IoT -
The barriers to scaling

Contents

A complex world of interconnected devices is creating new opportunities for many businesses, but by far the most exciting possibilities can be found in the manufacturing sector.

As it stands at the moment, 90 percent of manufacturing machinery and devices are unconnected, so the expectation is that IoT connectivity will deliver the next step change in productivity if implemented correctly.

"The internet has transformed the way we work, driving new areas of productivity. Those productivity increases are now going to be seen in industry, through IoT, but only if delivered in the right way"

Sir Tim Berners-Lee

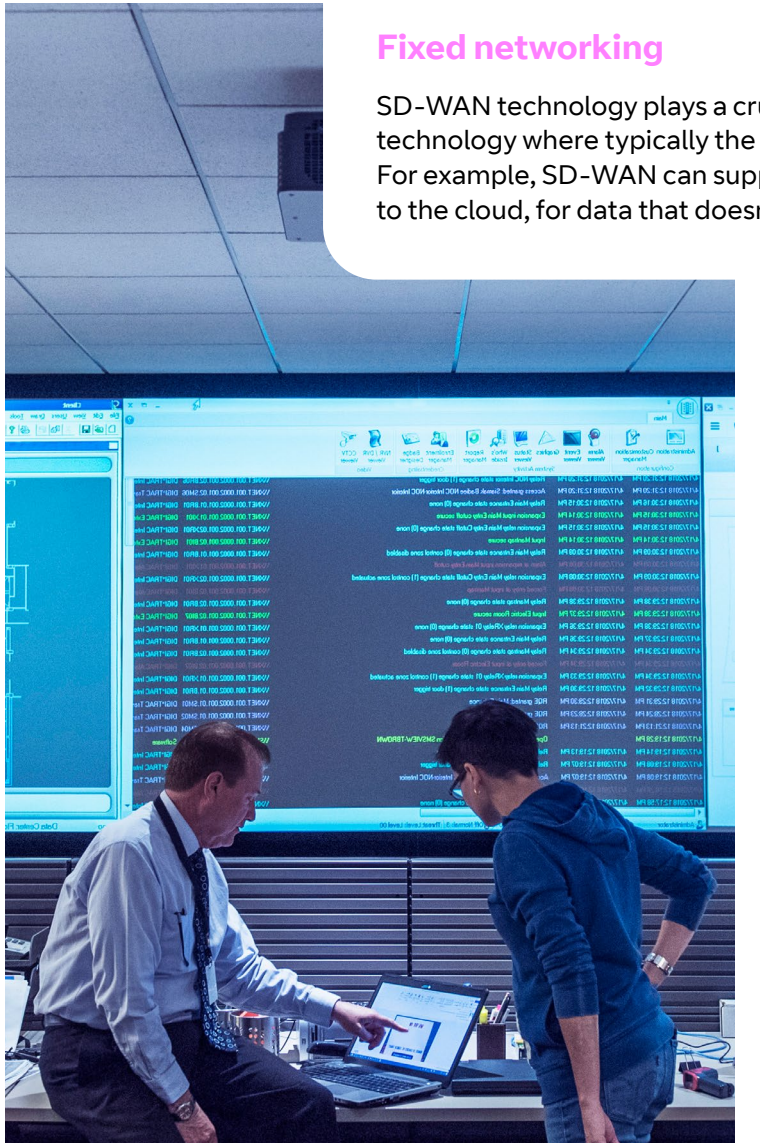
In this paper, we'll explore some of the challenges you may face when connecting thousands of devices, and provide some recommendations on how and where to process the data, how you can deliver the experience you need, and most importantly, how you make sure your devices and the rest of your organisation remains secure.

Challenge one – Edge infrastructure

Our customers tell us that they are planning to process up to 50 per cent of their data at the Edge[1]. They're driven by a need to reduce their connectivity costs and keep critical business and manufacturing processes running in the event of network issues. So, one of the main challenges when 'moving to the edge' is setting up an infrastructure that supports the implementation of IoT solutions. However, many businesses have legacy networks that can't do this.

[1] https://www.btireland.com/company/communications/bt-ireland-blog/authors/andy-rowland/the-disruptive-nature-of-the-edge

Considerations for a future-proof network infrastructure



Fixed networking

SD-WAN technology plays a crucial role when implementing Edge technology where typically the starting point would've been the LAN. For example, SD-WAN can support internet breakout from the factory to the cloud, for data that doesn't need to be processed at the Edge.

