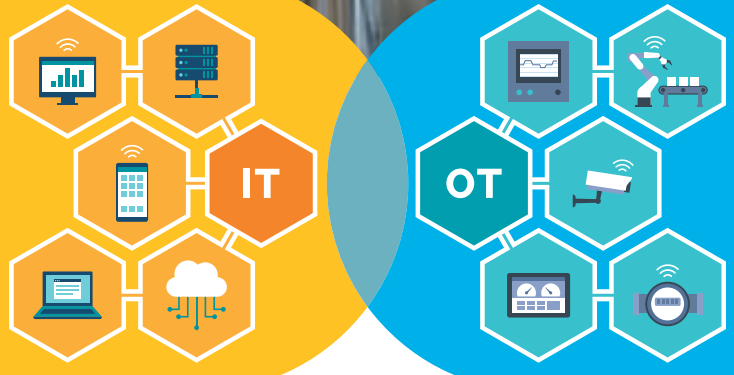
IT-OT Convergence: Trends and Opportunities

A study to understand where industrial companies are on their digital transformation journey



CFE Media and Technology™

Sponsored by EMERSON™

Figure 1: With the advent of industrial internet of things (IIoT), companies are increasingly connecting and creating new value and insights from their control system data.

Many IT and OT applications currently in service around the world rely on process control systems as sources of critical operational data input. This requires control systems to collect, aggregate, and generate large amounts of data for use beyond performing the controls. Companies are eager to connect and create value from their control system data, but require secure, streamlined,and contextualized access to do so.

Industrial process control systems are tasked first and foremost with providing safe, reliable, and efficient production for operating plants of all types. An additional benefit of these operations technology (OT) control systems is the volumes of valuable data they produce associated with process automation, field sensors, and devices. This includes real-time and historical process data, user changes, alarm data, batch information, system/device configuration, diagnostics, and much more.

Many end users access this data today for use in external systems deployed in the IT realm, such as a manufacturing execution system (MES), enterprise resource planning (ERP) system, and other data science applications, enabling the users to act for optimizing their operations.

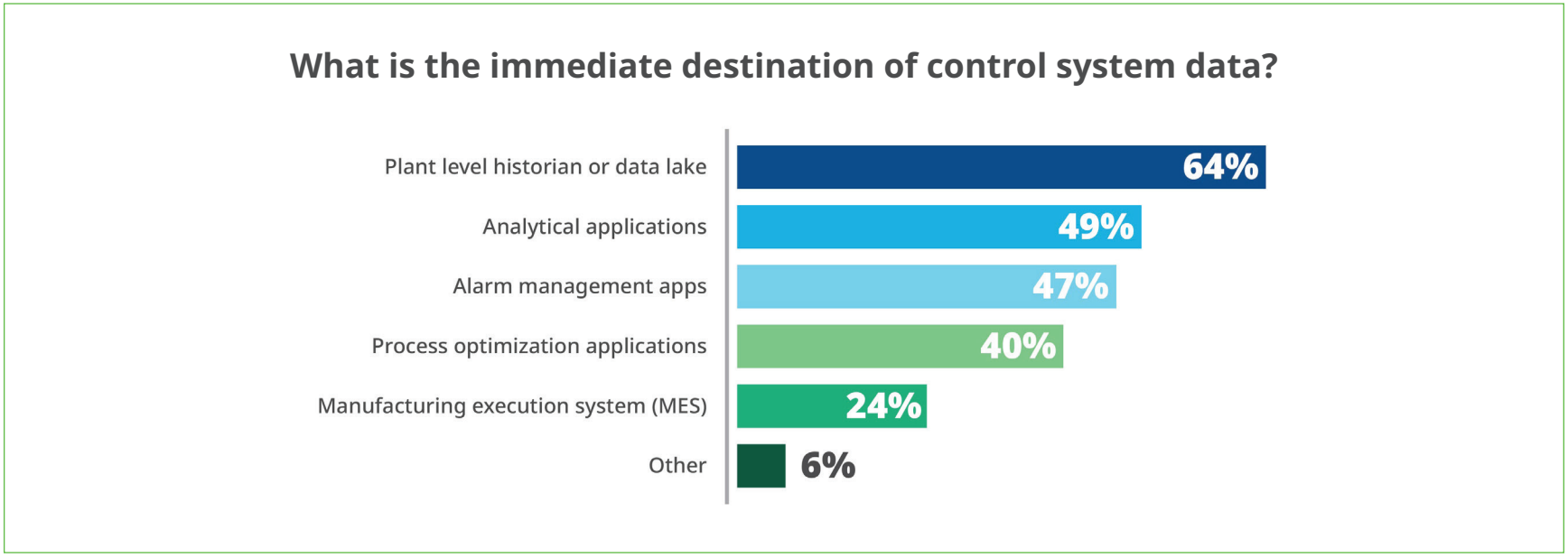


Figure 2: Typical destinations for data sourced from control systems.

As end users proceed on their digital transformation journey, they will need to harvest ever greater amounts of data and perform much more complex operations with it. To understand where end users are today along their digital transformation journey, what control system data is most important to them, and their current challenges, a study was conducted by Control Engineering on behalf of Emerson Automation Solutions to identify the types of data most often extracted from control systems, challenges with using and managing this data, and future plans for migrating data and associated tasks to the cloud.

