Extending Secure Data Connectivity across existing Industrial Control Systems into the Cloud

Published: October 2022

Abstract

In this Whitepaper, we are uncovering the potential of connecting existing Industrial Control Systems (ICS) to the Cloud for further data processing in the Cloud.

In modern OT scenarios, we often find restricted networking environments based on the Purdue Model (ISA-95) prohibiting any outbound or inbound connections. With the help of Edge technologies and Secure Data Connectivity, Automation Engineers can ensure a secure connectivity into the Cloud without compromising security standards on the shopfloor. By implementing a common Data Model, data from the OT and shopfloor can be harmonized and validated so Data Scientists, Machine-Learning Engineers and Cloud Developers can interact with the data efficiently.

This whitepaper includes practical guidance on how to implement a secure and reliable Data Connectivity solution based on Moxa's industrial hardware and Microsoft's Edge solutions in order to gain actionable insights from Industrial Environments.

This paper is written for Automation Engineers, Security Architects, Technical Decision makers and Project leads who are faced with the task of planning and implementing data-driven applications inside and across Industrial Control Systems.

Contents

Executive Summary

Work on ISA-95 began around the year 1995. So, Automation Engineers have almost 30 years of experience about how to exchange information between the Enterprise Business Systems and Manufacturing Operations Systems in an enterprise. In recent years security standards have been added, most notably IEC-62443.

In practice this has led to a sophisticated and highly efficient network of machines, sensors and IT systems like SCADA and MES. Industrial IoT and Industry 4.0 are now breaking the strict rules of mostly hierarchical data exchange of the Purdue Enterprise Reference Architecture (PERA) used in ISA-95.

The challenge for both IT and Automation Engineers is now to leverage the existing network infrastructure and securely integrate new data paths and without compromising security.

The introduction of an IIoT zone can help to explicitly establish boundaries of trust between the parts of the IIoT solution and the existing Industrial Control System and its existing network. Communication policies between the IIoT zone and the existing network zone can carefully guard the traffic and data paths when introducing new gateways into the existing network. In particular for parts of the network that should not directly connect to the Internet, Microsoft's Azure IoT framework offers Nest Edge, a way to link devices in lower levels of the Purdue Model to Edge Gateways higher up in the hierarchy, so only the top gateway needs to have a direct connection to the Cloud.

Moxa's portfolio contains both Industrial IoT Edge Gateways with Azure IoT Edge pre-installed and the switches, firewalls and routers to create a secure industrial network. In chapter 4, this paper covers how IIoT Edge Gateways with Azure IoT Edge pre-installed can be introduced into an existing IEC-62443-compliant industrial network in a practical yet highly secure way.

Conflicting goals and how to reconcile them between IT and OT are described as well as operational and logistical challenges when rolling out large-scale IIoT applications, so the benefits. The Use Cases for IIoT and Edge Computing are humongous and can help companies to increase production output, lower energy consumption, reduce CO2 emissions and save money. Security and network constraints should not limit companies to leverage the opportunities of what can be achieved using Cloud and Edge technologies but should rather support and align with the Digital Transformation of the whole company.

1. Digital Technology in Industrial Environments

Digital Transformation has been moving into every industry in recent years. While initially point solutions dominated the market, increasingly IIoT is becoming part of an upgraded industrial IT & communications infrastructure that is changing both the way traditional Automation Systems and industrial networks are designed as well as how data is being handled for modern data-driven applications.

Like it or not, digital technology in Industrial Environments is here to stay.

IT departments have long been suspicious of Automation Applications not managed by IT and now are moving in to clean up the perceived mess – creating frustration with the automation colleagues.

But what specifically are the issues when adding digital technology in Industrial Environments? Automation Engineers, called OT engineers by IT, involved in Digital Transformation projects are often confronted with some of the following issues:

- **Productivity:** OT's primary goal is to keep the OT environment up and running, and fix problems fast. If new IIoT zones and Edge infrastructure are introduced into the network, existing and well-oiled business processes cannot be suspended or stopped at any given time. Also, for every minute the company does not produce, they lose money. How can be ensured that production can continue while at the same time introducing new infrastructure?
- **Reliability:** New devices interacting with existing PLCs and Control Systems may cause those to fail or work less reliably.
- **Safety:** Some of these changes may in fact jeopardize human health and life: what if a PLC or ICS system is busy sending data to the Cloud and misses to respond in time to a life-threatening situation for a shopfloor worker?

In addition, IT departments come with a whole new set of priorities and requirements to software that used to be maintained by the Automation Department alone – many of them related to good Cyber Security practices:

- **Security:** Adding IIoT devices into an industrial network increases the attack surface. For the last 30 years, companies have created highly reliable and secure, closed network environment for their OT space to protect their critical machine assets. In addition, these OT environments can execute full end-to-end production lifecycles independently and very reliably disconnected from the Internet. With the rise of data as the new oil, it also becomes apparent that these network setups limit companies to extract data of their machines to analyze and optimize their processes further, on a more global level. OT personnel is now faced with ideas to onboard new devices into the OT network that should be capable of sending data outside of the perimeter, mainly into the Cloud. How can be ensured that security is still a top priority, and no harm can enter the network from the outside?
- **Application Lifecycle:** IT departments are asking for a professional software update process. OT environments have been very constraint with what applications are actually allowed to run on the shopfloor. These applications usually fulfill a very specific task for a production machine or for a workflow specific to the environment it is running in. What this means is that these applications are seldomly changed and updated requiring sometimes very old versions of Windows or Linux to be able to run. While this approach might work in a completely closed off infrastructure environment, it also creates a big security risk when it comes to adopting modern application lifecycle principles like DevOps. On top, these applications cannot always be swapped out by newer versions as they could have specific dependencies on driver components, protocols, or other interfaces. How can these older applications also be integrated into a more modern application lifecycle process? Can these applications also be modernized, at least some of them and participate in the new developments of developing application software for OT?
- **Data Management:** IT departments are taking control of OT Data Infrastructure Management. Standardizing IT processes is one of the top priorities of many companies. One solution to the standardizing efforts is to establish a deployment workflow with the help of DevOps Tools and processes. This would also include the Edge infrastructure machines and other physical assets like IPCs on the shopfloor. In the past, the OT infrastructure had their own processes and tools to update software, often manually due to the closed down network. How can the new infrastructure be included into the standardization efforts?

- **Interoperability:** Industrial Control Systems are highly efficient and very reliable. At the same time, a lot of the Industrial Control Systems implement their own protocol or custom logic. In particular for more complex environments with many different sub-systems and a large variety of data sources, it is important to create common operating environment for the new set of Edge applications. In this common requirement, systems need to communicate using standard protocols, common Data Models, and interfaces to integrate with other systems, such as the Cloud or other Edge infrastructures.
- **"Y architecture":** While IIoT applications need data from industrial assets, existing control systems need to continue to work without interruptions. How can ICS and IIOT applications share data effectively?

