International Journal of Academic Research and Development
ISSN: 2455-4197; Impact Factor: RJIF 5.22
Received: 25-06-2019; Accepted: 28-07-2019
www.academicjournal.in
Volume 4; Issue 5; September 2019; Page No. 116-121

Edge computing in industrial manufacturing: Systematic approach to implementing sustainable edge architectures

Mahendraprabu Sundarraj
Enterprise IT Architect, Information Technology, Sunpower Corporation, USA, Texas, USA

Abstract
Edge can be defined as a way either to bring the content closer to the customer or to bring the computing power closer to data sources. The second definition – bringing the computing power closer to data sources – revolutionized the Manufacturing industry. Though Edge computing is not new to the production floor, Manufacturing companies were traditionally relying on local data centers for the compute power than the edge. Since the data did not travel far from the production floor (only up to the local data center), companies did not have high latency issues and were never motivated to explore ways to shorten the system latency. IoT and Cloud expanded these imagination boundaries by bringing in advanced analytics possibilities. Advanced analytics capabilities require the data to travel far from the production floor and brought the system latency as a new challenge. Edge is now seen as a possible way to reduce the overall system latency. Though there is much literature on Edge computing and its advantages, Very little has been written on practical and systematic approaches to avoid the common pitfalls encountered in implementing Edge. This article discusses the advantages of Edge Computing in Manufacturing companies, common pitfalls organizations expected to encounter in implementing Edge projects, and systematic approach to overcome those challenges and create sustainable edge architectures.

Keywords: Edge computing, Smart Factory, IoT, Industry 4.0

1. Introduction

Cloud IT infrastructure, which emerged in the 2000s, enabled every organization to scale their compute and storage capacity in no time. Ability to scale the IT infrastructure without having to worry about hardware and its maintenance helped organizations to increase the speed of innovation and advancement in their respective industry. It is not an overstatement to say cloud service providers now hold the IT technology roadmap of the world. In the same period in which cloud emerged, IoT (Internet of Things) was enabled thanks to the convergence of commodity sensors, embedded systems, real-time analytics, and wireless sensor networks [1]. Evolution of IoT (a machine to machine communication network) [1] was also strengthened and expedited by unlimited storage and compute offered by Cloud services.

Power of IoT and cloud combined, Manufacturing industry saw the rise of smart manufacturing, which led to real-time optimization of manufacturing and supply chain operations [2], ability to respond to product demands dynamically, and decreased production downtime. In the early stage of smart manufacturing, companies started to centralize most of their computing needs – real-time analytics, machine learning and intelligent decision making - in cloud. All production lines started streaming their data to cloud storage, data were analyzed in the cloud, and intelligent decisions were pushed back to the production line from cloud applications. Compute power of each machine – Edge computing - is mostly gone unnoticed.

Edge computing is not a new paradigm, and other industries started using Edge computing even before manufacturing industry did. Origin of Edge computing can be traced to late 1990s when Akamai used CDN (Content Delivery Networks) to increase the performance of web content delivery [3]. In Akamai's use case, Edge computing – placed near the data consumer - was used to distribute data faster from a centralized data source. However, Edge computing can also be seen as the compute power that is placed closer to the data sources help to collect the data faster and deliver to a centralized data store. Using Edge in both these ways mentioned above helps to create sustainable and scalable IT infrastructure for smart factories.

Smart manufacturing requires IT infrastructure to support real-time data processing and analytics. As the production volume goes up, the complexity of computing workloads will also increase, and IT infrastructure should be scalable to support. Scalability comes with the cost, and the Manufacturing industry is not known to spend a lot of money on IT. As per industry research reports, manufacturing industry spends much less (1.95% of the total revenue) [4] than its peers. Less spending is justified given that manufacturing and supply chain companies make very less margin across industries. Hence smart manufacturing IT infrastructure should be built sustainable – not just in terms of performance but also in terms of cost to scale. Edge computing presents an opportunity to create these sustainable architectures.

