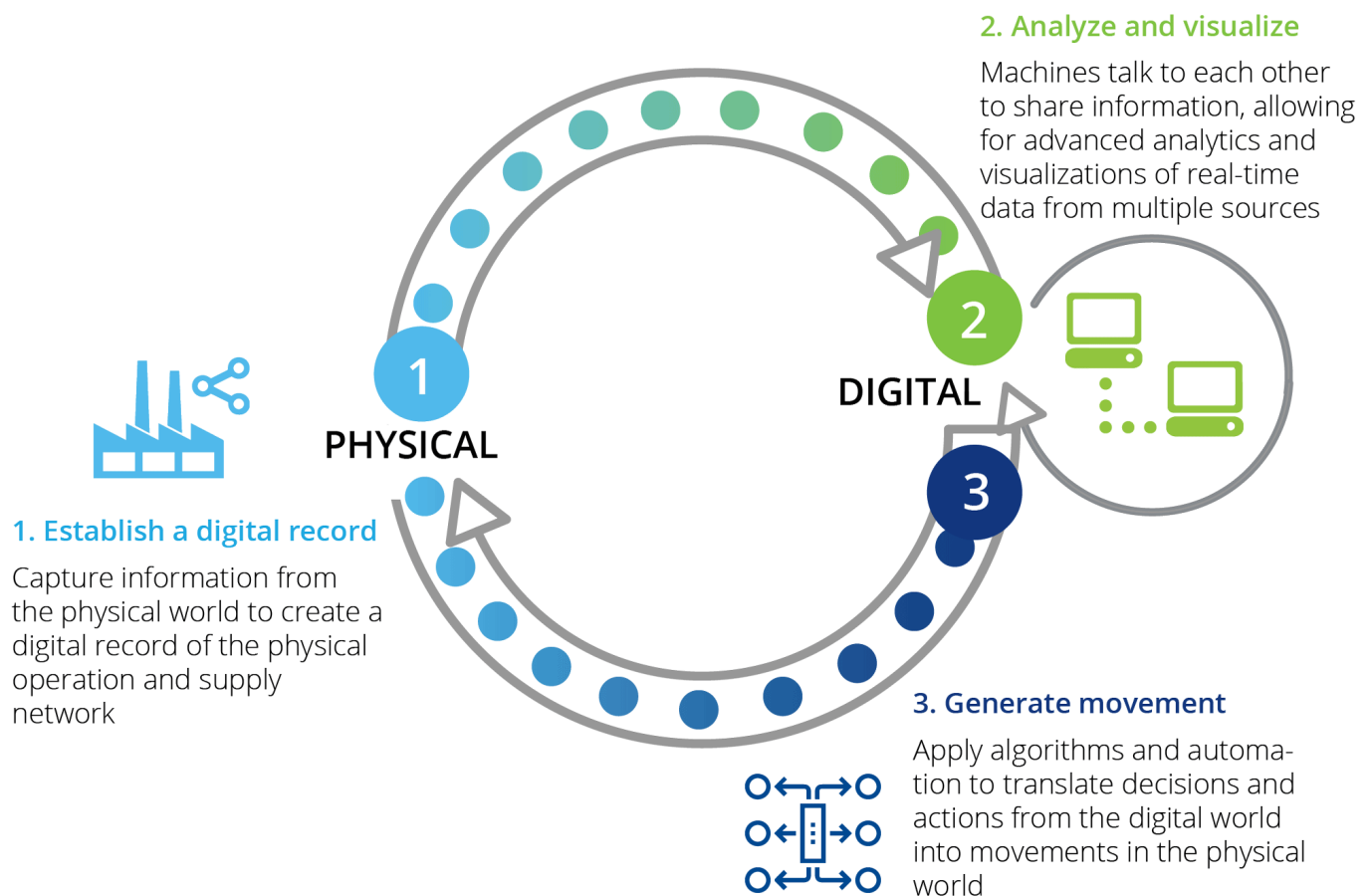# Digital Twins, IoT and ML - Introduction to industry 4.0

What is Industry 4.0

Traditional, linear manufacturing process uses embedded systems to control machines, for instance to manage speed of assembly line, or control CNC machine. Such manufacturing system has limited data concerning raw materials, process parameters and final products. In addition to that, collected data is analyzed periodically, which causes delays in process improvements. Industry 4.0 is a major transformation of the way production and manufacturing operates. It is the shift from linear process to integration and real-time data access. Cyber-physical systems enable collection, integration and advanced analysis of data from machines and systems and using insights to manage physical processes. Instead of operating reactively, as it has been always done, companies can learn along the way and use insights to adjust processes in real time. Such systems can self-optimize performance and self-adapt to new conditions in near-real time, and run entire production processes. Throughout the cycle data and actions flow continuously between physical and digital worlds. This cycle consists of three steps: • Physical to digital: Capture information from the physical world and create a digital record • Digital to digital: Use advanced analytics and AI to uncover meaningful insights • Digital to physical: Translate digital-world decisions to action and change in the physical world .

Figure 1. Physical-to-digital-to-physical loop and related technologies

2. Analyze and visualize

Machines talk to each other to share information, allowing for advanced analytics and visualizations of real-time data from multiple sources

DIGITAL

PHYSICAL

1. Establish a digital record

Capture information from the physical world to create a digital record of the physical operation and supply network

3. Generate movement

Apply algorithms and automation to translate decisions and actions from the digital world into movements in the physical world

