Industrial applications of Edge Computing

***Ashok Debnath**, Senior Technology Consultant,*
IoT, Data & Services Innovation, FORCE Technology (asde@force.dk)

***Michele Colli**, Industrial IoT Specialist,*
IoT, Data & Services Innovation, FORCE Technology (mic@force.dk)

Table of Contents

Executive summary

Internet use is trending towards bandwidth-intensive content and an increasing number of connected "things". To support the increasing needs that are coming with this trend, mobile telecom and data networks are converging into a computing architecture aiming at lowering data transport time while increasing data availability. International Data Corporation (IDC) predicts that "every connected person in the world will have at least one digital data interaction every 18 seconds — likely from one of the billions of IoT devices, which are expected to generate over 175 ZB of data in 2025 (Statista Research Department, 2021)"[1]. The "Internet of Things" (IoT) and the explosive growth of data generated by IoT devices is creating an unprecedented need for technologies like Edge Computing and AI to realize the value of such data, translating it in useful insights for supporting processes or optimizing systems.

Industry has understood the value of data for improving the efficiency of time-consuming processes as well as for meeting growing expectations for real-time customer engagement and responsiveness. This entails the ability to take advantage of the data generated by the myriad of sensors, embedded computers, industrial controllers and connected devices such as vehicles, wearables, robots and drones.

Given the volume of such data and the rate at which data is generated, the adoption of Edge Computing for processing it is a necessary step for industrial companies. Processing and analyzing data at the edge, close to its point of generation instead of in a centralized location (e.g. the cloud), has the potential to provide better support to those processes that relies on timely insights coming from the data they generate themselves, and that does not have to be aggregated with external data – what cloud platforms are mainly intended for. In addition to that, the robustness of such insights is supported by the offline nature of Edge data processing which, by processing data locally, does not depends on the stability of an internet connection – necessary when data has to be processed in the cloud.

If Cloud Computing is largely adopted especially when the focus is on storage capabilities and on scalability in terms of data processing, Edge Computing is preferred when there is a need for:

- Low latency, due to the presence of applications that require (near) real time data processing and feedback
- Offline capabilities, due to environments characterized by unstable network connections
- Low data transmission cost, due to the transmission of large amounts of – often unused - data
- Narrower bandwidth, due to network limitations
- High security requirements, due to the need for complying with strict data security policies

Edge Computing is transforming how organizations manage, process, and leverage data. Danish companies, due to a generally high degree of digital maturity, are rapidly increasing their engagement in Edge Computing, aiming at becoming more responsive to continuously changing conditions, whether they concern manufacturing processes or customer demands.

The purpose of this whitepaper to address the growing interest and curiosity around Edge Computing in industry, providing an overview of definitions, potentials and implementation challenges.

1. Introduction

Computing started with the advent of mainframes – room-wide computers performing bulk data processing - in the late 1950s. Mainframes are an example of centralized computing: they were too large and expensive to be on every user’s desk, hence all users were bringing all their data to the same – central – computer, which was intended to process all of it. In the late 1960s, minicomputers - due to their reduced dimension and cost - decentralized computing power, moving it (almost) to the single user. Research labs were able to control their own experiments, factory floors their own processes, and so on. The pendulum moved all the way to the distributed side with the arrival of the PC in the mid-1980s, finally reaching the user’s desk. In the early 2000s the Internet of Things (IoT) got momentum, allowing further decentralization by connecting sensors, controllers and machines and making it possible for different connected objects (or “things”) to exchange data over the internet. In the late 2000s, Cloud Computing answered the need for storing and processing “in the cloud” the exponentially increasing amount of data catalyzed by IoT, and Cloud Computing platform became popular thanks to Amazon, Microsoft and Google as a flexible and scalable alternative to expensive hardware for storing data. The computing pendulum today is swinging towards Edge Computing, which aims at further decentralizing not only data generation (i.e. what IoT did), but also data storage and analysis, which the cloud – acting as a platform that collects all data coming from different connected objects – is currently centralizing.