1. Optimising the network

The LAN needs to be able to collect data from multiple sources to support Industry 4.0 applications, such as plant and Enterprise Resource Planning (ERP) systems. So both the industrial and enterprise LANs may need to be refreshed. Direct connectivity between plant systems and the Edge server is possible if the plant has implemented a centralised, real-time Distributed Control System (DCS) or Manufacturing Execution System (MES). If not, then gateways have to be added to the LAN to collect plant data. These gateways will then feed an Edge server. Data may also come in batches (ERP), may be streamed live from a device (plant) or appear in high bursts (e.g. offloading data from a jet engine on landing). Consequently, the LAN and SD-WAN networks need to be highly flexible to enable efficient data movement and processing.

2. Vendor-agnostic

Many businesses use different hyperscalers for different applications, so future networks and Edge solutions need to be vendor-agnostic, providing businesses with flexibility and choice. Network edge infrastructure with modernised LANs, SD-WAN, enhanced security, access to cloud environments and compatibility with applications and technology, such as 4G and 5G can help organisations transform their network, as well as their business.

3. Intent-Based Networking (IBN)

The Edge sits between the LAN and the WAN so IBN is important in providing the best user experience. IBN starts to move away from the complexity of the technology and focuses on the customer and user experience. It verifies network-wide behaviours, predicts the results of changes, tracks compliance with the stated intent and policies, and provides guidance on remediation when there's misalignment.

Many organisations still manage their networks manually which can lead to delays in restoring connectivity after an outage. According to Gartner, a full IBN implementation can reduce network infrastructure delivery times by 50 per cent to 90 per cent, while at the same time, can decrease the number and duration of outages by around 50 per cent[2]. IBN enables digital transformation through automation and by aligning the IT and network services with business goals.

[2] https://www.gartner.com/en/documents/3599617/innovation-insight-intent-based-networking-systems

Mobile Networking

In the future, companies may not only want to rely on fixed networks but increase their flexibility and integrate 5G into their network infrastructure.

A great deal of research is focused on Multi-Access Edge Computing (MECs), where 5G applications can be run at the edge of the Radio Access Network (RAN). Researchers are also looking at how these MECs could be virtualised on an Edge compute box.

However, when implementing 5G, there are some important questions that may need to be addressed:

What are the trade-offs between signal penetration and bandwidth inside buildings?

Generally, with higher frequencies, the bandwidth increases (subject to available spectrum of course), as the penetration of the signal decreases. E.g. 4G at 1800MHz offers 22Mbps, whereas 5G at 24GHz offers theoretically 1.75Gbps. So for in-building systems, Wi-Fi 6 or small 5G cells inside the factory may be the best option. Companies need to weigh up their options and decide on a case-by-case basis which is the best solution for their business needs.

How do we decide between Wi-Fi 6 and 5G?

Tom Curry, principal architect, mobile infrastructure, BT suggests that: "Wi-Fi (incl. Wi-Fi 6) will remain the 'default' choice for most enterprise connectivity, but there will be some scenarios where extra investment in 5G cellular coverage will be justified."

For example, when:

- companies need additional power to ensure coverage, which may be the case in complex and challenging radio frequency environments (e.g. manufacturing) or, when companies are confronted with mixed indoor/outdoor deployments (e.g. ports, airports, logistics hubs), or
- dealing with mission-critical use cases for 5G where deterministic performance, i.e. coverage, speed and latency, is needed and where interference from other networks or users is unacceptable (e.g. use of military drones).

How do we manage quality of service?

Let's say your employees use 5G for business-critical activities, but there's a major sporting event taking place, so the public network is congested. Could a dedicated network slice be made available?