This paper has four sections. The first section discusses the Edge and its importance in manufacturing. The second section discusses the common pitfalls factories encounter in leveraging Edge. The third section puts forward a systematic approach to avoid pitfalls and maximize the benefits of Edge. The fourth section reviews the future direction and challenges to be solved.

Edge-importance in manufacturing

Definition of Edge is different from industry to industry, and it can be different per use-case within a given industry. Edge can be generalized as a device that generates or collects the data from a source (source should be close to Edge). In web content delivery, CDN (content delivery networks) can be considered as Edge or even an end-user laptop. Mobile phones or communication towers closer to consumers can be considered as Edge in telecommunications.

In the Manufacturing industry, Edge is generally umbrella term which combines Edge devices such as production machines, Sensors installed at factories, or Local Datacenters at factory sites. All Edge devices may not have the native storage and compute. Most of the new age smart production devices are coming up with the native storage and compute.

Cloudlets, a compute layer which is closer to the production line, helps to extract the data from production lines and export to cloud [5]. Cloudlet enables the legacy production tools and sensors, which do not have native compute, to stream data to cloud and help organizations to get all the benefits of cloud computing power without having to invest in smart robotic automation. Almost all cloud service providers today have released their cloudlet equivalent versions to get as much data as they can to cloud. Acquisition of Express logic (Real-Time Operating System vendor for IoT devices) by Microsoft [6] is an example.

Below diagram illustrates the what edge in a smart factory is

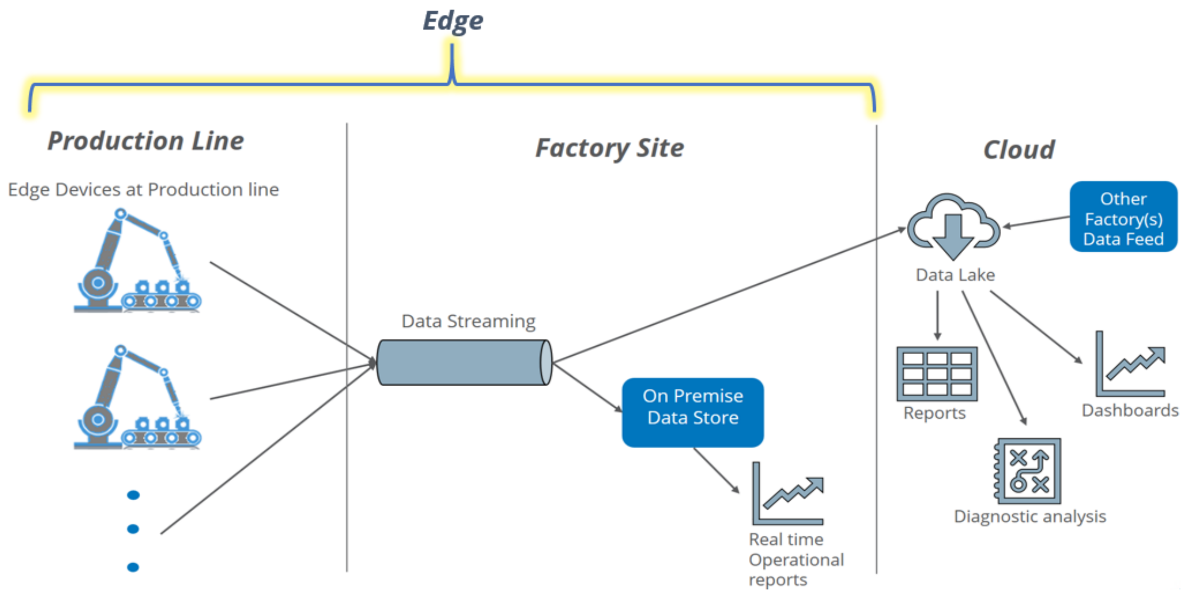

Fig 1: Edge in a Smart Factory

Production Line: Machine tools, Sensors, Bar code scanners, RFID scanners, Cameras, and any other data capturing device connected to the production tool are commonly known as edge devices. These devices may or may not have native storage and compute. If the factory is setup recently in the last decade, there is high possibility that these edge devices have native storage and compute. Conventional production lines can make use of cloudlets to compensate for the loss of native storage and compute.