Therefore, companies that are attempting to reap the benefits of Big Data, AI, Edge Computing and IoT are struggling to complete IoT projects in time and in budget, due to the complexity of the task.

When then the Business Case is also not clearly defined and supported by management, it is really no surprise that according to the "IoT Signals report Edition 3" curated by Microsoft (October 2021), 90% of companies worldwide are already "IoT adopters", but at the same time, only 25% of projects overall (not only limited to IIoT projects) are rolled out productively and in use.

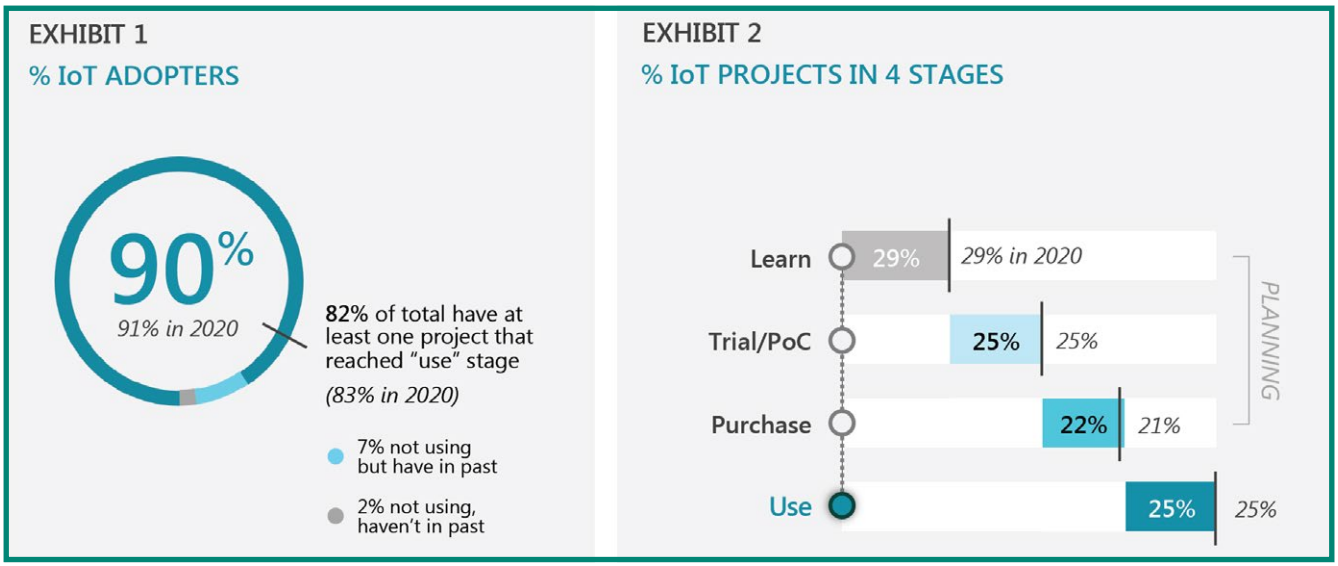

Figure 1: IoT Adopters and Productive Projects by IoT Signals Report 3rd Edition

One of the main challenges depicted in the report, next to the technical challenges we just named a few of, is lack of knowledge. Coming from different backgrounds and environments, OT and IT personnel often find it challenging to understand the challenges of one another when trying to implement an IoT project.

Consequently, OT engineers probably experience adding additional infrastructure into the network as dangerous. Rightly so – they are responsible for keeping the factory running. In the following chapters we would like to explain some of the concepts in more detail to help IT and OT better demystify the implications of realizing an IIoT project including Edge components.

2. Cybersecurity in Industrial Control Systems

For more than three decades, the same communication technologies and concepts have been used in the industrial segment as in systems in the Enterprise or Consumer Segment. Nevertheless, there are so many different implementations in this segment that it has become popular in recent years to divide it into an Information Technology (IT) and an Operational Technology (OT) part. The OT segment has generally been defined as "...hardware and software that detects or causes a change, through the direct monitoring and/or control of industrial equipment, assets, processes and events." [1] While this definition is good for understanding what an OT system is about, it lacks an explicit distinction from the IT system.

The Perdue Enterprise Reference Architecture (PERA), developed in the 1990s, can provide a more precise segmentation. As shown in Figure 2, the functions of an enterprise are divided into levels, starting with the process network at level 0 and ending with business planning at level 4. Although there are different interpretations of the high functional levels, the plant DMZ – typically at level 3 – can be seen as a simplified functional boundary defining OT for all levels below and IT at level 4.

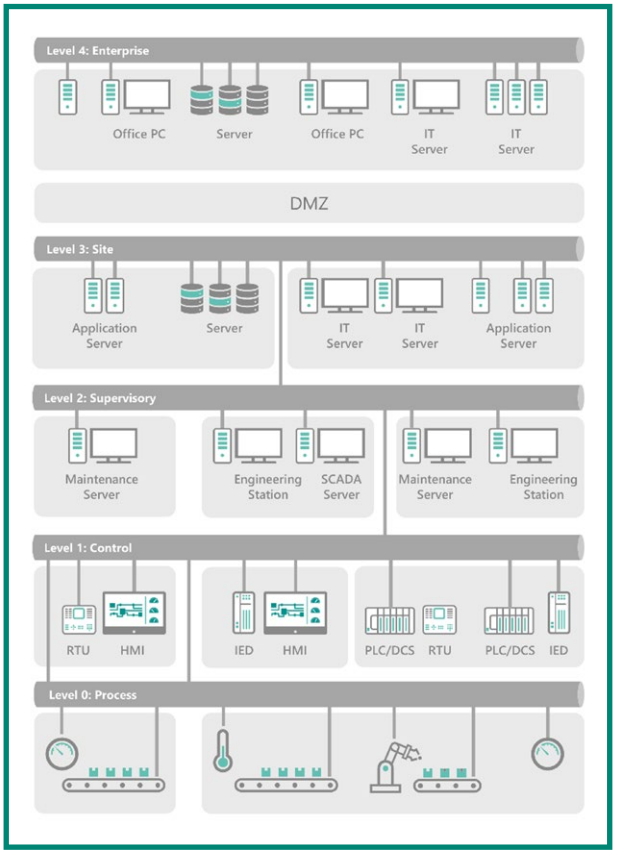

Figure 2: PERA Model

The division of enterprise operations into IT and OT is essentially due to the different objects that IT and OT managers must take care of. While IT professionals ensure the Confidentiality, Integrity and Availability (CIA) of data, OT engineers and operators look after the security, reliability and productivity of assets. The resulting OT triad differs from IT's well-known CIA triad primarily in the physical nature of these assets. Bridging between these different mindsets is daily business in various industry segments: Take Energy: For safety reasons, primary substation equipment must respond in real time to system faults, while the ever-growing share of renewable energy threatens the power grid with unprecedented levels of instability. Global maritime transport chains continue to face environmental challenges, while ever-increasing transport volumes are forcing ship

operators to place the highest demands on vessel speed, capacity and availability. Safety is key, and human lives are easily put at risk when OT technology fails. The physical world of OT is anything but virtual. It draws on centuries of experience in dealing with physical risk and safety.