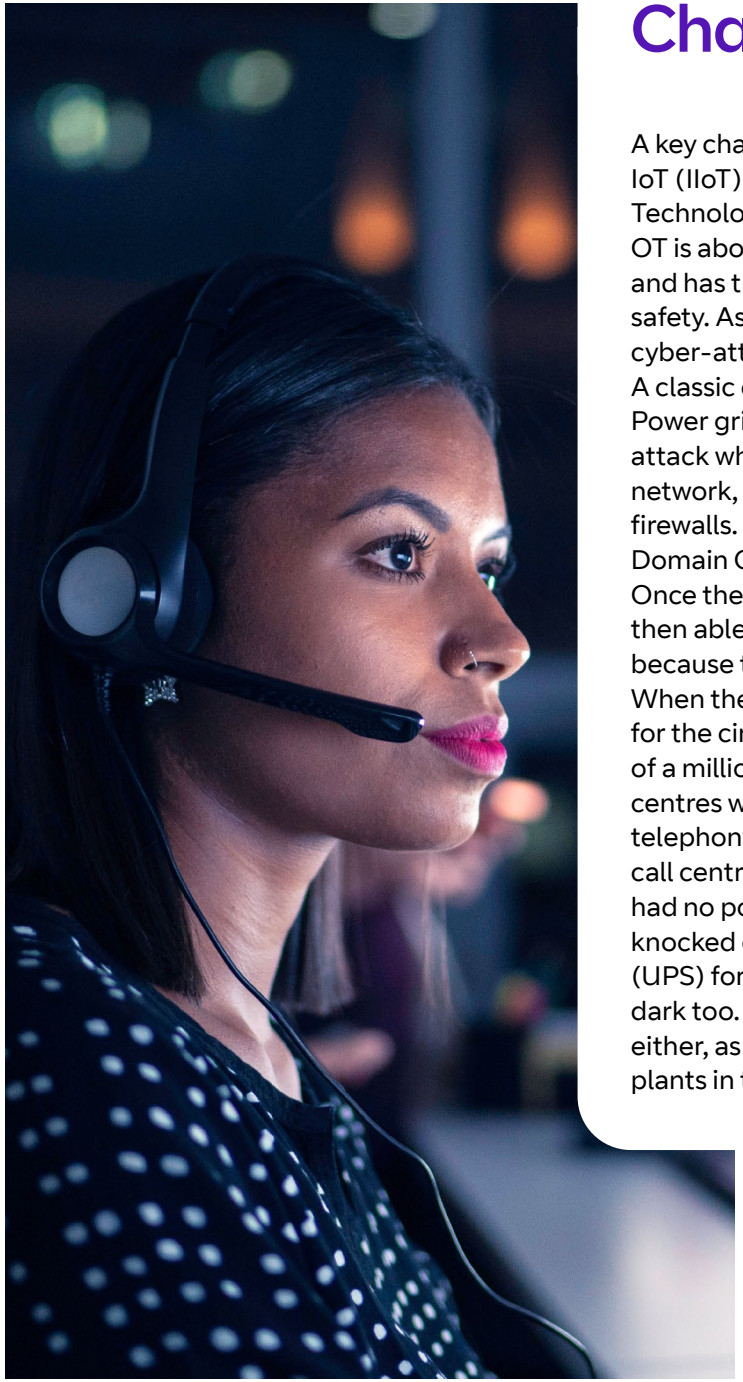
To make this work, there would need to be a high degree of automation in the network and we're not there yet. However, in the meantime, 'partial' slices may exist in a more static manner in 'controlled' environments, such as private or hybrid networks, where certain types of traffic within a customer site may be sent to a different core (e.g. a local onsite core versus a centralised/shared core).

How do we use private 5G spectrum for manufacturing and what are the challenges of managing a RAN as opposed to wi-fi?

There's an increased use of private 4G and 5G networks, co-existing with public 4G/5G to meet specific customer requirements, e.g. dedicated capacity as well as specific performance or data security requirements. Private networks may also be useful in very remote locations (e.g. oil rigs). The increasing availability of locally licensed spectrum in global markets will make it easier to deploy private networks. But private cellular networks are typically much more complex to operate than wi-fi networks[3].

To conclude, our researchers believe that the future network will be neither fixed nor mobile, but rather a smart hybrid core, divided into multiple slices, that can cope with traffic being delivered over any access type depending on the best available access and needs of the customer. This includes the concept of Access Traffic Splitting, Steering and Switching (ATSSS) that can make dynamic decisions about how to deliver traffic to the end-user device. For example, a device connected to wi-fi within an organisation may gradually move away from wi-fi and be simultaneously connected to the 5G network to maintain a high quality of experience — with traffic being delivered over both access networks until the wi-fi access disappears. In this vision, end users needn't worry about which access network they're connected to at any point in time. The device and the network collaborate to decide what is the best combination of access network(s) to use to deliver the service they need.