Factory Site: On-premise data centers at the factory site-level usually host the Manufacturing execution system and Operational analytics systems. They have more powerful compute engines than the production line edge devices have. These local data centers, given the proximity to the production lines, can also be considered as Edge. Factory site compute power is limited and scaling up the IT infrastructure is costlier than scaling up in the cloud

Cloud: Cloud can be broadly classified into two types – Public and Private cloud. Private cloud is optional. A global organization typically has a private cloud to host its ERP (Enterprise resource planning, CRM (Customer Relationship Management), PLM (Product life cycle management), and DMS (Document Management system) software tools. Public cloud is usually used to host Data lakes, Archived data marts, Visual analytics tools, any many digital applications. Today, Many organizations entirely rely only on public cloud services and may not have a private cloud at all.

As it was called out in the Introduction section, the Manufacturing industry spends significantly less on IT despite the need to scale up IT infrastructure continuously to stay competitive. Cost-effective IT infrastructure is critical for any organization to continue their smart factory journey and stay competitive.

As per Mahadev Satyanarayanan (Carnegie Mellon University), Edge computing provides critical advantages such as Highly responsive cloud services, Scalability via Edge analytics, Privacy- Policy enforcement, and Masking Cloud outages [5]. Each of these advantages is common to all industries.

Benefits of Edge Computing in Manufacturing can be broadly classified into five groups.

1. Availability
2. Speed
3. Cost-effective
4. Security
5. Quality

Availability: Production is typically level loaded, and production outages are costly. If any step in the production process is dependent on cloud computing, it can be a potential failure point to cause production outage if there are a network failure or cloud service interruptions. If architected correctly, Edge computing can mask cloud outages and save production outages. Most of the cloudlets – a type of Edge computing – has almost all features of cloud computing and can act as a failover node. In Oil & Gas and shipping industries, continuous connectivity with cloud may not be possible. E.g., a Cargo ship in the middle of pacific or Atlantic ocean en route to North America. Cloudlets which are locally installed can collect the data and stream to cloud when the connectivity becomes available. Cloudlet computing and storage capacity can be planned by calculating amount of data expected to be generated during expected outage time. Thus the cloud outages can be masked by Edge computing.

Speed: Collecting the data and streaming the data without additional preparation to the centralized cloud can be a costly exercise when the data volume increases. Once the raw data is collected in a centralized data lake in the cloud, further preparation would be needed to achieve the desired data quality before the advanced analytics processes start consuming the data. Overall latency of such a system can be calculated by adding three values: Time to extract and upload the data cloud (T1), Time to clean up and prepare the raw data (T2), and Time to analyze the data and make further decisions (T3). Overall system latency can be significantly reduced if the data preparation is done closer to the data source in small batches rather than in a big batch in a centralized data lake. If such a change implemented, there will be a small increase in the T1. However, T2 will be eliminated entirely. Once the data is uploaded into cloud, it will be immediately available for consumption, leading to much faster analytics thanks to Edge computing.

Cost-effective: Though cloud computing simplified the IT infrastructure management for any organization, it is not cheap either. As the data and compute complexity goes up, cloud cost will also go up. In order to keep the overall infrastructure cost under control, One can consider Edge computing as a viable alternative to some of the workloads run in the cloud. As explained before, if the data preparation tasks are handled at edge computing layer, we will not only achieve low latency but also save cost by cutting the cloud computing need. If the data is compressed at Edge before uploading to Cloud, we will not only decrease the time to upload but also save cost by uploading a smaller than original size file [7].

Security: Data compression and encryption at Edge before uploading to the cloud improves the data security significantly. Data security should exist at all stages of data lifecycle: data when collected, data while being sent, data when persisted in a data mart, Data when consumed. Achieving complete data security may not be possible without the help of edge computing. While collecting the data itself, Edge can embed access policies on the data to ensure security for the rest of its lifecycle.