Companies are directing their attention to Edge Computing challenging the often-dogmatic adoption of the cloud for every type of application, asking themselves:

- Can the cloud keep up with my data volume and response time needs?
- Is it necessary to and process all data in the cloud?
- Is there a way to select what is better to process in the cloud and what is better to process locally?

Applications that make use of artificial intelligence (AI) usually deal with vast amounts of data. This often entails significant costs in terms of transmission and storage in a central cloud service. However, in many instances it is not necessary to have a centralized storage of such data, as data is often not used by – or collected from - other parties. In this case, all (or a part of) data processing can be done at the “Edge” of the network, locally, not having to transmit data to the cloud through the network. This would help in significantly reducing costs and processing time (latency) as well as in reducing the risk of data security issues. The term Edge Computing refers, in fact, to the shifting of data processing closer to the data generation source.

1.1 Defining the Edge

Gartner defines Edge Computing as “***a part of a distributed computing topology in which information processing is located close to the edge – where things and people produce or consume that information***”[6]. In simpler terms, Edge Computing means running fewer processes in the cloud and in centralized data centers and moving those processes locally – on the “edge” - close to the devices generating data (see the figure below).

Edge	Core Infrastructure
• *On the device, generating the data* • *Dedicated computing device in environment* • *Micro-data center near where data generated* • *Hanging off device generating data* • *Gateway device*	• *Public cloud* • *On-premise data center* • *Off-premises data center owned by organization* • *Third-party off-site data center*

The definition of the edge is, however, perceived differently according to the user:

- Mobile network operators (MNOs) perceive the edge as the end of their radio access network (RAN).
- Data center operators may view the edge as the infrastructure deployed at key locations to minimize Internet latency.
- High-performance computing and edge server vendors view devices at remote sites, either within data centers or sitting in cages and closets, as their edge.
- Sensor technology vendors view their offerings as edge technology.

1.2 Basic Characteristics of Edge Computing

Edge Computing is characterized by some key features that must be considered in order to fully understand it and its capabilities and difference with legacy systems.

Connectivity

Connectivity is the basis of Edge Computing. Proprietary protocols and closed architectures, however, may lead to costly "lock-in" situations. This is why modern Edge Computing deploys open architectures that leverage different protocols such as WLAN, OPC-UA, MQTT, NB-IOT, 5G and others in order to avoid such "lock-ins" as well as to ensure interoperability with a variety of existing industrial application.

Constraints

Edge Computing products need to adapt – both from a hardware and software perspective - to harsh working conditions and operating environments, such as electromagnetic interferences, dust, explosions, vibrations, and current/voltage fluctuations. Moreover, industrial scenarios are characterized by very specific power consumption, cost, and space requirements which Edge Computing devices need to meet.

Distribution

The applications that run on sensor products are often tightly coupled to the hardware on which they run. Edge Computing resources de-couple applications from the underlying hardware and enable flexible architectures in which applications can move from one device to another. The main needs Edge Computing devices need to support are distributed computing, storage, and management of distributed resources, such as dynamic scheduling.

Data pre-processing and filtering

Transmitting and storing data to the cloud can be very expensive and inefficient. Data transmission architectures often rely setups where a remote server periodically requests data from a device, regardless its value changes. Edge Computing resources can pre-process data at the edge and only transmit when relevant, reducing data transmission and storage costs.

Edge analytics

While classic sensor products generally have limited processing capabilities and can only perform one specific function (e.g. generate data, transmit data, etc.), Edge Computing resources can both collect and preliminarily analyze data at the edge.

Consolidated workloads

Edge Computing resources are often equipped with hypervisors which – unlike proprietary real-time operating systems - abstract the operating system and the application from the underlying hardware. This enables an Edge Computing resource to run multiple operating systems and applications on a single device. This leads to workload consolidation, reducing the physical footprint of the computing resources required at the edge and can result in lower costs for device or equipment manufacturers that previously relied on multiple physical computing resources.