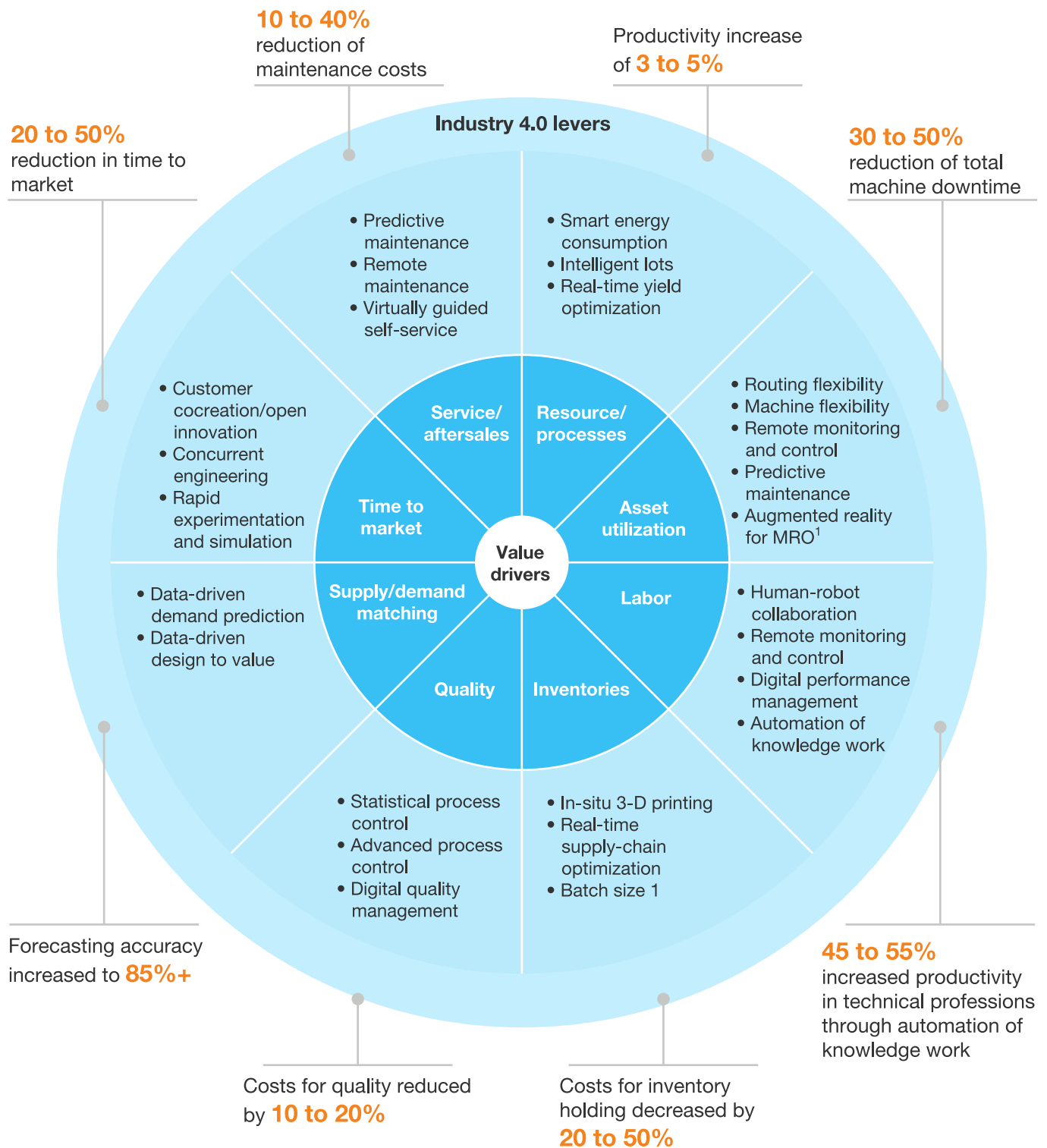Source: Center for Integrated Research.

Deloitte Insights | deloitte.com/insights

Source: https://www2.deloitte.com

In other words, Industry 4.0 is an integration of IoT – key connectivity enabler, IT technologies such as advanced analytics, artificial intelligence, augmented reality and physical technologies, like additive manufacturing, robotics and advanced materials. Such integration enables physical systems to communicate and cooperate both with each other and with human workers and act upon insights derived from data. Below table presents examples of technologies that are applied at each step of the physical-to-digital-to-physical loop.

How Industry 4.0 creates value

Thanks to digitization of operations, manufacturing, supply chain, and products companies can combine learnings from humans, machines, analytics, and predictive insights to make better decisions and create additional value. McKinsey has identified eight main value drivers that impact a manufacturing company's performance. McKinsey have found that applying Industry 4.0 leavers on each of these value drivers lead to improvements.

10 to 40%
reduction of
maintenance costs

Productivity increase
of **3 to 5%**

20 to 50%
reduction in time to
market

30 to 50%
reduction of total
machine downtime

Industry 4.0 levers

- Predictive
  maintenance
- Remote
  maintenance
- Virtually guided
  self-service

- Smart energy
  consumption
- Intelligent lots
- Real-time yield
  optimization

- Customer
  cocreation/open
  innovation
- Concurrent
  engineering
- Rapid
  experimentation
  and simulation

- Routing flexibility
- Machine flexibility
- Remote monitoring
  and control
- Predictive
  maintenance
- Augmented reality
  for MRO[1]

Service/
aftersales

Resource/
processes

Time to
market

Asset
utilization

**Value
drivers**

- Data-driven
  demand prediction
- Data-driven
  design to value

Supply/demand
matching

Labor

- Human-robot
  collaboration
- Remote monitoring
  and control
- Digital performance
  management
- Automation of
  knowledge work

Quality

Inventories

- Statistical process
  control
- Advanced process
  control
- Digital quality
  management

- In-situ 3-D printing
- Real-time
  supply-chain
  optimization
- Batch size 1

Forecasting accuracy
increased to **85%+**

45 to 55%
increased productivity
in technical professions
through automation of
knowledge work

Costs for quality reduced
by **10 to 20%**

Costs for inventory
holding decreased by
20 to 50%

[1]Maintenance, repair, and operations.