Though being modified and discussed[2][3] for the last 30 years the PERA model's layer structure stays valid mainly as its layers represent the value chain of control system integrators. Depending on the depth of added value sub-systems cover PERA levels from 0 to 2 or just 0 to 1. Although not the original intent of the PERA masterminds[4], this functional and ownership segmentation is one of the main reasons why these layers often designate the vertical boundaries for security zones in ICS today. It does not come by surprise that IEC 62443, the prevailing, international standard for cyber security in Industrial Automation and Control Systems (IACS), defines security zones amongst others as "…grouping of logical or physical assets based upon…responsible organization."[5].

Along with zoning comes the concept of conduits, which are defined as "logical grouping of communication channels that share common security requirements connecting two or more zones"[6]. Here it should be emphasized that conduits define not so much the way zones are connected as the fact that there must be no other connection between zones. Conduits are the conceptional prerequisite for controlled communication between zones, but do not specify how communication is controlled.

In addition to the segmentation of the value chain and security zones, the layer correlates with the type of industrial communication protocols typically used at each layer. While connectionless, low-latency, real-time fieldbus protocols e.g. GOOSE, TSN, Profinet RT dominate level 0 to 2, connection-oriented protocols based on TCP/IP are widely used on level 2 to 3. In short, the lower the PERA level, the less OT system designers value the insertion of additional control or monitoring interfaces.

Successful IIoT implementations consider the segmented nature of ICS considering functional, value chain and specifically security constraints. IIoT Gateways and Edge Computing Devices must be added via appropriate conduits within the cybersecurity context of IEC 62443-3-2. Therefore, IT companies should design Edge Computing Frameworks, where Edge applications can exchange data with each other across networks segments, but can only communicate with industrial assets within the zone and conduit concept of IEC 62443-3-2. Those can be a great tool to add IIoT applications into IEC-62443-compliant networks with very limited push-back from the automation department.

3. Edge to Cloud Connectivity within existing Industrial Control systems

To establish connectivity between industrial control systems and the cloud we can make use of recent developments in the IT space for managing hybrid infrastructure. These developments make it easier, cheaper and less error-prone to centrally operate both infrastructure and work-loads that are widely distributed among on-premise and cloud locations. They even provide mechanisms to centrally support offline scenarios where devices, such as sensors, gateways or servers, are only intermittently connected to the internet or don't have any direct uplink at all. These technologies lend themselves nicely to industrial connectivity applications where we often encounter the triple constraint of

- Having to run application workloads physically close to the equipment that we are monitoring or controlling,
- Working in an environment with heavy restrictions on which components or network segments are allowed to communicate with the cloud and when and
- Having to keep a growing number of physically dispersed IT devices and workloads manageable and secure from a central (i.e. cloud) location.

Let's look at these constraints in turn and how the services and technologies provided by the Microsoft Azure portfolio propose to address them.

What is Edge Computing?

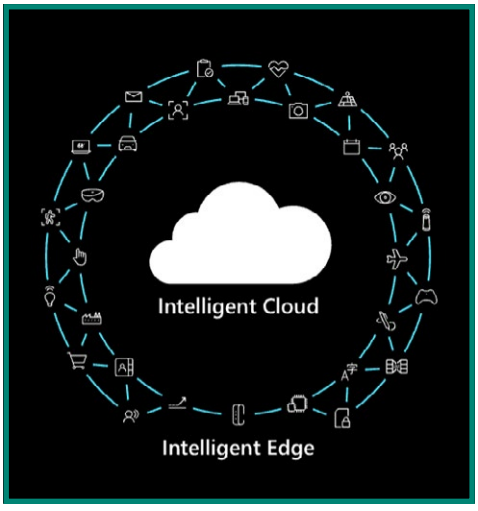

Figure 3: The intelligent Edge and Intelligent Cloud

When we talk about the **Intelligent Edge** we are talking about components and infrastructure that are located outside of centralized data centers. For the context of this paper, this includes any endpoints that directly interface with either end users or industrial equipment, but also the local infrastructure that is required to preprocess data and link it up to the **intelligent Cloud**.

Here, we use the term **device** as a catch-all term for any Edge Hardware Entity that we use in our local infrastructure. This ranges from individual sensors over network components

like switches, gateways and routers up until powerful servers in local data centers. It's important to note that a device may also simply be a logical entity used to describe a collection of physical sensors or even something more abstract like OPC-UA tags. This can be useful to map them to industry-specific Data Model concepts, such as machines or cells.

Gateways are Edge Devices that perform essential network functions like enabling wireless connectivity, providing firewall protection, and processing and transmitting the data of other Edge Devices.

A **workload** is any piece of software functionality that needs to be maintained and managed as a unit. Examples include algorithms that pre-process or transform data before it is sent on to higher business functions, but also interactive programs that may provide graphical user interfaces.

See What Is Edge Computing and Why Is It Important? | Microsoft Azure for details.

Running physically distributed workloads

Several factors specific to industrial scenarios require us to place workloads outside of Cloud Data Centers and on the Network Edge, thus physically distributing them.

For one, we need to deal with large amounts of high-frequency raw data. In most cases it would not be cost efficient to transfer all of it to the Cloud for further processing. In addition to that, we often face bandwidth restrictions that would prevent us from doing so.

Another reason for keeping workloads physically close to the equipment that they are influencing is that any roundtrip to the cloud introduces latency that can't be neglected when dealing with a closed control loop for high-speed industrial processes. It would also not be prudent to rely on an all-time stable Cloud connection for such control loops and the risk to production uptime it would introduce.

Lastly, we often see governance policies in place that disallow the transfer of critical data to the Cloud. This includes data considered critical intellectual property, such as specific machine configurations that optimize performance or historical data that gives insights into proprietary processes.

With IoT Edge, Azure Kubernetes Service and Azure Arc, the Microsoft Azure Cloud provides services that specifically cater to those requirements:

Azure IoT Edge allows to run and centrally manage containerized workloads on so called "light edge" devices. Such devices would include smaller, Raspberry-Pi-class gateways, but also moderately powerful industrial PCs, potentially enhanced by a GPU for AI applications. An IoT Edge device can run custom workloads for preprocessing or aggregating raw data, local instances of certain Azure services but even advanced artificial intelligence and machine learning algorithms. The latter is particularly useful for real-time visual recognition scenarios, e.g. in quality assurance or safety inspection applications.