1.3 Key components of the Edge ecosystem

There are some fundamental components that have to be taken into account when discussing the edge ecosystem: these are edge node, edge cluster/server, edge gateway, edge sensor, edge device, edge router, edge enterprise, edge data center and edge infrastructure.

Edge node

Edge node is a generic way of referring to any entity where Edge Computing can be performed.

Edge cluster/server

An edge cluster/server is a general-purpose industrial computer located in a remote facility within a factory and typically used to run industrial application workloads and shared services.

Edge gateway

An edge gateway is an edge cluster/server which, in addition to being able to host enterprise application workloads and shared services, also performs network functions such as wireless connection, protocol translation, network termination, tunnelling or firewall protection. Although some edge devices can also host network functions, edge gateways are generally an independent entity.

Edge sensor

The edge sensor is the entity that processes the data sensed from the environment right at the point of data generation. This happens to increase the quality, or the efficiency related to data generation. For example, in a video camera, a way to increase the efficiency of data generation is to generate data only when motion is detected, as the goal of a video camera is to capture image changes overtime. Capabilities such as motion detection can eliminate the need to transmit useless data (i.e. a series of identical images which do not witness any movement) to the cloud for unnecessary and costly processing and storage. This functionality requires Edge Computing right at the sensor – point of data generation. In most cases, the computing capacity of edge sensors is minimal and is used for a very targeted feature.

Edge device

An edge device is a special-purpose piece of equipment that integrates computing capacity into a product. While these are used to perform different functions, data from such products – or “devices” - can be collected and analyzed with a low latency level to ensure their safe and seamless operations.

Edge router

An edge router is an entity responsible for transmitting data packages across networks. In fact, the edge router acts as the demarcation point between external and internal networks. Some edge routers are characterized by a built-in computing module and are able to host applications.

Edge enterprise

An edge enterprise consists of a location which is collecting Edge Computing resources that can be used by the whole enterprise. In a distributed enterprise environment with many branch locations, computing resources can be shared among the branches to drive economies of scale and simplify management. In this model, instead of deploying Edge Computing instances in each location, these can be implemented in a shared site connected to the enterprise network. This entails higher computing capacity and capabilities, which can be used for applications that require more processing power and resources.

Edge data center

Edge data centers consist of physical spaces used for managing data close to the data user. These intend to provide users with an alternative to the cloud for managing data concerning particularly sensitive tasks such as disaster management.

Edge infrastructure

The edge infrastructure is intended as the collection of all the above Edge Computing entities. This aims at computing, storing and transmitting data through edge resources, facilitating the distribution of data processing, the reduction of latency and of data transport and storage costs.

2. Why Edge Computing for Industrial IoT?

2.1 Edge Computing or Cloud Computing

Edge Computing and Cloud Computing are not meant to exclude each other; on the contrary, Edge Computing is excellent for enhancing the use of Cloud Computing minimising its cost and maximising its efficiency. Having said that, in order to choose what to delegate to Edge Computing and what to delegate to Cloud Computing, it is paramount to have an overview of the key differences between them.

- **Data processing latency:** The primary difference between Cloud Computing Edge Computing is the location where data processing occurs. In Cloud Computing, data is processed on a central cloud server, which is usually located far away from the source of information and requires the time-consuming transmission of data over a network. Edge Computing mostly occurs directly on the devices to which the sensors are connected or on a gateway located in the proximity of such devices.
- **Data processing capabilities:** Cloud Computing provides superior analytic and storage capabilities compared to Edge Computing.
- **Data processing cost:** Cloud Computing entails costs related to the transmission of data over a network, the use of processing power on the cloud to analyze data and of storage space on the cloud to store it. Edge Computing, on the other hand performs such processes locally and entails costs related to the proprietary hardware and computing power for processing and storing collected data.
- **Connectivity:** Cloud Computing requires continuous internet access, while Edge Computing does not rely on a network connection.
- **Security:** When it comes to Edge Computing, data are stored locally, while for enabling Cloud Computing data needs to be shared across a network and be stored on the cloud.