Source: "Industry 4.0: How to navigate digitization of the manufacturing sector,"
McKinsey Digital, 2015
McKinsey&Company
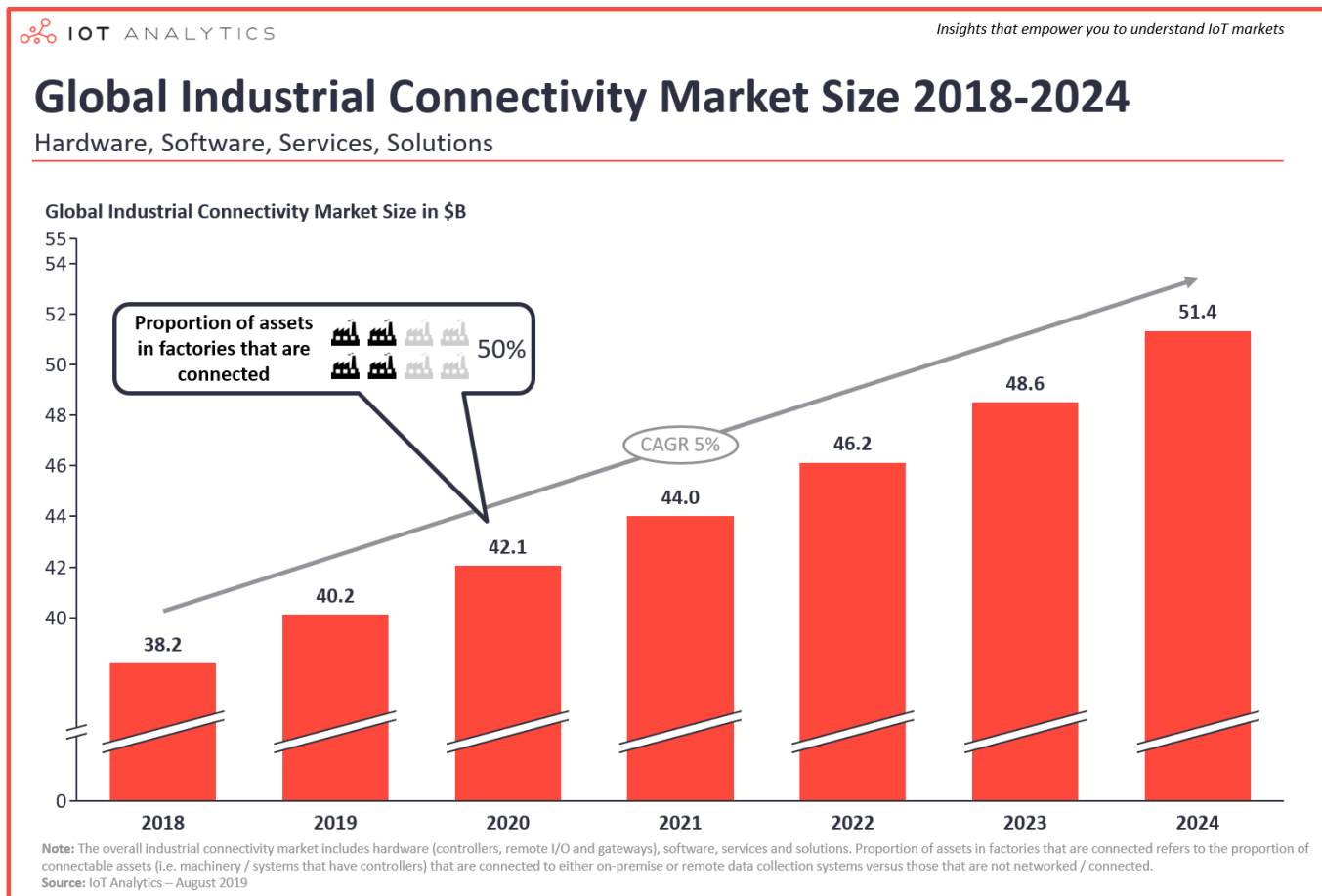
Source https://www.mckinsey.com/

Resources/process – Industry 4.0 mechanisms improve effectiveness of processes. ABB employed computer-based system to control and optimize cement kiln operations. System mimics behavior of 'ideal' cement plant operator. System uses actual measures to calculate process parameters required to meet target performance and adjusts process in real-time. Process optimization yields up to 5% improvement in throughput. Asset utilization – leavers like predictive maintenance can provide value by decreasing machine downtime or changeover times. GE10 offers predictive maintenance, where IoT sensors collect and report data on the machinery condition. Advanced analytics use this data to detect early signs of problems and indicate machines requiring maintenance. Predictive maintenance can reduce machines downtime by 30-50 percent and improve their life by 20 to 40 percent. Improving labor productivity – Industry 4.0 levers bring value by reducing waiting time or increasing speed of operations. Etalex, warehouse furniture manufacturer introduced robots to increase labor productivity. Company used robots to support workers in physically straining tasks. Human-robot collaboration increased throughput by 40%, without increasing employee base. Inventory optimization – Wurth's iBins uses intelligent cameras to monitor fill level of supply box. The box automatically reorders supply based on accurate fill data. Such real-time supply optimization reduce inventory costs by 20 to 50%.

Improving quality – Toyota uses advanced process control and data analytics tools for real-time errors tracking and correction. These measures result in minimizing rework and scrap. Applying Industry 4.0 leavers can decrease costs related to suboptimal quality by up to 20%. Supply/demand match – In order to maximize the value captured from the market company has to understand customer's demand. Industry 4.0 gives companies tools to use full value potential. One of OEM in automotive industry uses online configurator to identify the product options customers are willing to pay for. Such solution enabled significant reduction of relevant product options and decreasing time and costs of production.