Quality: Efficiency of Sensors and wireless networks are improved every day. However, data loss while collecting real-time event data is almost inevitable, and it can seriously affect data quality. Poor data quality will lead to much weaker analytics and false signals. Once the data is consolidated in a centralized data lake, data cleaning can be a complicated task given that Data lake is not close to the data source. However, data cleaning and extrapolation for the missing data can be a much easier task at the Edge level given its proximity to data source. If the centralized metadata repositories [8] are used to ensure data quality at the enterprise level, Data quality policies and expectations can systematically be passed to edge devices or cloudlets to make sure that data quality starts from Data collection.

Challenges in Edge implementation

Scalability: Though new age smart edge devices provide compute power, it is not unlimited. Maximum usage of Edge computing can risk the production operation if not planned carefully. Edge compute attached to the machine tool is primarily made available for the efficient operation of the machine tool. Leveraging this Edge compute in handling workloads such as data preparation and data upload were not the primary reasons for the existence of edge computing. Hence these workloads should not overstep on the primary purpose of Edge computing. Before planning to deploy any software on edge devices or connecting to edge devices, detailed design discussion should be held with the device vendor or the device owner. Below is the simple formula to calculate how much Edge computing can be used for any non-primary workloads.

Available Edge compute capacity (Ae) = (Maximum Edge compute capacity (Mx) – Maximum Edge device compute usage (Mu))

Maximum usable Edge compute capacity (MAe) = (Ae) X *(P)*

Here (P) is the percentage how much maximum available edge compute capacity can be used. This value (between 0 to 1) is agreed upon by all stakeholders – Edge device owner, Research and development team, IT Architects, Production Operations team, and Quality engineering team. Value (P) continues to be impacted by various factors such as production volume increase or decrease, network speed, software upgrades, and new business requirement to change the latency of data replication.

Failure to have a systematic and continuous process and monitoring of Edge compute usage would lead to scalability uses and unplanned production downtimes.

Poor Data Quality: Edge can help to increase the data quality significantly by structuring the data closer to the data source. However, if there is no centralized metadata repository, which holds the implementing and data governance policies at enterprise level – managing the data preparation consistently across edge devices and different factory sites would be hard. Metadata repository helps to create a federated data governance by enforcing the data governance policies across the enterprise including edge devices [8]. Edge compute implementation without a metadata repository may not yield results in terms of data quality.

Lack of Standards: Lack of standards in network protocols, data or machine log format, Integration patterns to integrate edge devices, and software (Operating systems, Databases, and APIs) installed in Edge devices can have negative impacts on maintenance and ability to scale. Standardization is often thought as a helpful exercise meant only for major manufacturing organizations who have multiple smart factory sites. However, an organization which has only one factory site can get significant benefits by having standards.

Lack of monitoring: If an edge system is not actively monitored, it can never be sustained. If monitoring of any part of edge devices is not possible, then the solution should not be approved to be implemented. Monitoring is never an afterthought; it should be part of the overall strategy and high-level capabilities that Edge is expected to offer. Most of the factories are level loaded, and they indefinitely run without any stop. Edge devices, given their proximity to production, can cause factory downtime. Lack of monitoring increases the possibility of unplanned factory shutdowns.

Lack of security: OWASP (Open web application security project) lists top 10 IoT security issues. OWASP 2018 report [9] lists issues which mostly caused by incorrect or insufficient edge implementation. Below is the list.

1. Hardcoded and insecure passwords for Edge devices
2. Insecure network protocols exposing the edge to internet
3. Insecure APIs or middlewares, which do not have efficient authentication, encryption, and input & output filtering, compromises the edge security
4. Lack of strategy to keep the edge software up to date on security firmware increases the edge security vulnerabilities
5. IoT devices are not often changed, and outdated components can become a security risk.
6. Lack of protection for personal information stored in Edge device
7. Lack of access control or encryption of sensitive data anywhere from edge to cloud
8. Lack of device monitoring and response capabilities
9. Insecure default settings
10. Lack of physical hardening

All the above security issues are mostly caused by lack of security strategy at the time of edge implementation. It is easier to overlook the security issues when organizations rush to get the advantages of their existing edge computing. It is easier to address the security issues at the time of implementation rather than later.