The data aggregated on the IoT Edge device (Telemetry data) can be sent to the Cloud, in particular to Azure IoT Hub, a scalable device management service for bidirectional communication with IoT Devices. A continuous connection to the Cloud is therefore not required and voluntary.

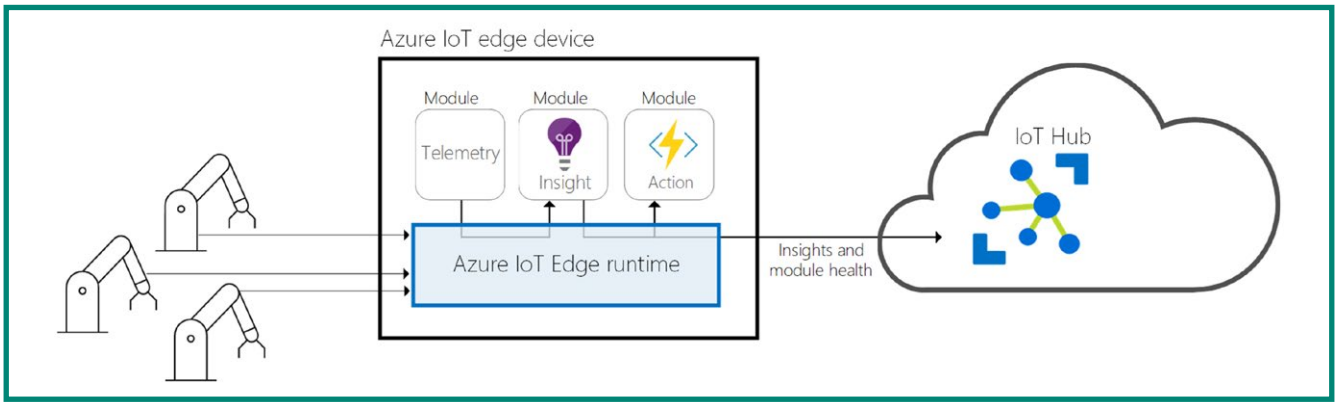

Figure 4: Azure IoT Edge simplified architecture

Like IoT Edge, Azure Arc allows to run and centrally manage workloads on Edge devices. However, it goes beyond the scope of IoT Edge by doing this not only for containerized workloads but also for entire virtual machines and other selected Azure services. As such, it targets more "heavy edge" devices. Those would be powerful, potentially multi-node server racks located on the shopfloor or a directly adjacent Data Center. This makes managing on-prem VMs, databases and entire Kubernetes clusters as seamless as running them in the cloud.

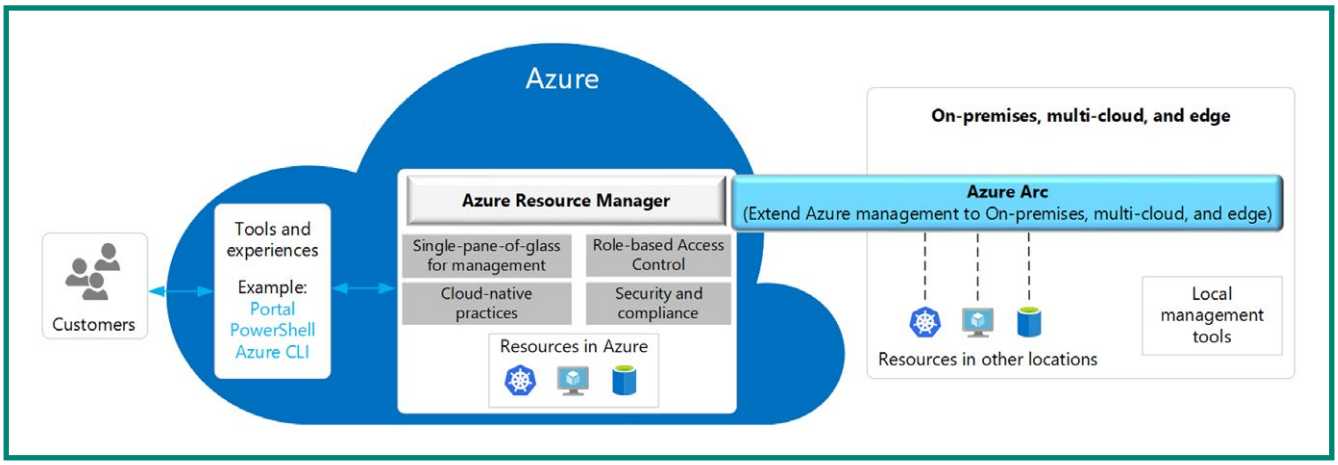

Figure 5: Azure Arc

In combination, these technologies allow us to implement a common Data Model that is working as a bridge between low-level and proprietary industrial protocols and the higher-level data abstractions required by cloud applications. A typical example would be to use machine-specific Edge Modules to collect data from an OPC UA server, harmonize it with the data from other sources in a local **data hub**. Edge workloads or „apps" can use this harmonized data to do the local analytics and processing or send it to the cloud.

Working in environments with restrictions on connectivity

Typical IIoT scenarios, connectivity is often very limited due to very restrictive networking setups to protect sensible data and processes.

In Chapter 2, we described the PERA Model which effectively provides a segmentation into different layers of the network infrastructure. Each layer can only communicate to the next immediate upper layer and data should be aggregated on the different levels. Conduits ensure secure communication between layers and Security zones whereas in every layer a gateway takes care of the respective data aggregation.

In terms of cloud connectivity, only the top a gateway would be able to connect to the cloud using a secure connection and certificates (TLS 1.2+). The lower layers, (e.g. layer 1-3) would also communicate via certificates using a CA certificate on every parent gateway and a child certificate on every immediate lower layer gateway or leaf device.

In addition to high requirements towards security connectivity and communication between layers, we also find other network related constraints such as limited bandwidth, high latency and spotty connectivity in these scenarios. This makes it even more important that data upload can be stored for a longer period of time and forwarded whenever a connection to the cloud or an upper layer is reestablished.

Environments where a nested setup or a highly constraint network setup can be found include vessels, oil rigs and other offshore installations, underground mining and construction sites, temporary or mobile facilities (e.g. in case of humanitarian crises).

Adhering to the PERA Model using a Nested IoT Edge setup:

IoT Edge can be configured to run in restricted environments to integrate with the PERA Model. As per our example, we would add a "light edge" device, such as a Gateway to every layer of the pyramid and configure it to only talk to their direct parent-level devices, not the Cloud.