2.2 Advantages of using Edge Computing

When it comes to Industrial IoT, to connect thousands or even millions of devices directly to the cloud – to perform Cloud Computing - is often not feasible due to cloud-related issues such as costs, latency, security, and bandwidth. Edge Computing, by facilitating local ("at the edge") data processing tackles these issues, often providing a number of collateral advantages.

Improve performance by decreasing latency

Since the collection and processing of data is – in an Edge Computing scenario - a matter of milliseconds, Edge Computing acts as a fundamental support for improving the performance of several industrial processes which rely on low-latency responses, such as feedback loops in industrial control systems or autonomous cars. For instance, a safety control system operating an industrial machine must stop in a matter of milliseconds in a dangerous situation, if a human is too close for instance. The processing of human recognition by a sensor and the processing of the decision to stop the machine should not be delayed by transmitting data to the cloud, process it in the cloud and transmitting it back to the machine.

Reduce operational costs by reducing the volume of transmitted data

As Edge Computing performs part (or all) of the data processing locally, the volume of data that is eventually transmitted to the cloud is smaller. This results in a smaller bandwidth as well as in a small data transmission and storage cost. For instance, manufacturers monitor equipment parameters for enabling condition-based maintenance, looking for anomalies in the behaviour of the monitored equipment. Edge Computing can be used to identify – on the device – if the monitored parameters are exceeding a pre-determined warning threshold and, only in that case, send data to the cloud, generating the alarm situation.

Improve resiliency/availability by addressing network instability

Critical infrastructures require a high level of availability and resiliency to ensure safety and continuity of services. Edge Computing, by processing data locally, ensures these attributes. Gas leakage detection systems, for instance, must be able to operate without having to rely on the internet connection.

Improve data security by limiting network-related leaks

Countries are increasingly instituting privacy and data retention laws. The European Union's General Data Protection Regulation (GDPR) is a prime example. Any organization that has data belonging to an EU citizen is required to meet the GDPR's requirements, which includes an obligation to report leaks of personal data. Edge Computing can help these organizations comply with GDPR, as local data processing and filtering through an edge gateway can reduce the amount of sensitive and private information that is sent over a network. For example, instead of storing and backhauling surveillance video, a smart city can evaluate the footage at the edge and only backhaul the meta data.

Increase computational efficiency by reducing datasets dimension

Deploying machine learning algorithms on the edge, taking advantage of an gateway, means that computational processing will be performed on smaller datasets and not on larger and aggregated datasets. While this will limit the representativeness of the machine learning analysis outcome, it will also reduce the need for resources to perform such analysis.