Time to market – short time to market gives first-mover advantage. Industry 4.0 leavers can help to speed up development process. Example can be Local Motors. Company produces cars through 3D printing, with design crowd sourced from an online community. Company was able to reduce the development cycle from six years, being industry average, to one year. Additionally, they achieved significant R&D cost reduction. According to McKinsey, possible time to market reduction ranges from 30 to 50 percent. Service/aftersales – maintenance and repair are important drivers of service costs. Reduction of these costs for the customer opens potential for additional value creation. Secomea uses remote maintenance to provide service to its customers. Company uses software that allow technicians to establish remote connection to industrial equipment at customers premise and carry out diagnosis. Remote and predictive maintenance can reduce maintenance costs reduction by 10 to even 40 percent.
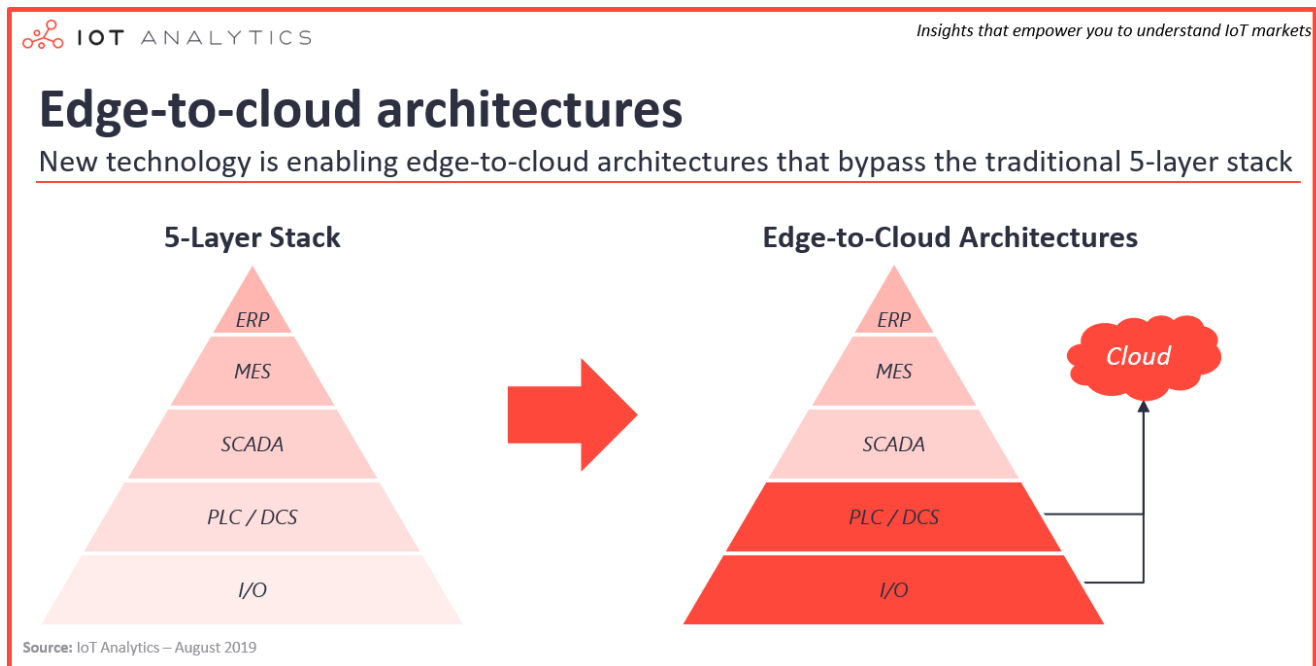
Market trends

Over 50% of industrial assets in factories will be connected Industry 4.0 is not a theoretical concept, but a real trend affecting how companies operate. An increasing number of manufacturers build industrial data collection solutions. According to IoT Analytics' report, by 2020, over 50% of industrial assets will be connected to some kind of data collection system. Percentage of connected assets is going to rise and will be key driver of industrial connectivity market growth.

Source https://iot-analytics.com/

Edge-to-cloud connectivity

In last couple years ERP, MES and SCADA systems moved to the cloud. Due to this fact new connectivity architectures, enabling direct edge to cloud connectivity emerged. Traditional communication through SCADA and MES systems depends on OPC servers. Advancements in connectivity technology (connectivity protocols), computing hardware (low cost edge computing that can run applications enabling edge to cloud connectivity) and software (software designed to run on edge devices as small as a Raspberry Pi) enable new architectures where edge devices can establish direct connection with the cloud.

Source https://iot-analytics.com/

Edge computing

Over recent years enterprise computing has been cloud-centric. There are many reasons for that. Cloud solutions are often cheaper, more powerful, easier to implement, maintain and scale. Storing data centrally and off-premise simplifies collaboration, remote working and flexibility. But business needs are changing. Emergence of 5G, increased popularity of IoT resulting in creating large amount of data and real-time analytics makes businesses look for alternative computing solutions. Despite many advantages, latency of makes hosted solutions insufficient for some use cases.

For time-critical processing users are moving towards edge computing that moves data storage and processing closer to the point it is needed. Placing computing power close to the sensors reduces amount of data sent to the cloud, which simplifies security and decreases network response times. Increased popularity of AI and advanced analytics Advanced analytics and artificial intelligence become more capable. Cloud solutions and improved computing capabilities makes them more accessible. Companies realize that advanced analytics and AI can create significant value when applied to manufacturing industry. Examples include predictive maintenance, digital quality management, and AI-driven demand forecasting. AI plays a key role in the smart factory, helping manufacturers predict demand and allocate resources. Wide adoption of digital twins