A systematic approach to Edge Implementation

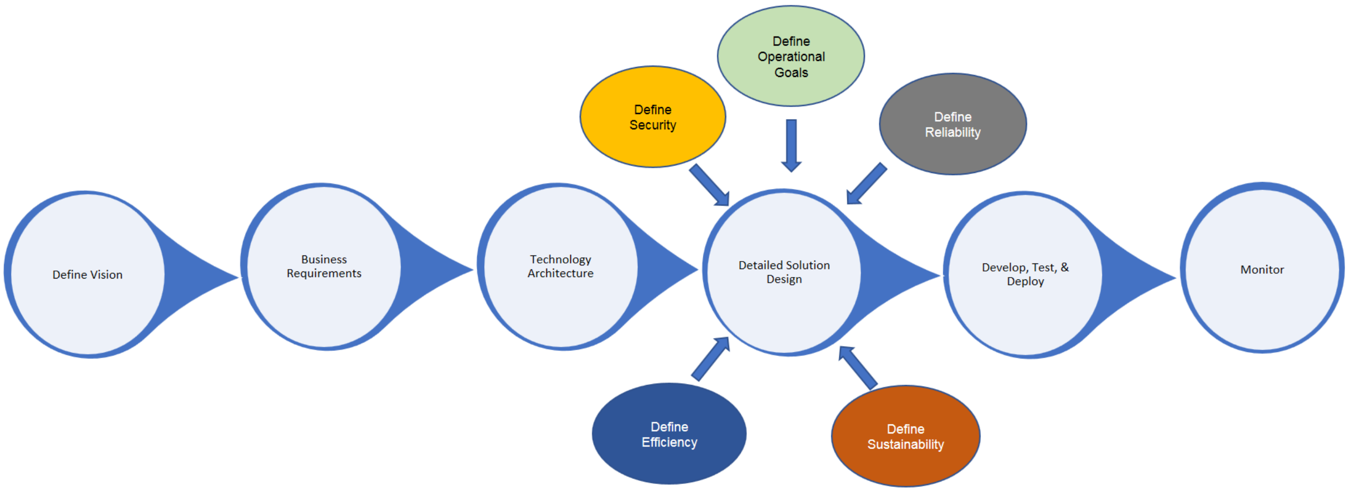

Fig 1

Above picture proposes the systematic process for enabling edge computing in industrial manufacturing. This process can be used for the first time edge implementation project or enhancing the existing edge compute further. In the case of enhancement projects, few steps such as Define Vision or Technology Architecture can be considered as optional.
The process is designed in such a way to make sure that the edge computing capabilities are built sustainable, scalable, and modular. The core part of this process is “Detailed Technical Design” step which aims to address most of the common mistakes that happen in Edge implementation. Below table explains each step’s deliverable and the text follows explains the core step of the process in detail.

Table 1

Step	The deliverable from the step
Define Vision	Set of high-level goals which would set the scope and expectations for the project team
Business Requirements	List of detailed business requirements, minimum criteria for successful implementation
Technology Architecture	High-level solution design which illustrates technology building blocks needed to achieve the business requirements defined in the previous step
Detailed Solution Design	Detailed solution design after the due discussion on the below. 1. Operational goals 2. Reliability 3. Sustainability 4. Efficiency 5. Security

Develop, Test, and Deploy	Workin system without any defects
Monitor	Continuous monitoring of the implemented Edge compute setup to achieve high availability, less cost, and high efficiency

Detailed Technical Design

Define Operational Goals: Edge compute implementation is not a one-time activity but a continuous activity. Scaling up or down the existing infrastructure or redistributing the load among devices is almost a daily activity. Onboarding new devices just in time are necessary, but at the same time, it should not affect the ongoing factory production. Testing and onboarding new devices automatically and programmatically managing devices are few important operational goals of edge projects. These operational goals should be defined, and detailed technical design should have the solution building blocks to achieve these goals.

Define Reliability: All failure points of the technical solution should be identified and documented. Security challenges such as Emergency security threats and hardware-specific hacks are to be defined. High availability goals and disaster recovery plans are to be defined according to the detailed technical design.