Parent devices then act as proxies, caches and container registries, providing the same APIs and services to lower layer devices as would be provided by the respective cloud services (at least a subset of it). All of this functionality is provided out-of-the-box without having to manually implement mechanisms to route leaf device requests through multiple layers of gateways and do the same for the response. Simplified, a nested Edge Setup would look like this:

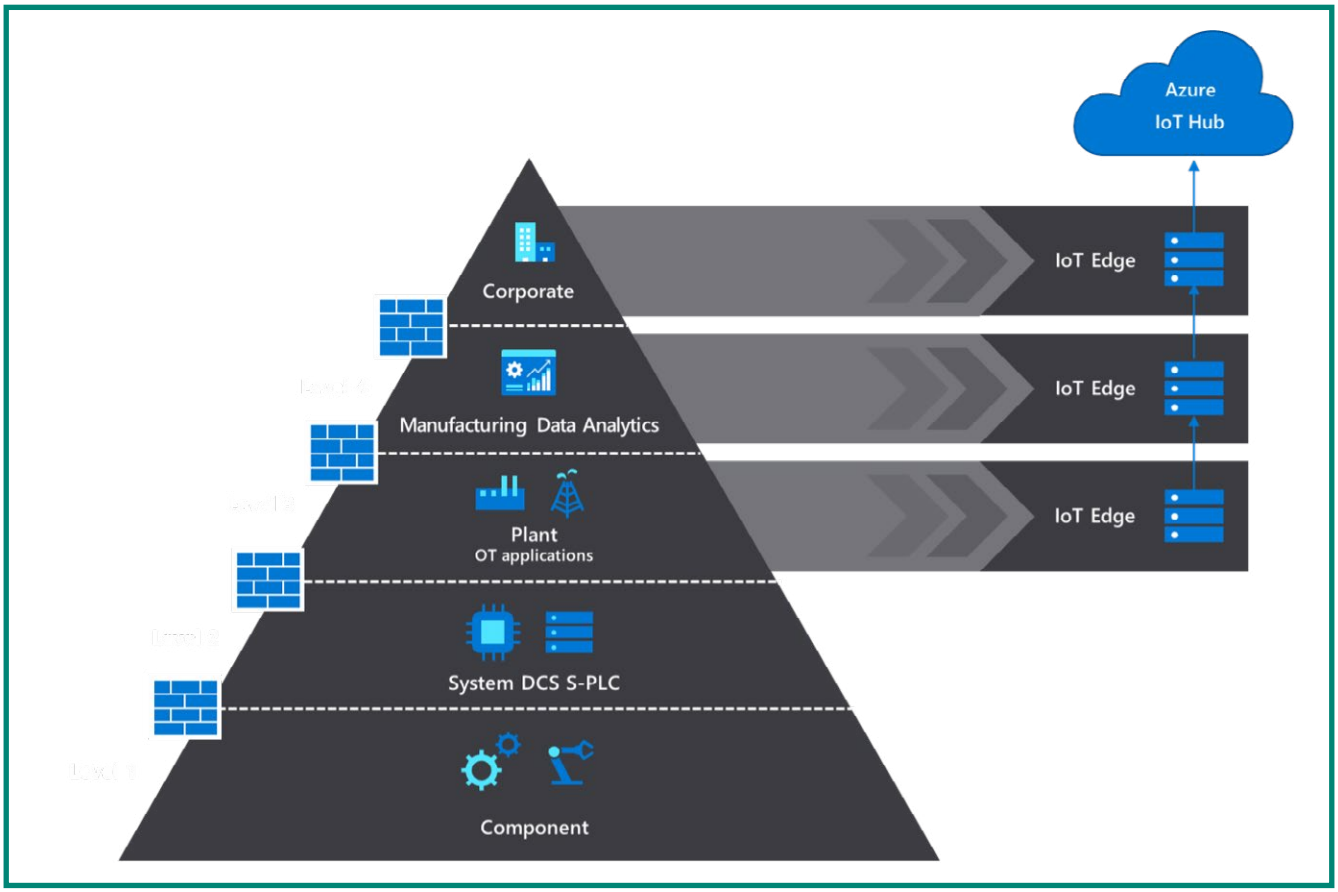

Figure 6: Azure IoT Edge in a nested Edge Setup

By using an automatically synced local container registry, this even applies to software updates that need to be applied to lower layer devices.

A benefit of this setup is that no inbound connections to lower layered devices need to be opened and the same already existing firewalls can be used. Outbound ports need to be opened in order to send data via MQTTS to the next layer and to pull images from the local container registries via HTTPS, but apart from that, the hierarchical Edge configuration allows for minimal changes and intrusion into the existing network setup.

See Hierarchical Deployment of IoT Edge with ACR Connected Registry - Microsoft Tech Community for details.

Manage large numbers of physically dispersed devices

Introducing new devices, such as Edge Gateways, into an existing network can prove challenging especially from a provisioning and management perspective. With every new gateway introduced, customers are faced with an additional layer of complexity. New devices, such as Edge Devices, typically run under a Linux or Windows operating system. These Operating Systems need to be patched, updated and overall, maintained such as every other MPU class device in the OT environment. On top of that, applications on the device need to be managed as well, which has been explained earlier.

In an OT scenario, we normally see three types of behaviors how customers handle Device Lifecycle management:

- Customers maintain a limited number of Operating systems as base images by themselves, normally managed by the IT department. This approach ensures homogeneity among all of the devices used but at the same time, devices optimized for a specific Use Case might not be allowed to be used.
- Customers let OT and IT manage devices on their own and specific needs and allow new Operating systems to be introduced into the OT environment. This approach bears the risk of a large complexity due to an inhomogeneous landscape. At the same time, different systems can be introduced allowing specific devices. Normally we see this behavior with smaller customers, but not exclusively so.
- Customers do not maintain device lifecycle management (DLM) for all devices on their own but use device lifecycle and update management from a Gateway manufacturer such as Moxa or a Cloud Provider such as Microsoft Azure (or both) to ensure optimal patch – and security updates of these devices. This approach is easier for customers in general but now always well accepted due to concerns of customers losing control of their own devices. It is important that device manufacturers and cloud providers provide flexibility when it comes to the control of the device fleet of customers and therefore allowing customers to choose the level of control of their devices.

On the other hand, introducing a large number of Edge devices at the same time can prove challenging as well because the new devices need to be onboarded onto the network, but also registered, be it in a Cloud or local inventory so that they can be tracked and interacted with.

If done manually, without an established workflow, onboarding a device in an existing environment introduces delays and potential for human error that become untenable when regarded at scale. In the worst possible case, a device might even need to be reset to factory settings, requiring again a manual interaction of the OT personnel.