3. Edge Computing use cases

3.1 Manufacturing

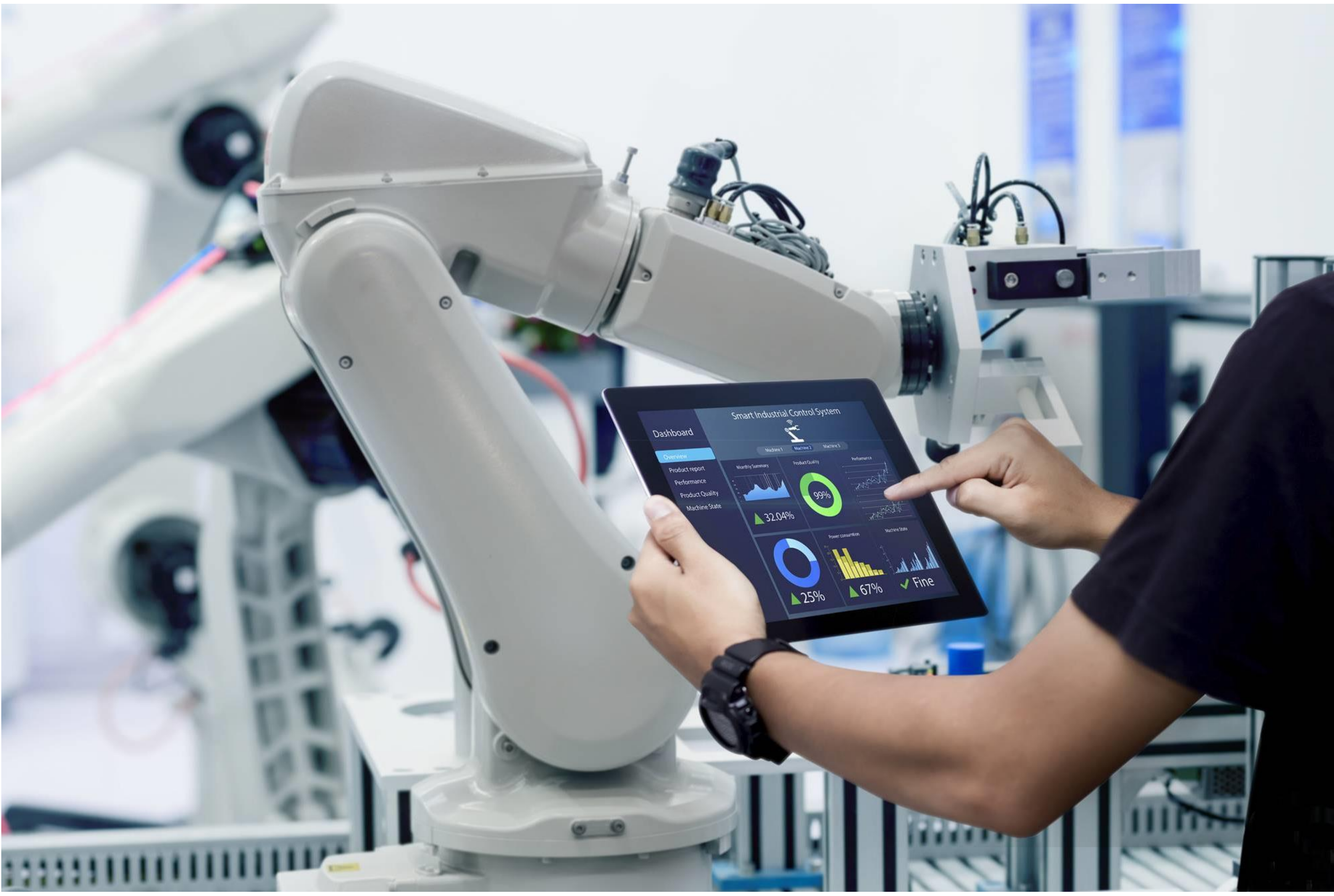
In the manufacturing field, Edge Computing is particularly relevant when it comes to controlling the collaboration between different machines interacting (and communicating in real-time) together or between machines and humans (e.g. collaborative robots, which need to continuously sense the surrounding environment to keep the humans safe during collaboration activities). These activities are strongly dependent on the continuous availability of data as well as on a low latency. Edge Computing would then provide a particularly valuable support in production environments where network access is non existing, intermittent, or limited by cellular or satellite bandwidth, such as in oil and gas rigs. Quality control activities or parameters monitoring for enabling condition-based maintenance are two additional examples of how Edge Computing could support manufacturers in their daily operations. In fact, a timely response to quality or operational issues is essential to reduce product defects and improve efficiency. The increasing diffusion of machine learning algorithms for analyzing monitored quality and equipment condition data and adjust the operating parameters and plan maintenance interventions accordingly will highlight the relevance of Edge Computing even more.

