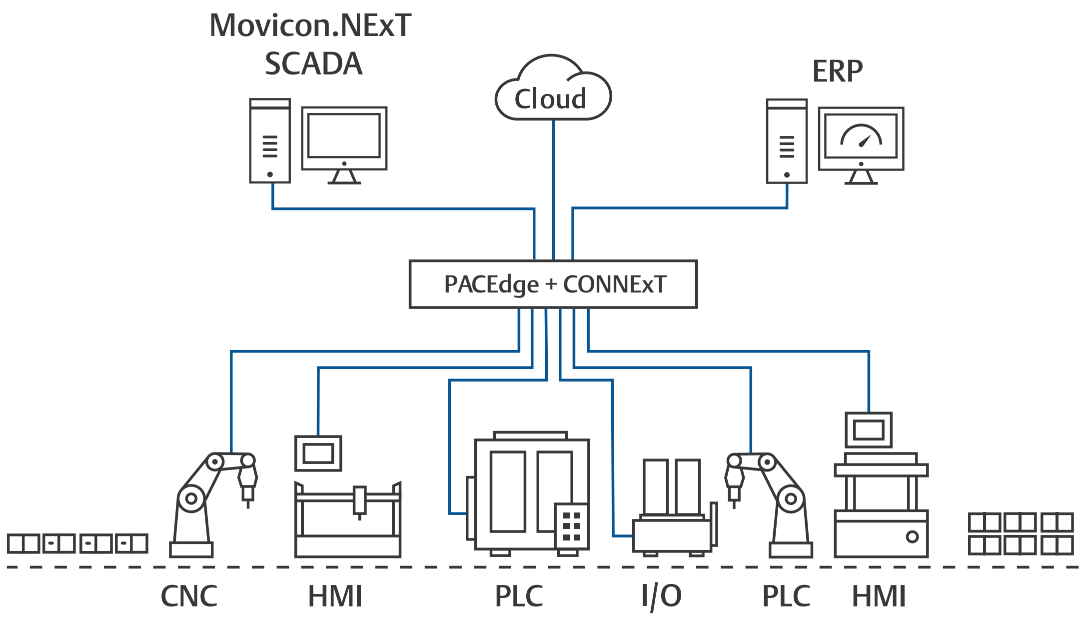
For these and other related reasons, a cloud architecture fits well with the needs of organizations when implementing IIoT data projects. The cloud is the enabling infrastructure of many IIoT projects and the combination of these two technologies enables innovative interactions among humans, objects and machines, giving birth to new business models based on intelligent products and services.

### Pros:

- Reduced costs: pay on demand, saving on software and hardware (purchase, configuration, installation, maintenance and disposal).
- Reduced support: the management of the computer infrastructure is the responsibility of the cloud provider.
- Ability to process big data sets efficiently with CPU processing power scaling based on the needs of the analytics.
- Greater flexibility: easy and timely adaptability of needs-based contractual conditions.
- Greater accessibility and mobility: data can be connected and accessed from wherever and whenever using any device capable of hosting a web browser.
- Greater security: data are kept safe on different servers with backup and disaster recovery options.
- Quicker development: platforms are immediately operative, only an internet connection and access credentials are needed.

### Cons:

- Data storage can be expensive. The cloud models work best when the overall infrastructure actually reduces the storage and capacity requirements needed locally at the individual plants. For example, if the site only stores 3-6 months of data while the cloud can store 10 years of data, the manufacturing systems' infrastructure on site will run faster, more efficiently, and cost less than if each plant had to store the same 10 years' worth of data locally.
- Can't be used for mission critical operational requirements as connectivity and availability cannot always be assured even with a high-performance service level agreement (SLA).
- Because cloud solutions rely on data residing on servers of the chosen provider's infrastructure, it is not always possible to have local backup copies of data in the event the cloud infrastructure goes down or is not accessible.
- Bandwidth can become costly when connecting to remote sites, so processing data locally and sending only what is needed to the cloud may be necessary to reduce the overall cost of implementation.




## How Movicon and PACEdge Bridge Data Between the Edge and the Cloud

IIoT connectivity solutions come in many sizes; indeed, there is great flexibility in choosing a right-sized implementation for each application. However, one particularly compelling solution uses Emerson's Movicon software on a variety of hardware platforms as a bridge between edge data and the cloud.

Emerson's Movicon.NExT™ platform brings a modular approach to connectivity, supervisory control, HMI, SCADA, IIoT, and plant analytics. It goes beyond traditional HMI/SCADA and IIoT solutions by offering extensive openness, scalability, security, and integrated connectivity. Depending on the need, the software can be hosted on an edge controller, a site-located PC or server, or a cloud computing resource. Because the software runs on Linux or Windows, it is possible to scale application deployments from small, embedded edge devices up to larger server-based systems.

Emerson's Movicon Connext products are enabled with standard drivers to connect with virtually any device in the automation, infrastructure and process control sectors. Built-in data routing and gateway functions make it easy to supply this data to cloud and other IT systems. The server offers numerous communication protocols and is easy to configure and manage connectivity, enabling creation of data collection systems capable of connecting to any enterprise-wide system. This technology makes it easy to securely collect and publish data on the cloud, manage business information flows towards ERP/MES business managerial systems, or simply connect field devices to software applications. Movicon works with many cloud provider offerings through standards adopted by both IT and OT users.

When required, Movicon can also log data directly to cloud databases such as Microsoft's SQL Azure. The Advanced Visualization and Web Clients connect directly to both OT and IT data sources providing an integrated experience to end users regardless of whether they are on the IT or OT side.

Emerson's PACEdge products provide another IIoT connectivity option that runs in parallel to existing PLC / SCADA systems. The Emerson PACEdge application enablement platform facilitates the execution of data gathering, analytics and more, running on edge hardware. With the PACEdge multi-container Docker infrastructure, a graphical development environment, and flow-based programming tools, users can easily create digital transformation projects for monitoring machine health and conditions, energy efficiency tracking, throughput improvements, and much more. PACEdge is natively compatible with common IIoT protocols and can be deployed on any class of edge hardware, providing complete flexibility.

## Modern Automation Software Connects OT and IT

Stranded data is an all-too-common reality at manufacturing sites and production facilities everywhere. It is the unfortunate result of legacy technologies incapable of handling the data, and traditional design philosophies focusing on basic functionality at the expense of data connectivity. Only recently has the importance and value of big data analytics become mainstream, so end users are working to build this capability into new systems and add it to existing operations. Edge-to-cloud data connectivity delivers value in many forms of visualization, logging, processing, and deeper analysis.

Any IIoT solution for bridging data between OT and IT relies on digital capabilities that can interface with traditional automation elements like PLCs, or can connect directly to the data sources in parallel to any existing systems. These edge resources must be able to pre-process the data to a degree and add context, and then transmit it up to cloud systems for further analysis.

Emerson's Movicon.NExT, Connext, and PACEdge products running on a range of hardware options are a flexible and scalable way of creating SCADA and IIoT solutions of all sizes. Users can deploy Movicon products where needed to connect with automation systems and other intelligent assets and route the data to cloud databases and other cloud-based resources.

For more information, and to find out how Emerson's Movicon products can help with your application:

**Download the Application Note: Emerson Movicon™ Bringing Data to Microsoft SQL Azure for Data Management and Analysis in the Cloud**

**Download the Movicon.NExT free trial demo.**

**Contact our experts for a consultation.**

**Contact Us**
**www.emerson.com/contactus**

EMERSON™