For many years, accessing control system data relied primarily on OPC—first with OPC Classic, and more recently with OPC UA—to provide connectivity to the process data. However, the survey showed that a significant number of users also utilize other methods such as direct data queries and database file connections. This is likely because it isn't always possible to connect with all the desired source data using OPC. The non-OPC methods can be intrusive, putting control system operation at risk, they provide different 'silos' of data, and are difficult to secure. In fact, users highlighted that their top challenge was security, followed by the lack of context for the data.

With the advent of industrial internet of things (IIoT) technologies increasing need for edge analytics, deep analytics, and other advanced initiatives, end users require even greater data access and connectivity so they can gain insights into their operations. In the survey, users responded that typical data destinations for control system data are plant/enterprise-level historians, on-site or cloud-based data lakes, visualization software, analytical applications, and alarm management systems (*Figure 2*).

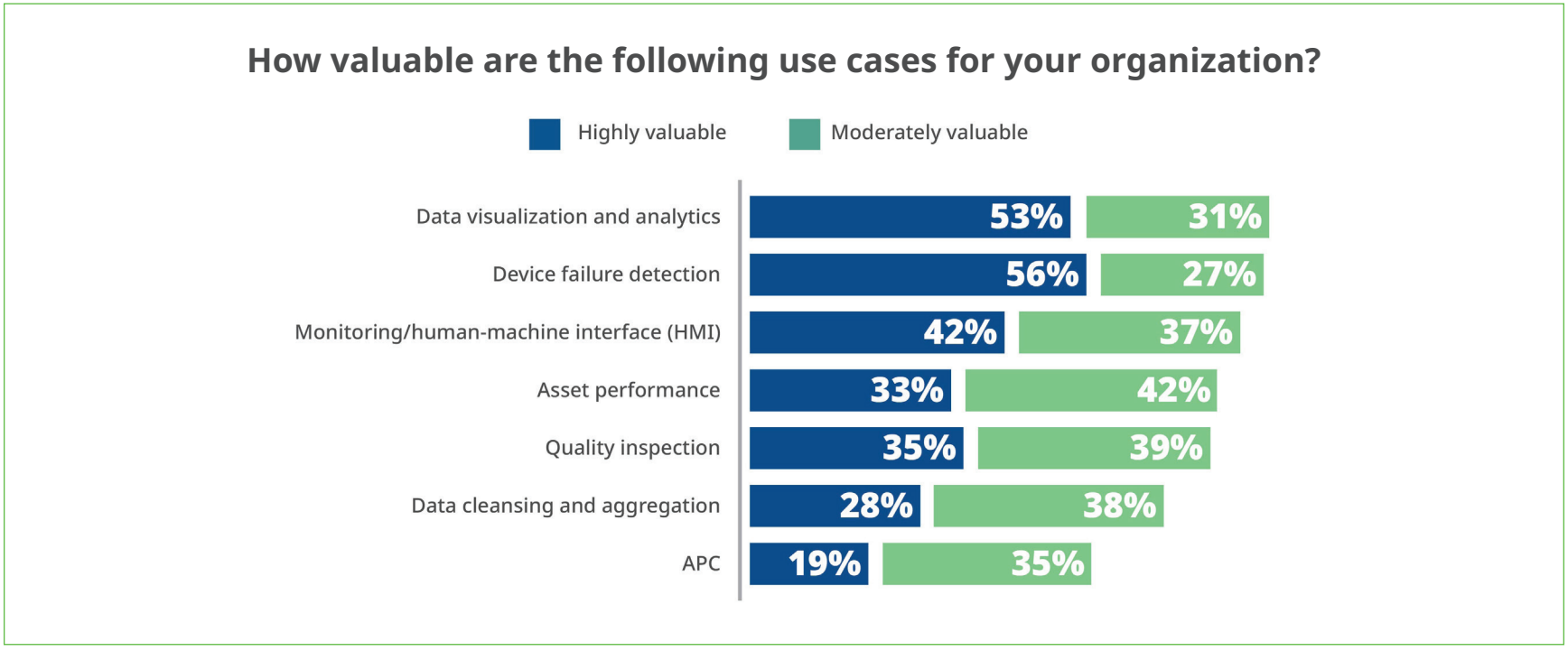


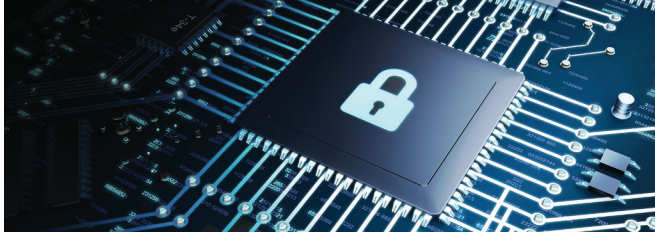
Figure 3: Most valuable analytics use cases.

End users want to connect with all of their OT source data securely and easily, and make it readily available for IT access with full context. However, 50% of respondents said that their companies will not allow cloud-based solutions to directly access their control systems, presenting further challenges. This ebook examines trends and user needs for IT-OT convergence on security, data, and context perspectives, and then explores potential solutions.

Delivering the data users need

Per the survey, external access to process control system data sources—especially to support data visualization and analytics, device failure detection, and process monitoring—is essential for supporting an organization's digital transformation strategy (*Figure 3*).

Each application requires defining what control system data needs to be accessed and ensuring the quality of that data. This means that the amount of data needed from the control system increases as more complex analytics are added. When ensuring data quality is important, this often requires doubling the amount of process values—for example using both the actual value and the associated value status—and is dependent on whether quality data from the control system is available.