In simple words, digital twin is a virtual replica of physical object, process, or product that is updated in real-time. Digital twins can range from single piece of equipment to entire production line. Digital twins are combined with IoT sensors providing data like temperature, RPM, pressure and other critical process parameters. Companies can use data to run process simulations, react in real-time to parameter changes to prevent downtimes and optimize production process and supplies stock.
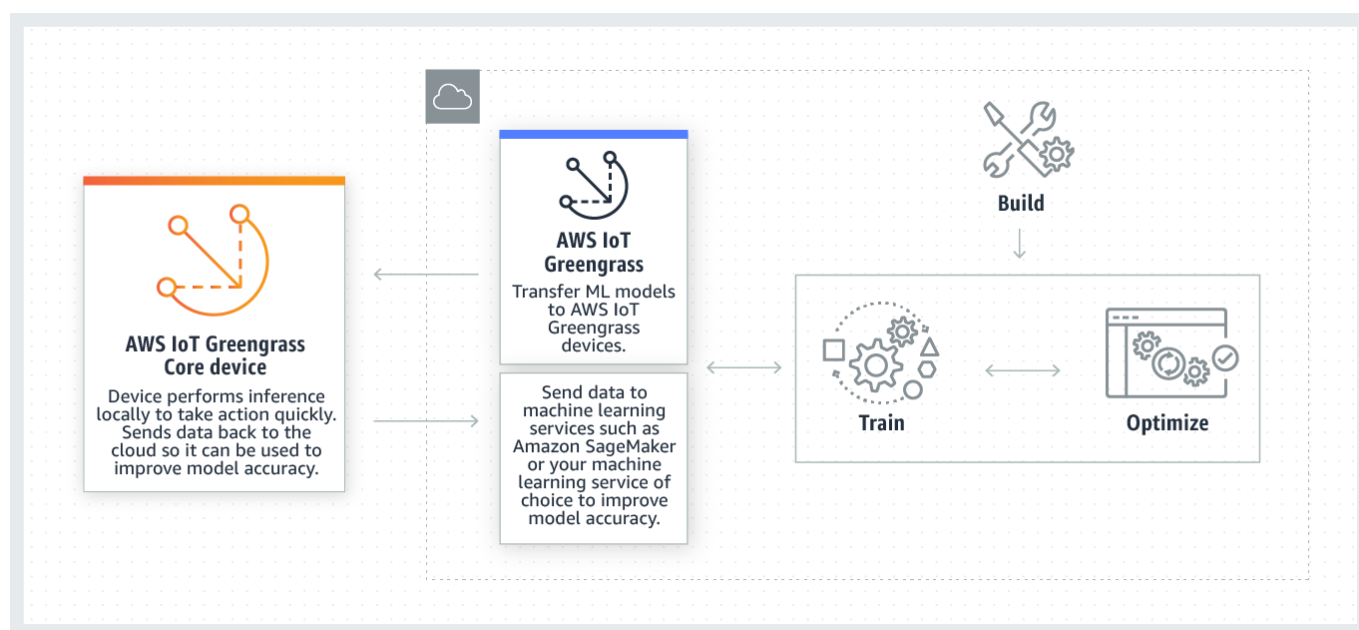
Problem statement

How to have IoT edge but still use benefits of the cloud ?

Edge computing will keeep your data secure, and you don´t necessarily need a cloud IoT platform. Industry 4.0 is all about connecting machines, so your manufacturing processes can react more quickly and intelligently to changing conditions. Connecting assets will help you achieve greater levels of agility and automation, however it will also increase your risks. Thanks to IoT edge you don't need to expose all your data to the cloud to gain the benefits.

Edge computing will also make your big data small. In order to use machine learning to optimize your manufacturing process, make real-time decisions and perform predictive maintenance you need proper data. Unfortunately even in a modest Industry 4.0 project, the amount of data is huge and requires very fast network which, even if exists in your are it brings up the costs for such bandwidth.

Additionally edge computing will allow you to operate with a very low latency. Sending data to a remote cloud datacenter for analysis gives a risk of unpredictable latency. If you need to act quickly based on your algorithm outputs, you may lose the opportunity even before you get the results back.

Thats why fog computing is getting more and more traction and sevices like AWS IoT Greengrass are getting more popular. AWS IoT Greengrass enabled device performs ML inference, and you can run a model locally to make real time decision but also sends data back to the cloud so that it can be used by services like Amazon SageMaker to further improve model acuracy, and store historical data anabling more accurate predictions. Such approach gives you all the benefits of both cloud and edge based solution. The main problem with such hybrid approach is that it relies heavily on stable and fast connectivity from the edge to the cloud.

Source https://aws.amazon.com/

How to leverage scale for learning a model for IoT ?

If you already invested in machine learning in one of your factories you can already see the benefits. Your energy consumption is much lower, you have much less stops in production thanks to predictive maintenance and you a nevert short of parts thanks to AI based werehouse stock management. Regardless if you use AWS SageMaker or your own deployment with MXNet or just a plain TensorFlow in Kubernetes all of them provide mnuch better results if you have more, and higher quality data you use to train your model. However leveraging scale, by providing more data sources is hard. Usually a solution is custom build, and coupled with particular machines you use in the factory. Just changing a vendor of a simple hydraulic press may reduce in different protocol used