[3]https://www.computerweekly.com/feature/Private-5G-networks-Are-they-the-right-choice-for-you



Challenge two – Cyber security

A key challenge is cyber security. Industrial IoT (IIoT) connects two worlds: Operational Technology (OT) and Information Technology (IT). OT is about availability first and security second and has traditionally relied on segregation for safety. As these air gaps are removed, cyber-attacks on OT will become more common. A classic example was the attack on the Ukraine Power grid in 2015. This started with a phishing attack which allowed the attackers into the IT network, but the OT network was secured behind firewalls. The next step was to access the Windows Domain Controllers to harvest workers credentials. Once they had user IDs and passwords, they were then able to access the VPNs the grid workers used because there was no two-factor authentication. When they were in, they reconfigured the firmware for the circuit breakers, plunging almost a quarter of a million people into the dark. The control centres were also caught unawares due to a telephony denial of service (TDOS) attack on the call centres, so people couldn't ring in to say they had no power. To add insult to injury, the criminals knocked out the Uninterruptable Power Supply (UPS) for the control centres so they were in the dark too. These type of attacks haven't gone away either, as evidenced by recent events at aluminium plants in the US and Norway[4].

Two recent trends we have seen are hactivist groups going after safety critical systems whilst also playing the long game[5].

Over the course of 2019, Microsoft watched the activity of APT33, an Iranian Advanced Persistent Threat (APT), as they conducted password spraying attacks against thousands of organisations. Initially targeting a wide range of organisations, the group then appeared to narrow their attempts to around 2,000 organisations and increased the number of accounts compromised at each organisation. Of those accounts they tried to exploit, the top 25 were related to organisations' manufacturing or supplying SCADA/ Industrial Control Systems(ICS) equipment. This suggests that the Iranian APT is playing the long game, attempting to compromise the producers of industrial control systems in order to target the organisations actually using them. Taking into account the known activity of Iranian APTs and their proficient use of malware, this presents a very real threat, and one which may not be realised for some time. The effects of using malware on industrial control systems would shift the motivation of the Iranian threat group from destruction of information and operational capability to actual physical destruction.

Ben Goodman from ForgeRock told Information Age that the Edge environment resembles the Wild West, "almost a no man's land"[6], there are lots of players, but little standardisation in place which means that "... to a certain degree, every project is bespoke, so that's very challenging," says Goodman. So, it's important for businesses to start thinking about cyber security when implementing IIoT at the Edge.

[4]https://www.bbc.com/news/technology-47624207
[5]https://www.technologyreview.com/2019/03/05/103328/cybersecurity-critical-infrastructure-triton-malware/
[6]https://www.information-age.com/cyber-security-for-iot-and-edge-computing-123485616/

Considerations for best practice in cyber security

Best practice can generally be divided into two areas: Good housekeeping and advanced threat analytics.

Good housekeeping – locking out those unwanted guests

Imagine you're buying a new house; one of the first things you make sure is that no one can enter without your permission. The same principle applies to IIoT at the Edge.

Then, you need to classify which devices and/or machines are connected to the network. It's important to identify vulnerabilities in case they're exploited by hackers to steal, destroy or modify sensitive data. Companies need to profile their assets and determine their importance to the business as part of a risk-based security strategy.

Micro-segmentation of the network becomes even more crucial in the context of Edge so that the impact of any future attack is limited.

By understanding the profile and purpose of specific assets and networks, companies can implement segmentation based on:

1. isolating specific business processes
2. safety risk
3. age of technology (e.g. old devices that can't be patched)
4. requirements for third-party access and limiting visibility of parties to specific assets.

Advanced threat analytics – the smoke detector

It's crucial to have good visibility of device and network behaviours to detect anything abnormal. One option is the use of visual and graphical analytics tools to detect behavioural outliers.

Visual analytics tools are used to spot data patterns and let users combine their business knowledge with AI to explore very large sets of structured and unstructured data.

A big advantage of visual analytics is that the analysts remain in the problem space rather than having to think about speaking the language of the database.

Graph analytics identify threats to the network. They model relationships that exist or can be derived from the data and let resulting graphs be visually explored. They allow the user to filter and style the data at scale. Also, graph analytics are underpinned by AI-based big data analytics techniques to preserve the most significant data aspects before pushing them to analysts (for example, clustering and outlier detection).