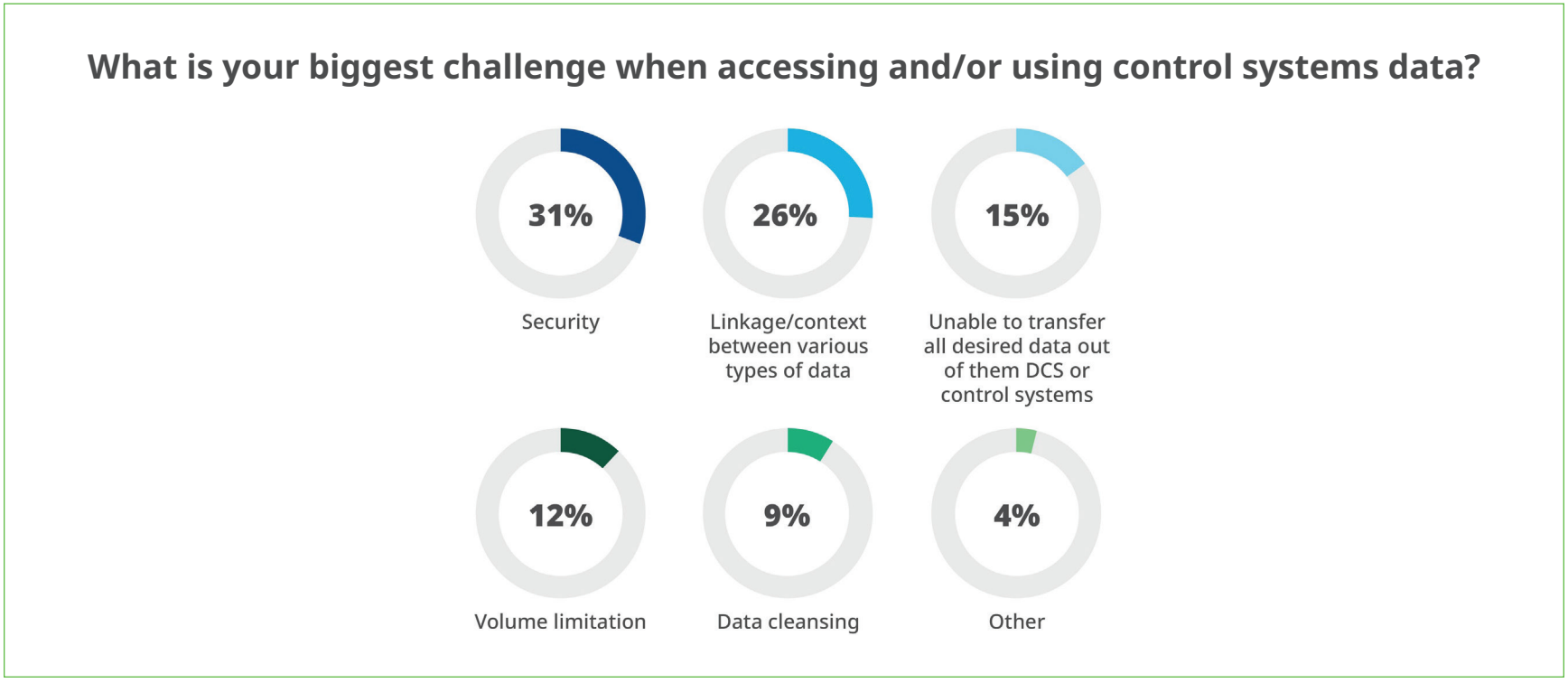
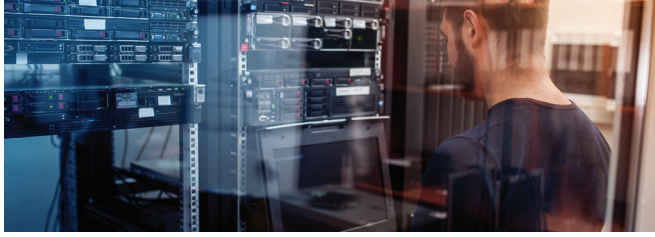


Figure 4: End users are faced with many challenges when accessing and/or using control system data.

In many cases, data flow and usage remain entirely on-site. However, users are increasingly evaluating the advantages of moving some or all these functions beyond the enterprise level network to the cloud, which will demand even greater connectivity and security capabilities.

There are many needs and possibilities for process control system data connectivity, but streamlined access can be difficult, risky, and/or expensive to create, configure, and maintain for many reasons. In some cases, the data may be impacted by unacceptable latency or fidelity. *Figure 4* shows some leading end user challenges.

To address these and other issues, and as part of a comprehensive on-premises or cloud-connected solution, users are tasked with addressing networking issues such as creating a DMZ, ensuring a properly configured firewall is in place, defining ports, adhering to network policies, creating and maintaining usernames and credentials, and more. A complete solution can consume months of planning and implementation, and users are burdened with ongoing updates and maintenance.