to send sensor data, different data model, and different standards followed, which results in a need for a lot of additional development in order to make your solution working again. Even bigger issue is to use data from two production lines for the same product from 2 different factories located in different parts of the globe.
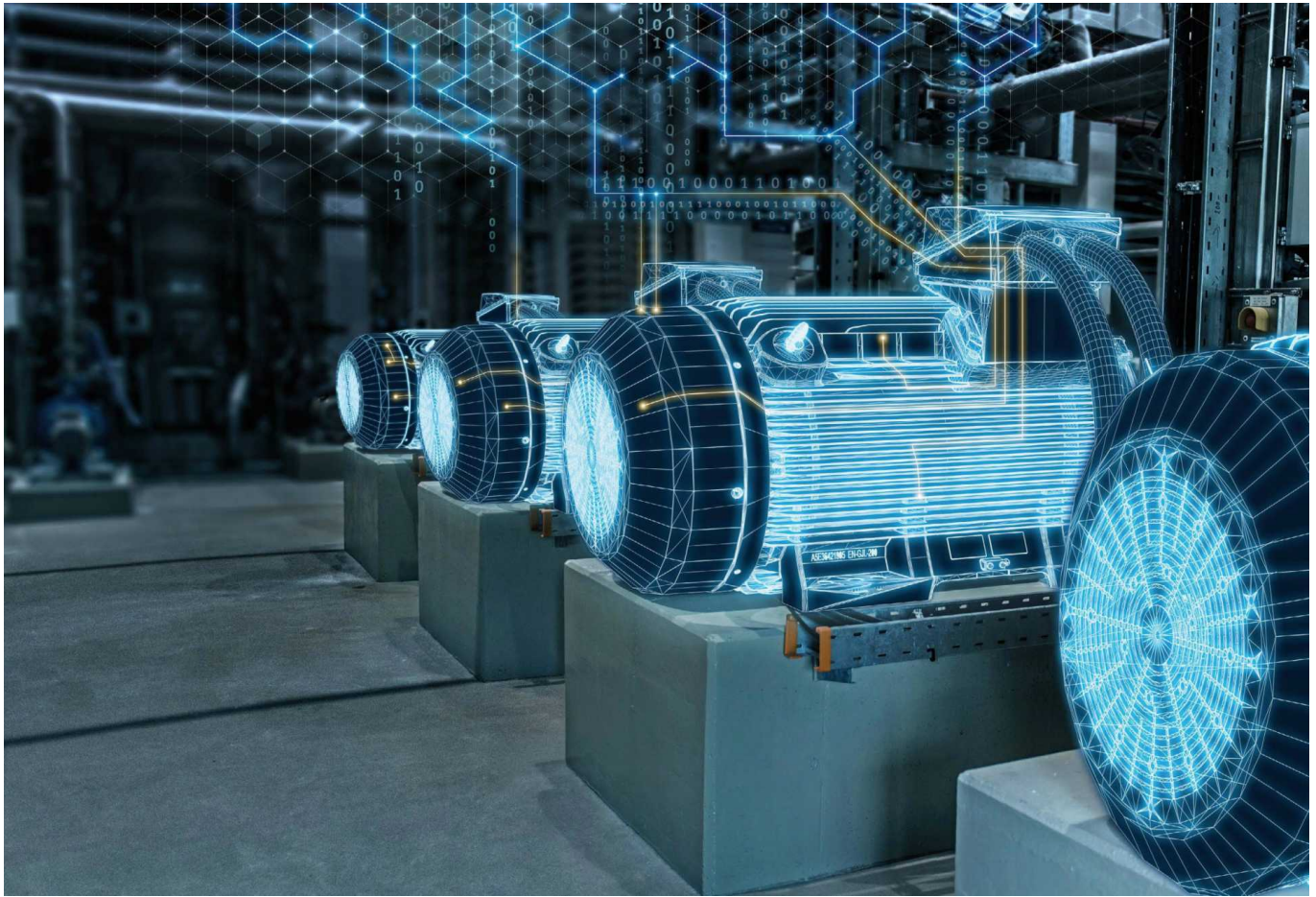

Source https://www.dqindia.com

How can I benefit from digital twins having more than one similar production line ?

In an ideal world a production process of let's say a piece of furniture for the same brand is constant. You have several suppliers having exatly identical factories, with the same machines, the same floor plan, identical environmental conditions etc. Such case would allow you to utilize a digital twin to model the whole production line, with full process, each production nest and a quality control in the same way. Then you can gather all the sensor data and send them via a fast internet connection to the cloud where you use all of them to thain a model and in seconds get insights, recommendations and predictions back you all your factories simultanously.

Unfortunately the real world is different. Even making the same IKEA desk is different accros even 2 plants. They use different tools and machines, they have different floor plan, production volume, quality checks. Moreover one of them is in a middle of nowhere with only a slow mobile connectivity. Also the situation is very dynamic, and if someone has a good idea on how to improve the process it's usually not replicated in other factories due to lack of communication.

Source https://www.themanufacturer.com

MTL - digital twins at scale

MTL Concept

Mesh twin learning is a concept we figured out in PGS Software in response to our customers needs. Three main points behind it are:

- we use digital twins to develop a digital representation of a physical production line environment
- we have IoT Edge deployed, utilizing fog computing principles having small machine learning model being run next to devices or on the factory IoT gateway
- instead of sending a full real time data data stream we export a machine learning model and send the data in batches in some later time or not at all
- we keep a digital twin representation of the whole production line in the cloud, and have the models from different factories compete there
- we are using reinforcement learning in the cloud to train even beter models
- the best models are then pused to production back on the edge
- in order to speed up the advancements process we are using micro optimization strategies by introducing small changes in single production parameters and observing the results, choosing the correct parameters and optimizations step is based on the simulations run in the cloud utilizing a digital twin and prediction models
- we utilize automation and managed services from cloud providers, which means there is no need of building a new IT department to maintain such solution