Define Sustainability: Scalability in Edge comes with the cost, which can slow the speed of continuous improvement in factories. Hence Detailed technical design should consider sustainability goals and select the technology building blocks. For example, Changing the file format at Edge device from a simple log file to optimized columnar format file can save high ETL cost in data preparation. Similarly, compression of data not only increases the speed of data transfer but also decreases the cost of the network (internet) and storage. These sustainability-driven design decisions are better taken at the time of detailed technical design.

Define Efficiency: Efficiency at different levels such as Edge devices, cloudlets, Enterprise network, or cloud should be defined, and detailed design should meet those efficiency targets. Knowing these efficiency goals at the design stage is essential to make sure the foundational architecture is scalable and sustainable. For example, Data transfer from an edge device to cloudlet can easily be enabled using an open-source ETL tool. ETL tools work typically better in batch mode data transfers and point to point data connections. If there is an expectation that real-time operational analytics would come in future and factory would open similar production lines, then relying on an on-premise ETL can become a bottleneck to scale up to new efficiency goals.

Define Security: Data security at all stages of data processing is critical. Access controls, secured credentials, and encryption should exist in all stages, including edge devices. Usually, Edge devices and software are not updated frequently. However, it is crucial to keep the software up to date. If the security goals are defined correctly, Detailed Design can include features to monitor edge devices and install security updates remotely without disrupting the factory production.

Future Challenges

Edge computing, with all the advantages, still has scope improve further. Below are some of the critical areas where the improvements are expected to come and benefit industrial manufacturing.

1. **Security:** Though the security can be ensured to a great extent at the industrial network or Edge level, there is lot to learn how to handle the data security in the cloud. Most of the manufacturing companies are still in the early stage in migrating their workloads to the cloud, and best practices are yet to emerge.
2. **Open Standards:** Open standards in Edge is still a distant dream. Though there are few non-profit organizations made significant progress in defining open standards, adoption across companies is still in early stage. Given the cost to change, adoption is slow; new technological improvements help to accelerate adoption would bring in sweeping improvements in industrial manufacturing.

Conclusion

Much has been written about edge compute and the opportunity it brings. This paper reviews all the advantages in the Industrial manufacturing context and lists the common failure points in Edge implementation. The author has used the systematic process proposed in this paper in multiple edge implementation projects across the globe – Philippines, Mexico, and Malaysia – with significant success.

References

1. Internet of Things. https://internetofthingsagenda.techtarget.com/definition/Internet-of-Things-IoT. Accessed on, 2019.
2. Ersue M, Romascanu D, Schoenwaelder. Sehgal A. Management of Networks with Constrained Devices: Use Cases, 2014. IETF Internet Draft https://tools.ietf.org/html/draft-ietf-opsawg-coman-use-cases-02
3. Dilley J, *et al*. Globally Distributed Content Delivery, IEEE Internet Computing. Khalid Kark, Anjali Shaikh, and Caroline Brown "Technology budgets: From value preservation to value creation" Deloitte. USA, 2017. [Online]. Available: https://www2.deloitte.com/insights/us/en/focus/cio-insider-business-insights/technology-investments-value-creation.html Accessed on: Dec. 5, 2018
4. Mahadev Satyanarayanan. Carnegie Mellon University, The Emergence of Edge Computing, IEEE computer

society.
5. Microsoft acquires Express Logic, accelerating IoT development for billions of devices at scale, https://blogs.microsoft.com/blog/2019/04/18/microsoft-acquires-express-logic-accelerating-iot-development-for-billions-of-devices-at-scale/, Accessed on 17 Aug, 2019.
6. Mahendraprabu Sundarraj. A Sustainable Method to Handle Dark Data in a Smart Factory, Software Quality professional magazine, American Society for Quality
7. Mahendraprabu Sundarraj. Data governance in smart factory: Effective metadata management, International Journal of Advance Research, Ideas and Innovations in Technology.
8. Security issues in Edge computing and the IoT. https://resources.infosecinstitute.com/security-issues-in-edge-computing-and-the-iot/#gref. Accessed on 19 Aug, 2019