While AI is becoming an essential part of digital business, we need to be aware of the "dark side of AI": Adversarial AI which refers to machine learning techniques that try to mislead models through malicious inputs, so everything appears fine when it's not. Although adversarial AI is still in its early stages, there are a number of examples.

In 2017, researchers at Kyushu University showed that by changing only one pixel it was possible to fool current deep learning algorithms[7]. Researchers have also discovered methods for altering the appearance of a stop sign so that an autonomous vehicle will classify it as a merge or speed limit sign[8].

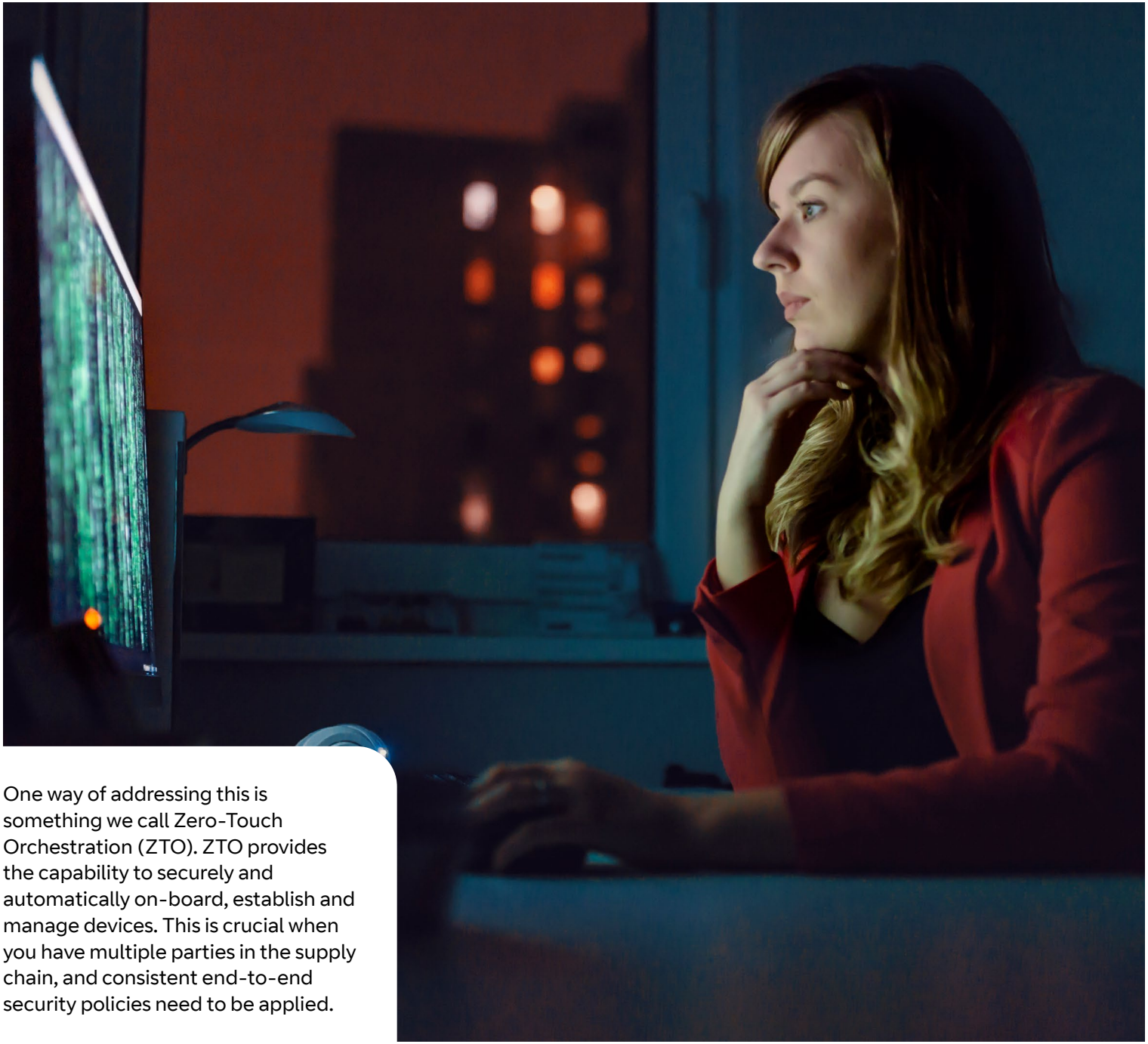
To address adversarial AI, designers need to identify potential vulnerabilities by simulating potential attacks on the system. Countermeasures such as additional features or different learning classifications may be needed. But, because proactive approaches are not necessarily better than reactive ones, companies should always consider investing in both.