In MTL we use four different levels of digital twins: single component, asset, system and process.
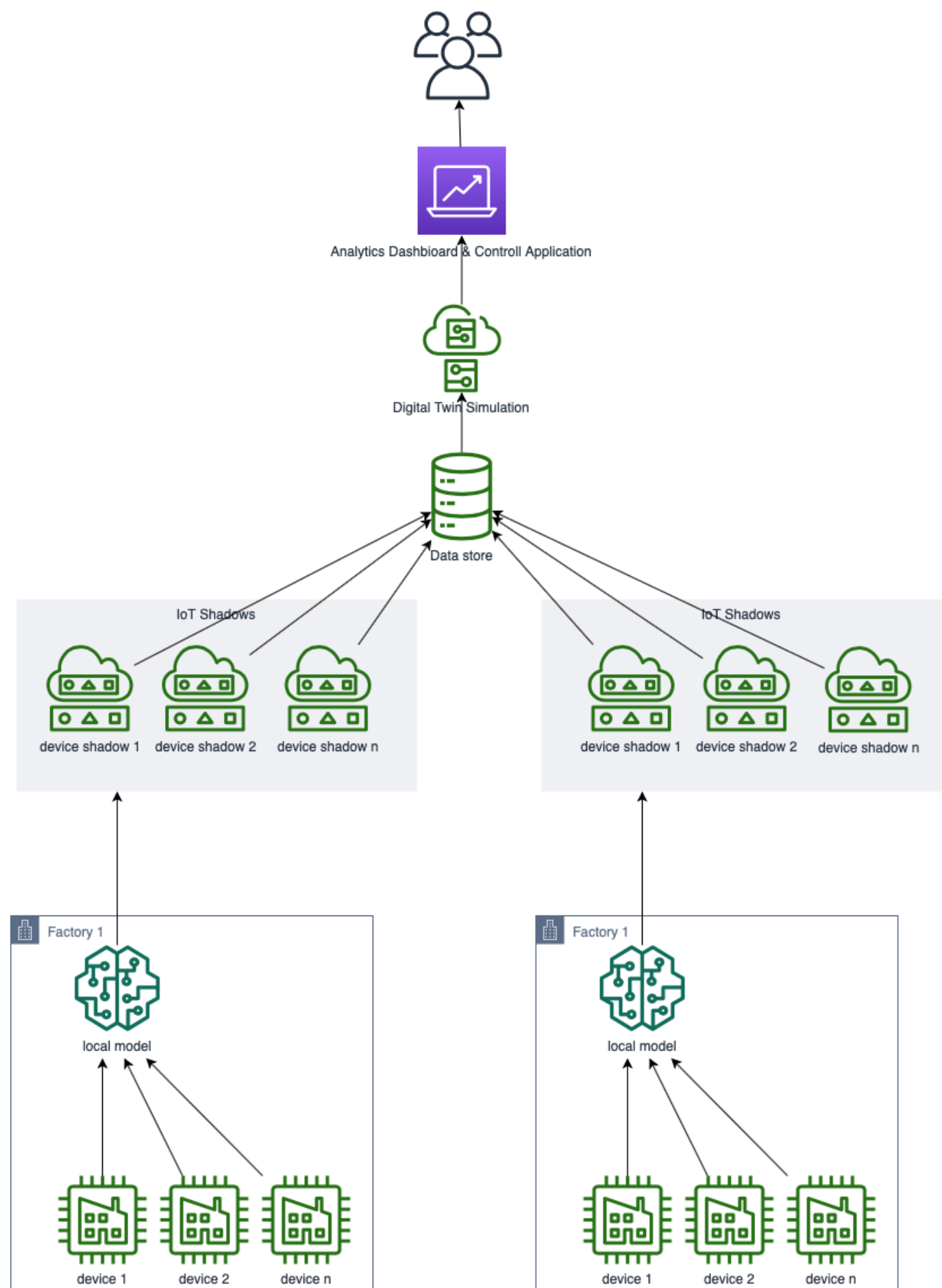
Component This is a digital twin of an individual component within an asset, like a bulb in a scanner or a blade in a turbine. Data at the individual component level allows for data driven decisions for operations and maintenance of that component, the sub-system is sits within and the overall asset.

Asset This digital twin is a model of an entire asset – a piece of manufacturing equipment, a motor or a pump – that gives a view of its workings to optimize performance and enable condition-based maintenance and as more and more data is gathered, predictive maintenance.

System This is a model of a collection of assets that perform a function in a manufacturing setting – such as a production line in a factory. The digital twin delivers data that demonstrates exactly how the ecosystem of assets function together.

Process A process digital twin is the highest level model. It provides a business-level view to measure operational characteristics that underpin business operations across the enterprise. It gives end-to-end visibility to optimize throughput, quality and performance of the process. It enables organizations to visualize and simulate alternative approaches to re-engineer entire processes.

The generic concept looks like this:

In one of next chapters in this white paper we show an implementation example with AWS, as it provides a lot of ready made and managed services making such implementation possible as well as a PoC that we developed internally utilizing AWS Deep Racer cars.

MTL Business benefits

The main benefits of using Mesh Twin Learning are:

- if you have 2 or more simmilar production lines you can get a benefit of scale in production optimizations introduced by machine learning
- you get a continous improvement of your production process being done automatically thanks to micro optimizations, also you speed it up with every addtional production plant connected to the system
- you get a fast feedback cycle and all benefits of a cloud infrastructure without increasing your data flow

MTL Challenges - Why no-one is doing it yet ?

Lack of IoT standarization

- Open standards are key enablers for the success of wireless communication technologies (like RFID or GSM), and, in general, for any kind of Machine-to-Machine communication (M2M)
- Without globally recognized standards,the expansion of RFID and M2M solutions to the Internet of Things cannot reach a global scale
- Clarification on the requirements for a unique global identification, naming and resolver is needed. Lack of convergence of the definition of common reference models, reference architecture for the future networks

This however can be addressed by our approach of basing on the model exports to share learnings between sites instead of raw data.