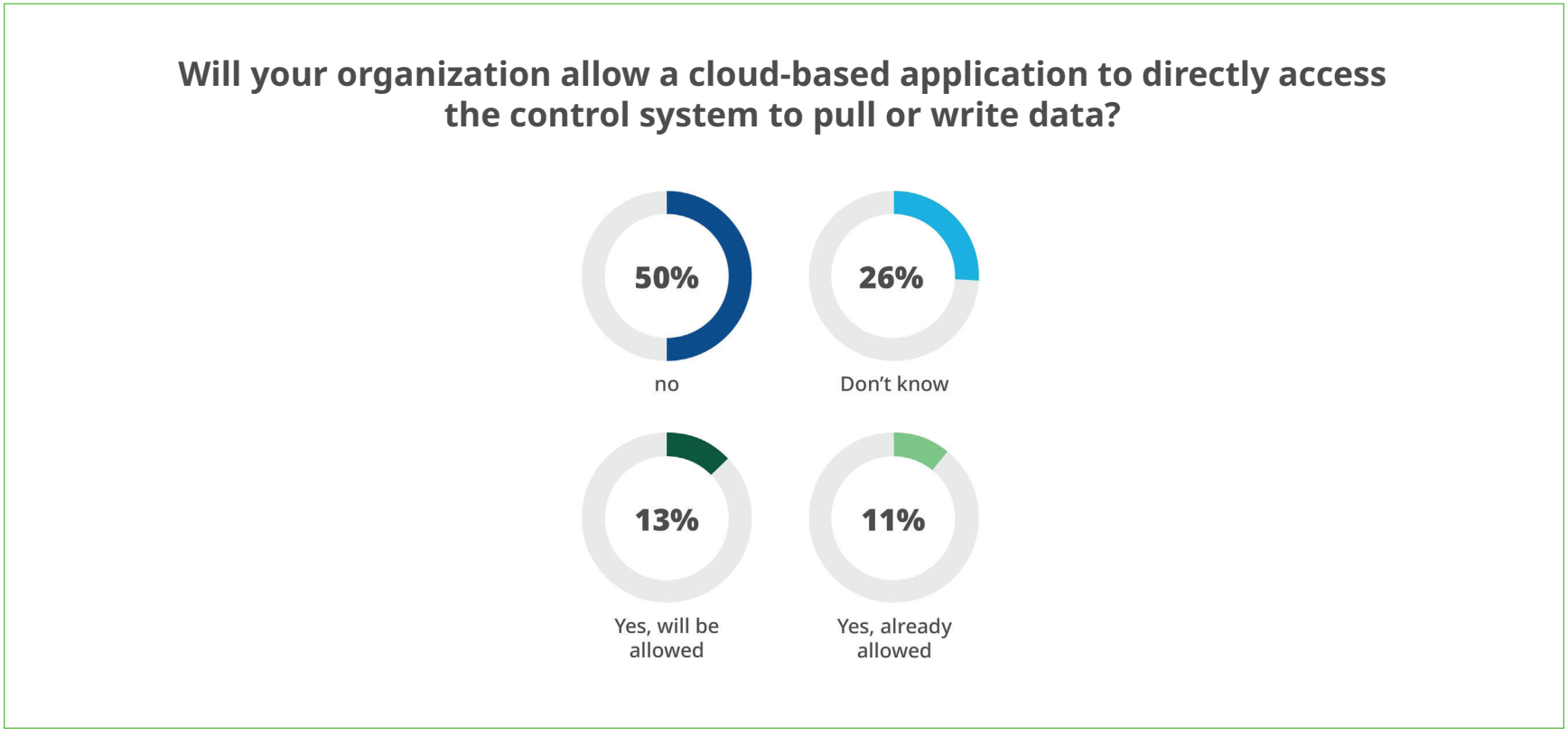
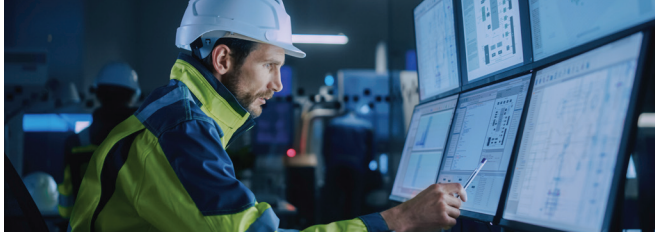


Figure 5: Most control systems do not directly connect to cloud-based applications.

Often, proving the complete compatibility and security of a custom approach can be difficult or impossible. Due to this and other difficulties, only 24% of those surveyed would allow cloud-based applications to directly access their process control system for the purposes of reading or writing data, 50% do not (*Figure 5*).

Another consideration regards exactly what types of data can be accessed, and the quality and context of these data. The most common data accessed today are alarms and events, followed closely by "live" process values such as temperatures, pressures, flows, and the like. But the reality is that many other data types exist, with many required by various data analytical applications.

A process control system includes numerous data types contained in various locations and formats. Around half of users in the survey also access diagnostic and configuration data. These values are applicable to continuous and batch users alike. In addition, batch processors will also need access to recipe and sequence records.