[7] https://www.bbc.com/news/technology-41845878
[8] https://towardsdatascience.com/breaking-neural-networks-with-adversarial-attacks-f4290a9a45aa

Challenge three - Orchestration

One of the greatest challenges is how to on-board, scale and secure millions of devices automatically. For manufacturing alone, there are over ten million sites globally[9]. Analysts report that manual provisioning and on-boarding of each device can take more than 25 minutes on average; so for 100 devices per factory, that would be about 240,000 Full-Time Employee (FTE) years of effort, assuming eight working hours a day for 220 days a year! In today's IT industry, this is incredibly costly, impractical and prone to mistakes which could lead to security breaches.

This is on top of the 25 per cent of total project budget often needed for manual maintenance and in-life device management.



One way of addressing this is something we call Zero-Touch Orchestration (ZTO). ZTO provides the capability to securely and automatically on-board, establish and manage devices. This is crucial when you have multiple parties in the supply chain, and consistent end-to-end security policies need to be applied.

[9]https://www.scmo.net/faq/2019/8/9/how-many-factories-is-there-in-the-world

Establishing a Zero-Touch Orchestration approach

The ZTO approach is based on four services:

1. Zero-Touch Attestation (ZTA)

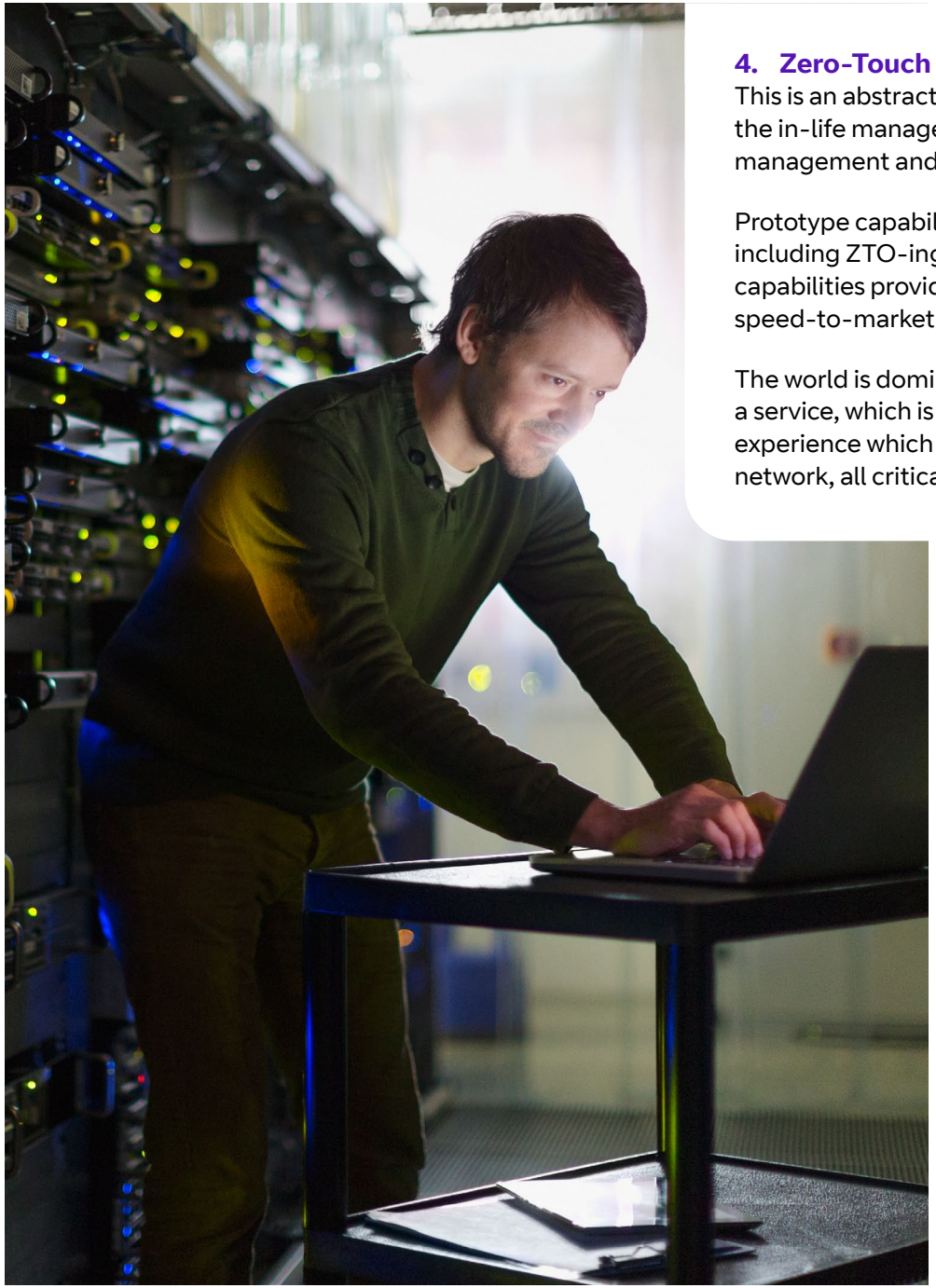
This interacts with third-party solutions to establish bi-directional trust between IoT management servers and remote endpoints with hardware root of trust. These endpoints usually come from a variety of individual vendors. There may well also be a number of different third party on-boarding services which the attestation service will interact with in order to securely on-board as wide a portfolio of devices as possible.