Complexity

- building MTL solution requires very good knowledge about latest cloud services, combined with experience in machine learning, complex visualizations and industry knowledge
- there are not many experienced vendors providing necesary competences and out of the box solutions will not exist any time soon due to too scattered market of IoT suppliers
- making a project including couple of components or a single asset digital twin is something you can build as an MVP in couple of weeks but a full system or process digital twin, having all equipment conected and all relevant ML components implemented can take 2-3 years
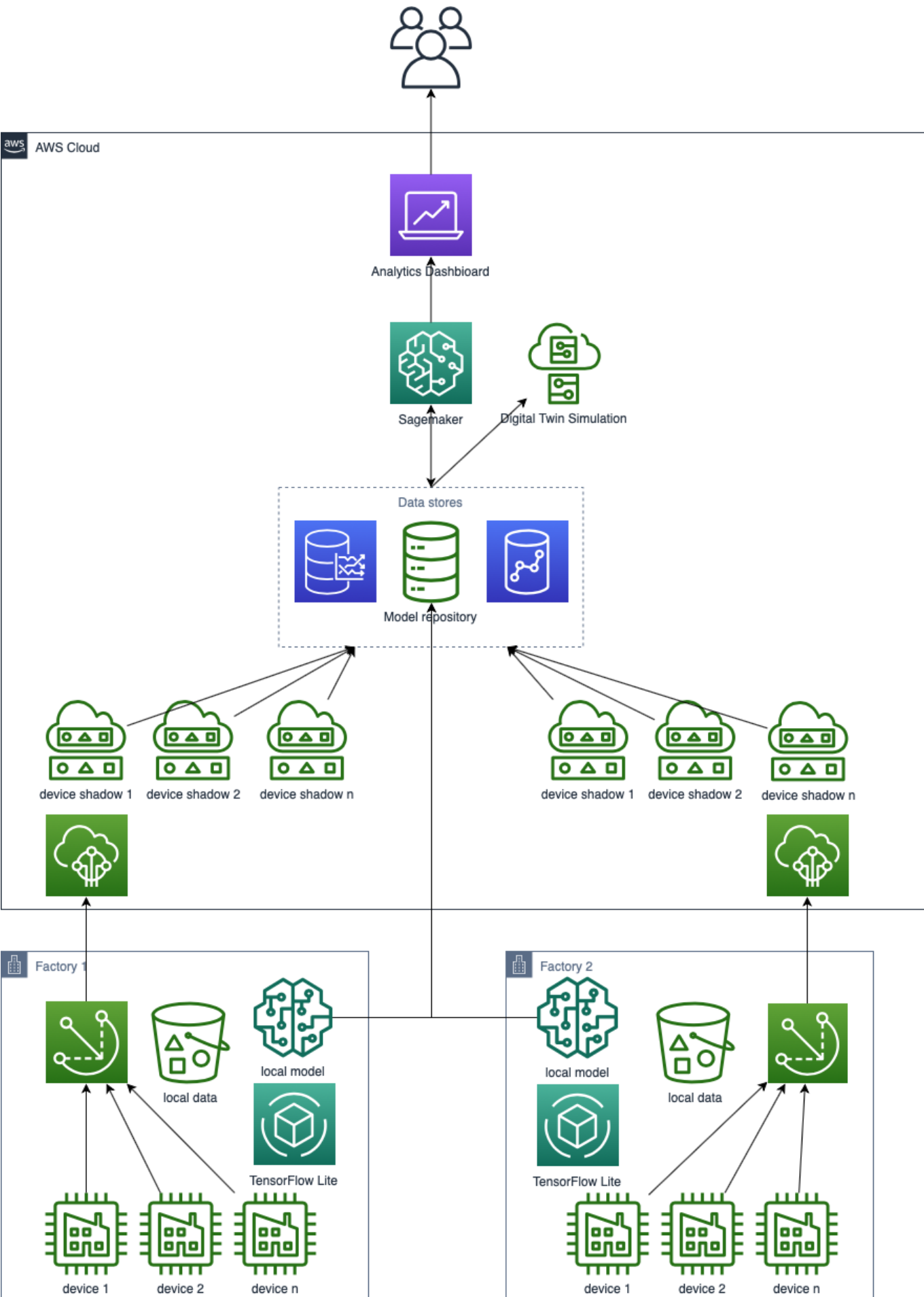
Thanks to using cloud services we can avoid a lot of this complexity, as many different components of MTL are offered by major vendors as managed sercices.

It's brand new

- MTL is a new concept, as it's enabled by the latest cloud services, and advancements in Machine Learning that were released in the begining of 2019
- implementation of this concept can give manufactureres a huge edge over their competition by being first using it in specyfic industry and all the work we saw is surrounded by a veil of secrecy
- there are 0 advertised commercial implementation, technology behind it is not yet established and there is no possibility to "copy & paste" a solution from someone else

Solution example

Ok, we talked a lot about the concept and it's benefits. Now let's see how it can be realized. AWS based implementation of an MTL concept looks like this:

The key component that make this solution possible are:

AWS SageMaker

This is a heart of MTL implementation on AWS. SageMaker is a service that provides every developer and data scientist with the ability to build, train, and deploy machine learning models quickly. Amazon SageMaker is a fully-managed service that covers the entire machine learning workflow to label and prepare your data, choose an algorithm, train the model, tune and optimize it for deployment, make predictions, and take action. In MTL concept it's the place where we are running models being important from different edges, evaluate their results and also use it to figure out micto step optimization plan.

AWS SageMaker Neo

SageMaker Neo enables developers to train machine learning models once and run them anywhere in the cloud and at the edge. Amazon SageMaker Neo optimizes models to run up to twice as fast, with less than a tenth of the memory footprint, with no loss in accuracy. This is another crucial component for MTL, making it possible to use a model trained in the powerfull SageMaker instance and re-compile it for anb exact device running on the target edge and makes it execution efficient. You start with a machine learning model built using MXNet, TensorFlow, PyTorch, or XGBoost and trained using Amazon SageMaker. Then you choose your target hardware platform from Intel, NVIDIA, or ARM. With a single click, SageMaker Neo will then compile the trained model into an executable.

AWS GreenGrass ML Inference

AWS IoT Greengrass makes it easy to perform machine learning inference locally on devices, using models that are created, trained, and optimized in the cloud. IoT Greengrass gives you the flexibility to use machine learning models in Amazon SageMaker or to bring your own pre-trained model stored in Amazon S3. This allows us to make MTL concept fully automated in terms of model management and deployment.

Digital Twin Simulation

Based on IoT devices shadows that we have in our repository we are able to build a full digital twin of an asset, device, system or even a full process. This allows us to:

- have an abstraction layer hiding a complexity of IoT devices comming from different vendors
- have a safe simulation environment where we can evaluate different models
- build a full near real time visualization if needed

Solution example results with Amazon Deep Racer

As we wanted to have also a public showcase of this concept we implementd above solution in house wiht AWS Deep Racer. This is a simple component level digital twin with just a few parameters nevertheless it nicely shows the benefits of MTL.

We managed to have 2 AWS Deep Racer cars utilizing each others experience and sharing what they learned with MTL.

In PGS we have a lot of experience with AWS Deep Racer and we are getting very good results in AWS competition.

> Add some info about Tomasz Panek and his results in AWS DeepRacer competition ...

Summary and Conclusions

As you can see the Industry 4.0 landscape is changing quickly. With new cloud services, and improvment in Machine Learning tooling we can achieve great benefits for all manufacturing companies, and there is no requirement for a big bang approach and a huge investment. Mesh Twin Learning is a practical way of solving many problems that factories currently have with introductio of IoT and additional services utilizing this infrastructure for a business benefit.

> add something like if you want MTL implemented for your use case contact PGS ...

Data Exploration Workshop 1 pager

> insert the DEW one-pager here