3.2 Transportation

Most mobile assets nowadays rely on a navigation system. When it comes to mobile robots, this is based on data that the robot collects and processes itself, on the edge. This enables the robot to move autonomously, delivering items from one point to another without bumping into obstacles. These navigation systems are typical based on lidar or camera technology. However, in order to receive orders (i.e. tasks to perform), the robot needs to communicate with a cloud-based data center/MES system. Such communication is also used to make sure that the robot is taking advantage of its navigation system effectively – to put it simply, it is always finding the best route to get from A to B. To do so, navigation data from each single mobile robot are collected centrally and compared with the navigation data from all the other mobile robots; insights obtained from such comparison – and used to improve the navigation abilities of the robot as well as to establish common routings – are transmitted back to the robot.

3.3 Surveillance

Surveillance cameras, whether they are deployed to ensure the security of a location or the quality of the output coming from a production line, work in similar ways: they often record a continuous stream of footage, which is then sent to the cloud for processing or storage. To cut down on the amount of stored data, the cloud server may, for example, automatically delete useless footage (e.g. where movements are not detected). Although this would optimize the required storage space, video data would still be transmitted to the cloud and processed taking advantage of Cloud Computing resources. This entails transmission and computing power costs regardless the utility of the video. Edge Computing would, on the contrary, perform the processing (and deleting of useless video) locally, transmitting to the cloud only useful video data to be stored, hence minimizing the related cost. Drones – another tool deployed for air surveillance - are performing visual search, image recognition, object detection and tracking. As for surveillance cameras, to take advantage of Edge Computing allows an increase of cost-efficiency related to the use of computing power (locally instead of cloud-based) – whether related to the recognition of vehicles, structural issues on bridges and buildings, fires or radiation spread - and data transmission.



3.4 Energy

The energy sector is one of the most prominent adopters of Industrial IoT, driven by regulatory compliance and by the business potential linked to efficiency improvements. These are strongly based on the use of data, which is used for monitoring and managing the various functions related to an energy grid distribution infrastructure as well as for maintaining it, taking advantage of real-time equipment diagnostics, as well as for billing the customers according to their actual consumptions. Edge Computing is particularly relevant in this sector due to the high volumes of data and the wide and various geographic range across which data is being generated

3.5 Consumer products

Customer behaviour data is a popular data source used to generate business insights. Companies (and their marketing and sales departments) are increasingly taking advantage of data that is generated by the use of connected devices such as smartphones and wearables, but also cars, laptops, televisions and washing machines. Another example concerns fraud detection, which increasingly relies on real-time data processing at the point of payment to ensure fraud is detected and acted on immediately. To achieve this, data processing needs to be carried out as close as possible to the data generation point, hence the relevance of Edge Computing.

4. Challenges for implementing Edge Computing

The implementation of Edge Computing entails some challenges that users need to be aware of:

- **Scalability:** Each Edge Computing location requires its own monitoring system to ensure an overview of its health and status, including power consumption, cooling performance and network status. A highly scalable monitoring infrastructure should be considered to be able to have an overview of the entire Edge Computing environment and of the interdependencies between different Edge Computing locations.
- **Control Access:** Edge Computing locations may be challenging to be accessed physically in order to control the IT equipment and assess, isolate, and solve issues. Access points and well-defined access practices should be established.
- **Security:** If Edge Computing reduces data security issues concerning the transmission of data through a network, local data processing increases the security responsibility of the devices where data are processed. The ability to identify and shut down compromised devices in the event of an attack becomes paramount.
- **Network Bandwidth:** As more data is stored and processed at the edge, bandwidth requirements should be adapted – if this is generally larger the closer you get to data centers and narrower the closer you get to the devices, in the case of Edge Computing network bandwidth is more balanced between the two endpoints.

5. Conclusions

The ability to collect, analyze and ultimately take advantage of large amounts of data is, nowadays, a must for production companies. If cloud seemed to provide all the answers, its extended adoption also highlighted some of its limitations – the main ones being related to latency, cost, security. Edge Computing is emerging to address such limitations. Our investigation highlighted how industry – in particular industrial domains that transmit, store and analyzes large data volumes, need low latency levels or have to comply with particularly high security standards – should start considering Edge Computing, not as a substitute to the cloud, but as an enabler for a more efficient use of the cloud and for a more effective use of the data. The cloud will continue to play a critical role in aggregating important data and performing analyses on a massive set of information. Nevertheless, the ability to analyze data closer to the source – Edge Computing - will support companies, when needed, in minimizing latency, improving privacy and security and lowering data processing costs.