Therefore, we would strongly recommend to customers to establish a provisioning workflow early on. This workflow would need to include the possibility to onboard the device into the network by an Automated Workflow and register the device automatically in a device registry with at least manual interaction as possible.

Device Provisioning Service and IoT Hub for Zero-Touch Provisioning

With Azure IoT Hub and the Device Provisioning Service, MPU and MCU class devices can be onboarded with zero interaction by providing the device with the necessary information to automatically execute an Automated Onboarding Workflow when local or Cloud data is being established.

Once the device is configured and initiated, it will reach out to the Device Provisioning Service and proof its integrity by either showing a symmetric key, a certificate chain or a TPM Endorsement Key. Afterwards, the Device is onboarded onto IoT Hub and configured with the necessary software components.

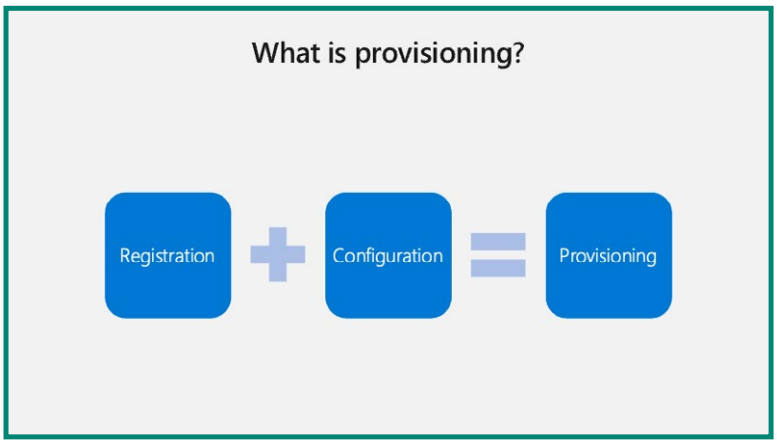

Figure 7: Provisioning definition

It is important to understand that while Microsoft Azure is offering these tools to integrate with companies' onboarding workflows, customers still need to build a fully functional provisioning workflow that works for them out of the box. The necessary SDKs for zero-touch provisioning need to be installed and present on the devices and backed into a device onboarding workflow. In some cases, device Manufactures such as Moxa are offering pre-configured devices for easy provisioning and registering onto IoT Hub that can easily integrate with customer's requirements.

After registering the Device onto Azure IoT Hub, customers benefit from a bidirectional communication (northbound and southbound). For example, devices can be remotely configured to restart, change parameters in one of their software applications and also, run firmware updates on the device. Using the Nested Edge configuration as prescribe earlier allows these functionalities as well. The top layer device would route the new configuration instructions southbound, towards the lower layers.

4. A practical approach to IIoT in existing ICS systems

A highly secure, but still practical approach to adding Industrial IoT applications to an existing ICS system environment relies on the concept of security zones derived from IEC-62443. As described in chapter 2, areas with exposure to similar security threats are kept within a network segment or zone that can only be accessed via defined connections between zones, called conduits. Those conduits only allow network traffic or data exchange between zones necessary to keep the system running (least privilege principle).

When adding an IIoT gateway to an existing network infrastructure based on security zones and conduits, like in chapter 2, a new "IIoT zone" is added to the network and conduits to those zones and devices are created that are necessary to make the IIoT application run.

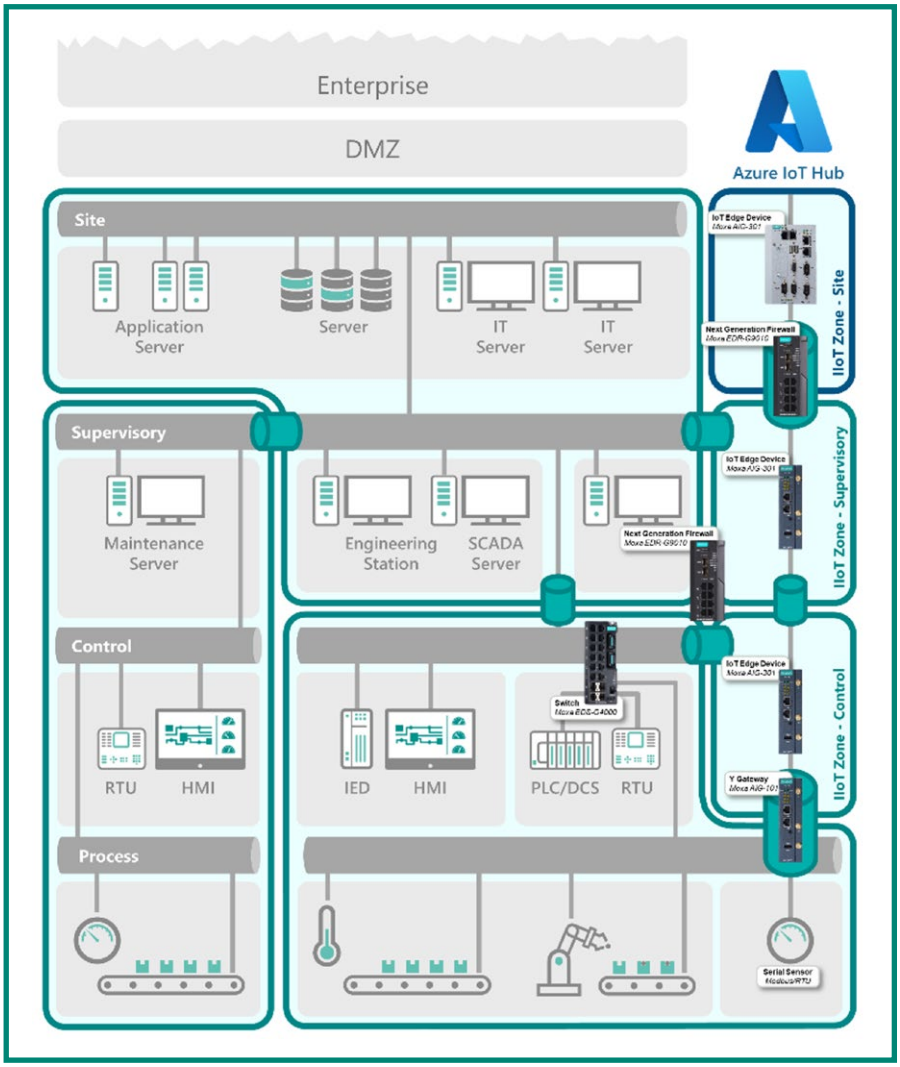

Figure 8: PERA Model with dedicated IoT Zones

As IIoT applications on the one hand and existing networks on the other hand are typically managed by different teams, often from different companies, the two teams have no reason to and should not trust each other ("zero trust" principle).