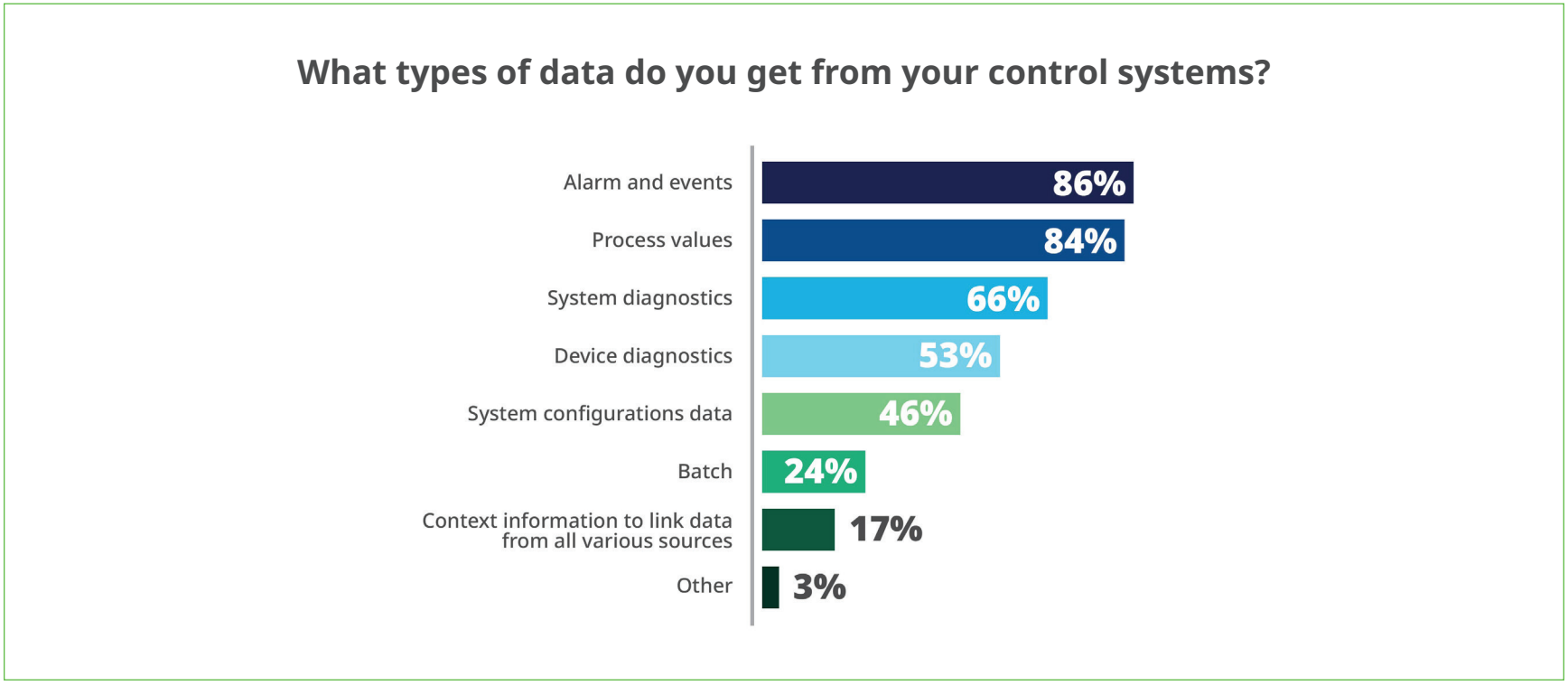


Figure 6: Users require many types of source data from control systems, accessible in a useful format with full context.

When multiple types of data are accessed from a control system, it can be useful to correlate them together—to understand the data in context. However, accessing and applying data context, which is embedded in the control system configuration, can be difficult and adds another layer of complexity.

Figure 6 shows the leading types of data needed as part of a data replica.

To access this and other data, most process control systems support OPC UA, a secure and high-performance industry-standard protocol, enabling users to access certain process data, along with alarms and events (*Figure 7*).

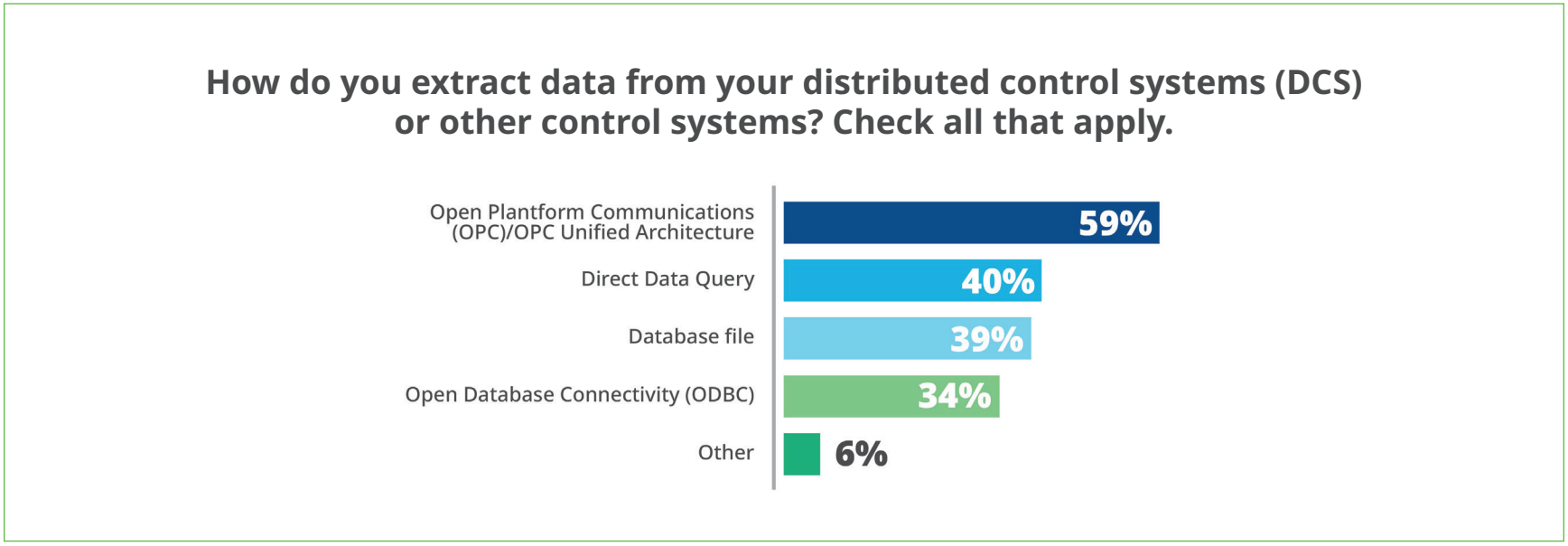


Figure 7: Commonly used ways to access control system data.