2. Zero-Touch Bootstrapping (ZTB)

This automatically prepares and provisions endpoints for in-life operation. ZTB automatically provisions the cloud management server and also prepares and encapsulates the protocols, firmware, applications and device management agents for each device. These are based on its type, resources, purpose of use and other capabilities like chosen cloud device management platform. Applications, configurations and installation scripts are then sent to the Edge device automatically, all secured bit-by-bit, using the secure channel established by ZTA.

3. Zero-Touch Connection (ZTC)

Maximises the automation experience, in that devices remotely, automatically, and most importantly, securely, connect to the network infrastructure without getting provisioned in the network cloud or on the device.



4. Zero-Touch Device Management (ZTD)

This is an abstraction layer between the various networking protocols and device protocols to simplify the in-life management of the components; including chain of ownership, digital twins, decommission management and hand-over.

Prototype capabilities exist which have been demonstrated on a full end-to-end IoT deployment including ZTO-ing: private networks, customer Edge compute and end Industry 4.0 IoT devices. These capabilities provide low cost, secure and rapid onboarding, late binding capability and increase speed-to-market.

The world is dominated by digital natives. People expect to be able to buy and manage anything as a service, which is vendor- and platform-agnostic from a single shopping cart. People want a digital experience which delivers a repeatable experience to deploy secure IoT devices and intent-based network, all critical for smart manufacturing.

In this digital marketplace, telcos and vertical industry partners can package their own products and services with products and services from other partners and sell and support these new offerings to enterprise customers in a frictionless, self-service digital ecosystem. This approach offers scalability, global repeatability, and monetisation for all parties and, above all, secure trusted endpoints.