The new IIoT zone and its conduits are the automation team's way to make sure the IIoT device cannot do any harm to their control system. On the other hand, device hardening and a secure connection to the Cloud are the IIoT team's way to protect the IIoT device from its network environment. The latter task is relatively easy to achieve when using the latest hardware from Moxa like AIG-300/500 and the latest Azure IoT service from Microsoft.

So for the remainder of this chapter we'll focus on how the automation team can add an IIoT zone and set up appropriate conduits in three steps:

- Analyzing existing networks & systems: Identify Data Sources & Sinks in the respective sub-systems & Network Zones
- Building the IIoT zone: setting up the IIoT Edge Gateway(s) in the IIoT zone(s)
- Building the conduits: connecting the IIoT zone to required Data Sources & Sinks

Analyzing Networks & Systems

A modern industrial network is built as a collection of network segments or zones, connected to each other via conduits. Often those conduits are firewalls with firewall rules guarding the type of traffic that is allowed and prohibited between two zones.

The key questions are:

- Where are the data sources and sinks required for the IIoT application and what interfaces are available?
- How can the existing network infrastructure be used to connect data sources and sinks?
- Where can the IIoT Gateway(s) or Edge Computer(s) be placed?

Diagram 1 shows an existing network diagram with the assets marked that are relevant as data sources for our IIoT application. Zones and conduits affected are identified and a potential location for the IIoT gateway is identified.

The IIoT gateway should be placed, such that it can easily be connected to the data sources and at the some time to the Internet. If the gateway has a direct connection to the Internet, e.g. via a 4G or 5G mobile network, it is easy. If the local Internet access needs to be used, the automation team has to provide "a port on a local switch" with Internet access.

In security terms this port on a local switch is the conduit from the IIoT gateway to the Internet zone. In practical terms, the local switch should establish a VLAN to the DMZ and from there to the Internet. It's easy to set up with managed switches (using port-based VLAN) and it gives the automation team the control they need to carefully channel the IIoT gateway's traffic to and from the Cloud without any possibility from the IIoT gateway to the access and interfere with local assets. Often a dedicated physical port on the IIoT gateway is reserved for this so-called North-bound traffic, while a second physical port on the gateway is used for the so-called South-bound traffic to exchange data with local assets. In reality a single physical port is sufficient, when routing is set up appropriately, but most implementations follow the two-port logic. Maybe this is done, so the connection to the Internet can easily be cut by physically removing the corresponding LAN cable.

Building the IIoT zone

Once the relevant assets and potential network connections between them are identified, the so-called IIoT zone has to be established. In its simplest form the only asset inside the IIoT zone is the IIoT gateway. The conduit to the Internet has already been discussed in the previous section and it is a vital building block during the deployment of the IIoT infrastructure.

As described in the previous chapter, IIoT gateways can be deployed using the so-called zero touch provisioning concept. The gateways come pre-configured and enrolled into the customer IoT Hub's list of allowed IoT devices. When the devices arrive physically at the site, the local technician only needs to install them mechanically, connect power and LAN cable(s) and the devices connect to the Cloud automatically and receive their complete configuration details. Even software modules running within IoT Edge can be deployed and changed after the physical installation at site.

In short: everything can be managed from the Cloud, once the IIoT gateways are connected to the Internet. This is called Device Lifecycle Management (DLM) and represents an important pillar of modern IIoT infrastructure. That's why ideally that part of the IIoT zone should be established before the gateway is installed. In practice this boils down to a LAN port on a local switch with a VLAN routed to the Internet.

If multiple IIoT gateways have to be placed in different IIoT network zones, e.g. in a nested Edge Architecture, the conduits between those gateways/IIoT zones have to be established in a similar fashion as described in the next section: only MQTTs traffic between those two devices are allowed.

The next step is to establish the South-bound connectivity via the appropriate conduits. This often means setting up appropriate firewall rules in case of Ethernet-based data sources, but there are other options.

Building the data paths: connecting IIoT infrastructure to OT data sources

For the automation network management team there is one relatively simple way to connect two assets in two different zones, if the conduit between those zones is a firewall and the two devices can exchange data via TCP/IP, like when using Modbus/TCP, OPC UA, REST API, MQTT and many other commonly used standard and proprietary protocols today.

Conduit type 1: Industrial Firewall
A firewall implements policies via rules that allow or prohibit certain types of packets between devices on different physical ports of the firewall. So for instance, when both teams agree that the IIoT gateway should be able to read data from a Modbus/TCP power meter in a different network segment, the firewall can be set up to allow packets between those two devices. The technical details behind such a firewall policy can be very granular, in order to make sure that the IIoT gateway really can only read data from this device and only access it via Modbus. Practically this can be achieved by specifying the IP and/or MAC addresses and ports of the devices involved and in the case of the Moxa EDR-G9010 by applying deep packet inspection to check the protocol and within it the type of command, e.g. reading or writing data via Modbus and only in one direction.

Conduit type 2: VLAN to the DMZ / Internet
An additional network feature can be combined with the firewall: the VLAN. A virtual LAN (VLAN) is an overlay network over an existing network that isolates traffic between a subset of devices from the rest of the network. Most if not all managed switches support VLANs and in combination with layer 3 routers and/or firewalls the data exchange between the IIoT gateway and the Cloud can be securely channeled through a VLAN to the Internet.

Also connections to industrial assets that are not in the same or a Neighboring Network Segment can be established by setting up VLANs from the IIoT gateway to the other device, potentially crossing a number of firewalls or routers. Still the VLAN makes sure that traffic between those two devices is kept separate from all other traffic.

Conduit type 3: IDS/IPS systems
In the context of network zones and conduits, intrusion detection / prevention systems (IDS/IPS) are an advanced version of the firewall of conduit type 1. In addition to the firewall functionality, IDS/IPS systems can monitor the traffic and detect a variety of intrusion events. An example of how the Moxa IDS/IPS system on Moxa's EDR-9010 can securely allow access to legacy industrial assets is the virtual patch feature: If for instance an HMI or IPC is still running an old "unpatched" version of Windows and can not easily be upgraded quickly, the virtual patch feature of the EDR-9010 can help. The vulnerabilities of old Windows systems are typically known and so are the traffic patterns of the respective exploits. If the EDR-9010 detects such a pattern it can either warn the operations team (intrusion detection) or directly discard such packets (intrusion pre-

vention). This way devices and systems that the automation team would normally keep separate (air-gapped) can also be connected without unreasonable risk.