References

1. Coughlin, T. (2018, November 29). *175 Zettabytes By 2025*. Forbes. Retrieved in March 2021 from: https://www.forbes.com/sites/tomcoughlin/2018/11/27/175-zettabytes-by-2025/?sh=750a5d735459
2. Andrade, F. (2019, July 12). What is an edge device, and why is it so crucial to IIoT? Retrieved in March 2021 from: https://netilion.endress.com/blog/what-is-edge-device-iiot/
3. Basecamp | IEC. (n.d.). Edge intelligence. https://www.iec.ch/basecamp
4. European Commission (2020). Building an ecosystem where IoT, edge and cloud converge towards a computing continuum. Shaping Europe's Digital Future - European Commission. Retrieved in March 2021 from: https://ec.europa.eu/digital-single-market/en/news/building-ecosystem-where-iot-edge-and-cloud-converge-towards-computing-continuum
5. Gartner (2021). Business and Technology Trends - Smarter With Gartner. Retrieved in March 2021 from: https://www.gartner.com/smarterwithgartner/
6. Definition of Edge Computing - Gartner Information Technology Glossary. (n.d.). Gartner. https://www.gartner.com/en/information-technology/glossary/edge-computing
7. Coughlin, T. (2018, November 29). 175 Zettabytes By 2025. Forbes. Retrieved in March 2021 from: https://www.forbes.com/sites/tomcoughlin/2018/11/27/175-zettabytes-by-2025/?sh=750a5d735459
8. Editor's Choice. (2021, March 1). What does the future hold for cloud and edge computing? Information Age. Retrieved in March 2021 from: https://www.information-age.com/what-does-future-hold-for-cloud-edge-computing-123494089/
9. i-SCOOP. (2021, January 18). The expanding and changing impact of IoT data on IT infrastructure. Retrieved in March 2021 from: https://www.i-scoop.eu/internet-of-things-guide/iot-it-infrastructure/
10. Longbottom, C. (2020, October 8). How to implement edge computing in 5 steps. IoT Agenda. Retrieved in March 2021 from: https://internetofthingsagenda.techtarget.com/tip/How-to-implement-edge-computing-in-5-steps
11. Matthews, K. (2020, July 28). How edge computing will benefit from 5G technology. Information Age. Retrieved in March 2021 from: https://www.information-age.com/how-edge-computing-will-benefit-from-5g-technology-123485756/
12. Meulen, R. (2021, February 4). What Edge Computing Means for Infrastructure and Operations Leaders - Smarter With Gartner. Copyright (C) 2021 Gartner, Inc. All Rights Reserved. Retrieved in March 2021 from: https://www.gartner.com/smarterwithgartner/what-edge-computing-means-for-infrastructure-and-operations-leaders
13. Security, H. N. (2019, June 19). 41.6 billion IoT devices will be generating 79.4 zettabytes of data in 2025. Help Net Security. Retrieved in March 2021 from: https://www.helpnetsecurity.com/2019/06/21/connected-iot-devices-forecast/
14. Taylor, A. (2019, April 23). Edge computing is in most industries' future. Network World. Retrieved in March 2021 from: https://www.networkworld.com/article/3391016/edge-computing-is-in-most-industries-future.html

15. Unthinkable Software. (2019, October 27). Cloud vs Fog vs Edge Computing: 3 Differences that Matter. Retrieved in March 2021 from: https://devisha-singh.medium.com/cloud-vs-fog-vs-edge-computing-3-differences-that-matter-68612cfc65e2
16. IBM. Why organizations are betting on edge computing. Retrieved in March 2021 from: https://www.ibm.com/thought-leadership/institute-business-value/report/edge-computing