Indeed, the survey indicates that well over half of end users (59%) take advantage of OPC UA, while others use means such as direct data queries (40%), database access (39%), and other methods.

Although users have found ways to connect with their different types of control system data, these non-OPC methods can be intrusive, risk control system operation, are difficult to secure, and provide different 'silos' of data.

While the limited availability of certain data may be acceptable today, there is generally a growing demand from users for increased external access to all these data sources, and more, in a contextualized format.

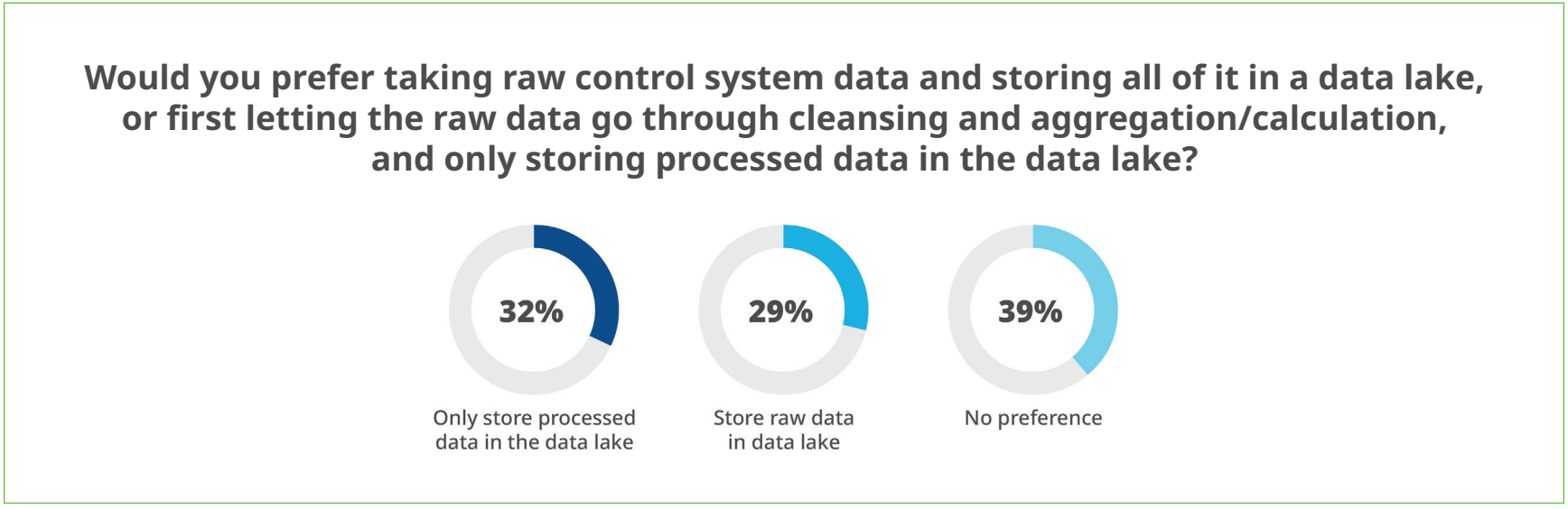


Figure 8: Raw system data vs storing data lake.

The data dilemma

One consequence of the integrated nature of a modern industrial-grade process control systems is that some of the desired data are not readily available for external consumption. Several factors contribute to this situation.

By design, a process control system must ensure uninterrupted performance by avoiding resource-heavy communications activities that could compromise operation. Because the process control system is critical for plant operations, its security is continuously enhanced and updated to defend against malicious access, but this in turn can make it more difficult for external applications to access data.

Furthermore, new industrial standards and devices keep adding data types and data models into a process control system, which challenges standard data transportation protocols.

Increasingly, larger amounts of process control system data are needed to provide real-time insights, guide localized actions, or fulfill regulatory requirements. The data demand in terms of volume and velocity can strain network performance and complicate security setup. For example, a medium sized control system can generate millions of data points in each second. It often requires premium computing and network resources to transport, process, and store those data with satisfactory performance.

Adding to the dilemma is how users want to access, use, and store the data versus how the data is presented to users. For survey respondents considering the transmission of process control system data into a data lake, about one-third would prefer data to be cleansed, aggregated, and pre-processed so only consolidated, and potentially easier-to-use, results are stored.