Conduit type 4: "Y Gateways"

Conduits are simple to implement in greenfield or other relatively new environments where all devices have Ethernet port and can communicate via TCP/IP (conduit type 1-3). Setting up data connectivity to older or very cost effective assets, with connectivity options like proprietary serial, Modbus/RTU and fieldbus protocols like Profibus can appear difficult at first glance. Moreover these assets are often already connected to local control system and do not necessarily support multiple hosts, i.e. providing data to multiple systems or devices. Specifically those devices often cannot be connected simultaneously to the local SCADA and the IIoT gateway.

In many cases special types of gateways can solve that problem elegantly: serial-to-Ethernet converters like the Moxa NPort portfolio, or fieldbus protocol converters like the Moxa MGate portfolio serve the purpose of connecting legacy devices to the network and do support multiple hosts. SCADA engineers have rich experience in connecting legacy assets to the network and to SCADA systems, and Industrial IoT project teams should not hesitate to leverage their expertise.

A "Y gateway" is not a technical term. It illustrates the fact that in many IIoT scenarios data from lower levels of the Purdue Model (the bottom leg of the letter Y) has to be shared both locally for traditional control systems and with Cloud-based IIoT systems (the two upper lines of the Y).

The Moxa AIG product line of IIoT Gateways is built for that purpose. A common scenario is the connection of many Modbus/RTU power meters to both the local energy management system (EMS) and the Cloud. The AIG-101 gateway is built with that purpose in mind: For instance, it can collect data as a Modbus/RTU master on the South-bound interface and makes it available on the North-bound interfaces as a Modbus/TCP server to the local EMS. In addition it publishes data to Azure Cloud. As polling/publishing frequencies and local processing needs differ between the local EMS or SCADA system and the Cloud, AIG gateways offers simple ways to set up those data paths independently.

Please note that in this case the Y gateway is the responsibility of the automation team, not the IIoT team. It serves as a conduit from the IIoT gateway in the IIoT zone to the power meters in the "power meter zone" and it protects the power meter zone from the IIoT gateway. The IIoT gateway cannot directly connect to the power meters, because it only gets the data from the Y gateway. And if there are any doubts about the integrity of the IIoT gateway, e.g. when it starts polling the power meters at very high frequency, the automation team can simply change the configuration on the AIG-101 and disable the connection to the IIoT gateway completely.

Nested Edge, Common Data Model, Apps

Now the IIoT gateways have been integrated into the existing network infrastructure by setting up new IIoT zone(s) and respective conduits and the data path is open, but not yet implemented or used by applications.

This is when the power of IoT Edge and the ability to share data and device management functionality across the entire network of IoT Edge devices can be leveraged.

A hierarchy of IoT Edge enabled IIoT gateways forms the basis for a network of Cloud-managed software container environments across the levels and network segments of an existing Industrial Control System.

Applications developed for these IoT Edge environments could now use the defined data paths directly. However, introducing a common Data Model can decouple applications from low-level data acquisition functionality, so app developers can build their code without worrying about whether the data has been collected via a specific Modbus register or an OPC UA tag. Defining a common Data Model for the IoT Edge / Nested Edge environment can be as simple as defining a JSON schema with an appropriate level of semantic description. Small software modules on IoT Edge can serve as connectors between the low-level physical interface and the common Data Model.

App developers can now focus on building their core logic without dependencies on the lower level infrastructure. Using Azure IoT infrastructure, those apps can be deployed from the Cloud as IoT Edge modules across the complete hierarchy of IoT Edge enabled IIoT gateways. That way the IIoT application layer on IoT Edge is decoupled from the IIoT infrastructure layer.

In order to maintain the highest level of flexibility as IIoT applications are rolled out more and more, there are two main design principles that companies should enforce with their suppliers and internal teams:

1. **IoT Edge**: All edge application software should run as an IoT Edge module
2. **IoT Hub**: Edge device management should leverage Azure IoT Hub services

As a result companies can deploy any IoT Edge-enabled hardware, even from different vendors, to meet connectivity and performance needs. And once installed edge applications can be added or removed, when the need arises. This is true both for internally developed software as well as edge applications from third parties.

Even when industrial assets are upgraded, changed, or replaced, the impact on the app layer remains limited, when the IIoT infrastructure, including the respective connectors for the common Data Model are adapted accordingly.

So how to get such IIoT infrastructure components operational in the field as easy and fast as possible? As described in the previous chapter, Microsoft provides an efficient framework to provision IoT Edge gateways, i.e., registering with IoT Hub via DPS (Device Provisioning Service) and configuring each gateway appropriately. Moxa IoT Edge gateways come with TPM (Trusted Platform Module) as the hardware root of trust that DPS uses to securely provision the gateways. Making this process even easier and faster, Moxa has implemented the tools and services to handle this provisioning process for their IoT Edge capable gateways at scale. In addition, IIoT Service Partners stand ready to offer zero touch provisioning service to customers. Then local customer staff primarily needs to plug in power and a LAN cable – or use LTE – and the IIoT gateways connect to the Cloud by themselves and from then on, "everything" is managed from the Azure Cloud. IIoT projects could not be much easier, faster, and scalable.

5. Conclusion and recommendations

A journey of a thousand miles begins with a single step, as they say. This is particularly true for the digital transformation journey. Step one is to add a first IIoT gateway to the ICS network, using the mechanisms described in this paper:

- connect to a firewall with one VLAN routed to the Internet (the North-bound conduit) and one policy that allows connections to an industrial data source (the first South-bound conduit)
- define a common Data Model / JSON schema and build a simple app that collects OPC UA or Modbus data and sends it to a simple Cloud Dashboard
- enjoy the fact that both Gateway and Edge applications can be managed entirely from the Cloud and tweak the dashboard until it brings some first value to the ICS team

After the first step, more steps will follow. The Azure IoT framework with its Nested Edge architectures allows the hierarchy of gateways to grow without the need to connect all gateways individually to the Cloud. Zero-touch Provisioning can make this process particularly easy for large scale roll outs.

The fact that software modules across the hierarchy can be added, removed and managed from IoT Hub allows companies to separate their Edge application strategy from the Edge infrastructure strategy. They can choose to build their first applications themselves and move to more sophisticated third-party applications later. Or they do the reverse and start with commercial off-the-shelf software first and replace with more specific and cost-effective self-developed solutions later, after learning what works and what doesn't work both technically and commercially.

Either way, the first step is made, and the digital transformation journey has started.

References

[1] Gartner Information Technology Glossary
[2] The Purdue Reference Model outdated or up-to-date?, blog, 2020-06-18, Sinclair Koelemij
[3] Is the PERA model still pertinent?, web, 2022-09-02, Ian Verhappen,
[4] Stephen Mathezer, "Introduction to ICS Security Part 2", 2021-07-16, SANS Blog
[5] IEC 62443-3-2:2020, 3.1.25
[6] IEC 62443-3-2:2020, 3.1.3