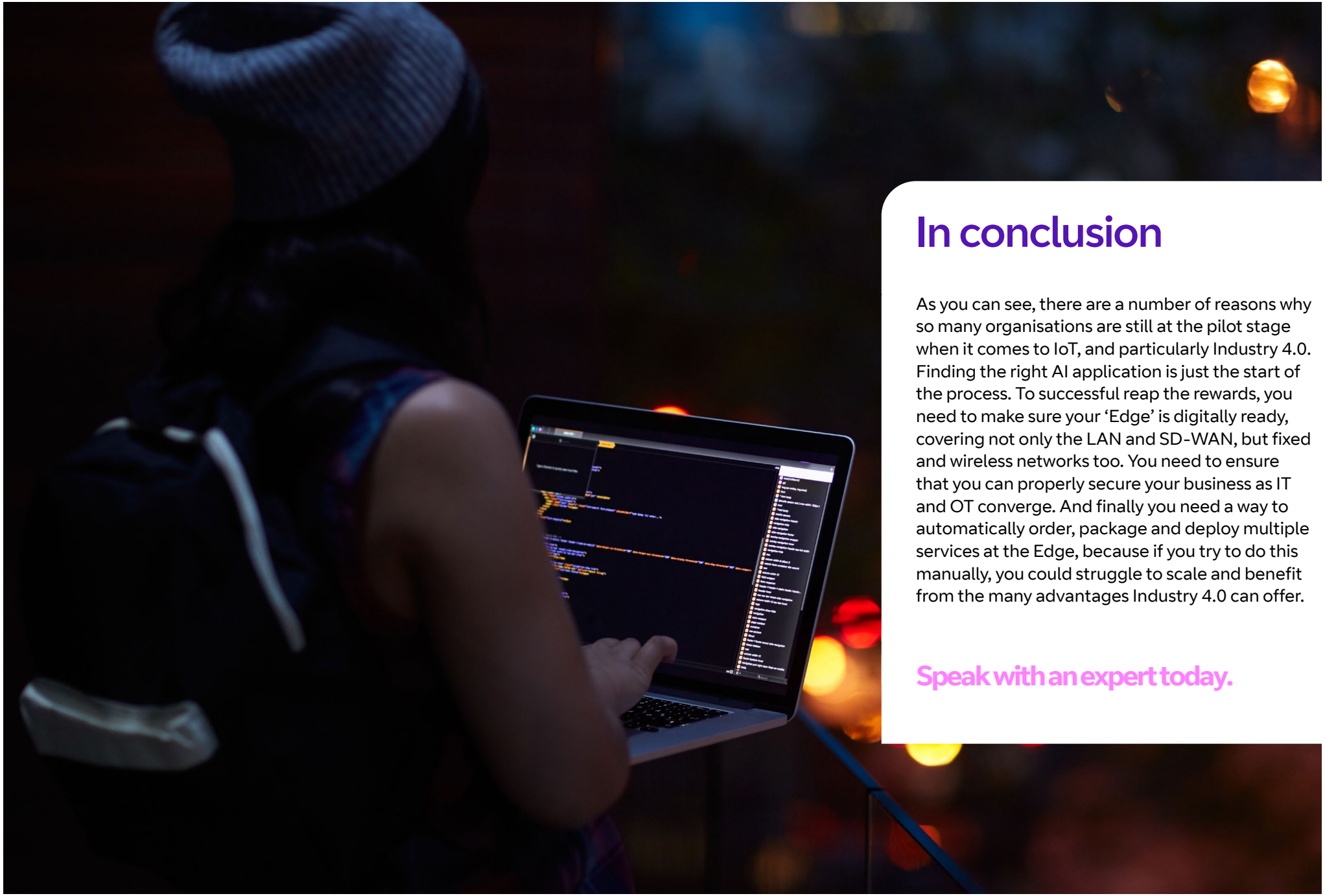
Customers want to achieve this in a single place, automated with late-binding capabilities, with security down to the chip of the devices.

So instead of prescribing pre-built bundles of unified Customer Premises Equipment (uCPE), SD-WAN and virtual firewalls customers can build up their own uCPE architecture and IBN in their shopping basket.

As an example, as a proof of concept, we deployed a Zero-Touch OSS, using Intel® Secure Device On-board[10] as one of the available Zero-Touch attestation choices (and now a FIDO Alliance standard[11]). We also built a frictionless trading Digital Business Platform, and a digital self-service shopping cart to reduce the hugely laborious manual work that would have been needed before to on-board and orchestrate IoT devices and networks.

[10] https://www.intel.com/content/www/us/en/internet-of-things/secure-device-onboard.html

[11] https://fidoalliance.org



In conclusion

As you can see, there are a number of reasons why so many organisations are still at the pilot stage when it comes to IoT, and particularly Industry 4.0. Finding the right AI application is just the start of the process. To successful reap the rewards, you need to make sure your 'Edge' is digitally ready, covering not only the LAN and SD-WAN, but fixed and wireless networks too. You need to ensure that you can properly secure your business as IT and OT converge. And finally you need a way to automatically order, package and deploy multiple services at the Edge, because if you try to do this manually, you could struggle to scale and benefit from the many advantages Industry 4.0 can offer.

Speak with an expert today.

Offices worldwide

June 2020