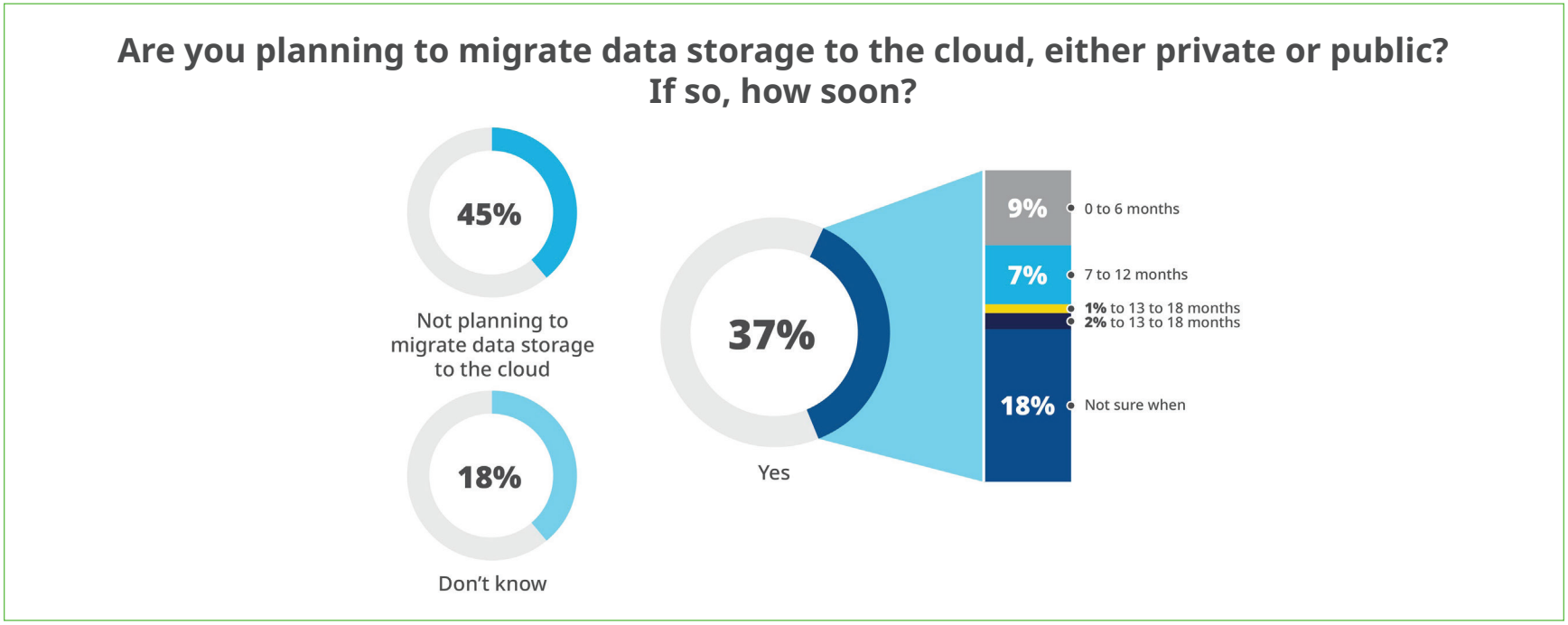


Figure 9: Expectations for migrating storage to the cloud.

Although cloud-based applications today may not be allowed to directly access the process control system, 37% of users are planning to move data storage to either public or private clouds, with half of them planning to move storage in the next 2 years. This will demand even greater connectivity and security capabilities (*Figure 9*).

A recent Gartner IoT study (Reference 1) indicates users have started to realize that not all data needs to be sent to the cloud or a core data center, which would be cost-prohibitive, bandwidth intensive, and likely to cause performance implications. Instead, many think it is beneficial to aggregate data generated from source devices, and normalize it prior to transmitting it to analytics platforms, whether on-premises on cloud-based.

When it comes to external applications accessing process control systems, most require data with context because siloed data output has limited value. For instance, an application might need to identify what alarms happened during a specific process state or batch, and what the associated process values were during the alarm periods. If only some of these values are available, or if they are all available but not related in any way,

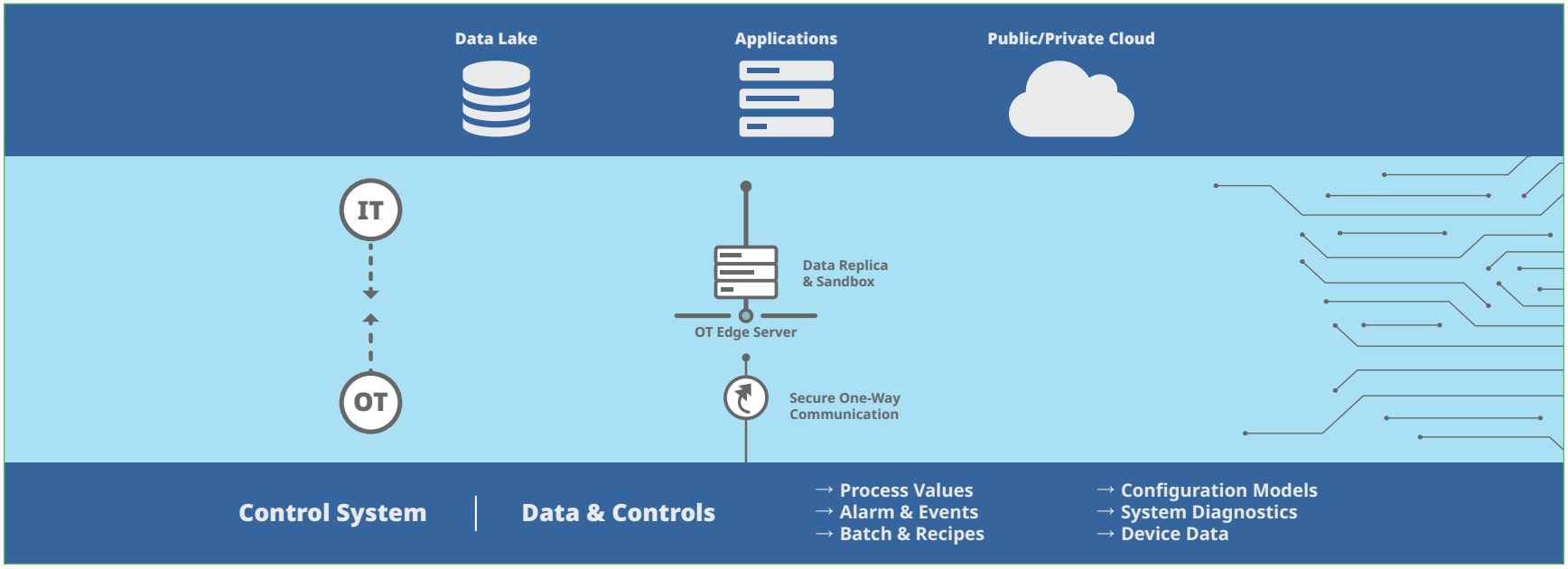


Figure 10: To satisfy end user needs for easy and secure control system data access, a data replica deployed at the OT edge and protected by a data diode can be an advantageous solution.

it is very hard to create insights from the data. It can be a complex endeavor to externally relate multiple data streams. Considering that not all the data is even available, it would still pose an unacceptable security and performance risk to attempt unpoliced external calls to access all available data. In extreme cases, this effort might provide time-delayed results lacking in fidelity, and it could even affect process control system operation.

While it may be possible to simply store certain raw data into external data lakes, the limits of what is available and the lack of context make this approach very hard to put into practical use.

Therefore, many users are looking for a more comprehensive option that provides contextualized data but does not increase engineering efforts. In addition, they also want:

- Consistent data structures throughout the system.
- Reliable and pre-filtered data streams that are free from error.
- Condition monitoring of equipment to enable predictive maintenance.
- Historical data access supporting troubleshooting and defect recognition.

Many users are not planning to migrate data storage to the cloud any time soon, but they still want improved on-site access to process control system data. In addition, 37% of those surveyed (*Figure 9*) are making some plans to migrate data storage to either a private or public cloud, and they will need improved connectivity options to do so.

Solving so many data access needs will require a unique approach addressing all these user concerns, and more.

A potential IT-OT data solution

One potential solution to this data dilemma is to separate a process control system's control functionality from its data provision functionality, enabling these two threads to run independently, yet with synchronized data (*Figure 10*).

This approach provides a number of advantages.

First, the mission-critical process control function is protected from unauthorized access. Equally important, the process control system is unburdened from servicing a wide variety of data requests, some of which could potentially be duplicates. Instead, it handles streamlined data access in the most efficient manner possible.

The data provision functionality streams large amounts of data needed for users' digital transformation strategies to a repository located at OT edge which acts as an 'onsite data replica' of the control system. The data replica can provide data access for applications from the enterprise-level network or the cloud. As an intrinsic control system data replica, the data provided will be comprehensive, contextual, and trustworthy. This eliminates the need for applications to directly access the process control system, reducing intrusion risks. Data provision enables on-site resources to perform big data analytics close to the source where timely responses can be taken.

Due to the large volume of control system data and the complex contextualization needs, a dedicated on-site edge data server would provide the ideal computing platform for filling the IT-OT edge gateway role. In addition to providing data, this server should be open to hosting popular analytical, calculation, or visualization applications. These can be developed by users or obtained through an application marketplace.

This architecture provides a desirable data sandbox functionality for users to explore, test, and try on the data for customized purposes, in addition to the already-poplular data cleansing and data aggregation needs.
Secure one-way data communication methods

could be used between the process control system and the data provision IT-OT edge gateway. An approach gaining traction in the process control industry is to address connectivity security concerns with the use of a data diode. A data diode provides one-way data communication of clearly defined data from the process control system thus guaranteeing secure communications. In this case, the control system automatically sends the pre-defined data, instead of waiting for a request from an external application.

A final point is that any IT-OT edge gateway system needs to be conceived as a complete solution, optimized for "low-touch" configuration by users. A complicated arrangement of do-it-yourself software requiring extensive IT experience and knowledge would be unmanageable at best, and could compromise performance and security at worst. It should be easy for OT personnel to perform worry-free deployment, setup, support, and maintenance of the IT-OT edge gateway, preferably using cloud-based access. One increasingly adopted approach is to use containers technologies.

In summary, any future data solution must be designed from its inception to bridge the divide between an OT-centric process control system and the multitude of IT-focused applications and data destinations, doing so in a secure and high-performance manner.

An example of a data provision IT-OT edge gateway in action helps describe the benefits.

A connected reality

Emerson has long provided the robust automation solutions end users need to run efficiently and safely. As end users progress along their digital transformation journeys, they are looking for more expansive data connectivity to support external databases and applications, and these need data far beyond basic process values, accompanied by much enhanced security. An effective IT-OT edge gateway will make it easy for end users to fulfill growing data access requirements via a simple infrastructure, providing secure and contextualized OT edge data for any external need.

References

1. Gartner, Market Guide for Edge Computing Solutions for Industrial IoT, by Santhosh Rao. **https://www.gartner.com/doc/reprints?id=1-28QJ8Y9N&ct=220112&st=sb**

Emerson
North America
1100 W. Louis Henna Blvd.
Round Rock, TX 78681-7430,
United States
+1 800 833 8314

Latin America
1300 Concord Terrace Suite
400 Sunrise, Florida 33323,
United States
+1 954 846 5030
+1 954 846 512

Europe
Blegistrasse 21
6341 Baar,
Switzerland
+41 41 768 61 11
+41 41 761 87 40

Asia/Australia
1 Pandan Crescent
Singapore 128461
+65 6777 8211
+65 6777 0947

Middle East/Africa
PO Box 17033
Jebel Ali Free Zone - South 2
Dubai, United Arab Emirates
+971 4 811 8100
+971 4 886 5465

www.emerson.com

EMERSON