Journal of Innovation

industrial internet® CONSORTIUM

OpenFog®

10 th Edition

INNOVATIVE TECHNOLOGIES

March 2019

Dear Reader:

Welcome to the Innovative Technologies issue of the Journal of Innovation! The pace of innovation is accelerating, creating disruptions across the board: individuals, businesses, governments and society. Traditional boundaries between these entities are completely redrawn. Digital technologies which began as a differentiating advantage—including the Industrial Internet of Things—are turning into something expected from every business.

This edition of the Journal of Innovation covers a range of Innovative Technologies pertinent to the Industrial Internet of Things, from Custom Application Specific Integrated Circuits (ASICs) to Intelligent Realities:

- Michael D. Thomas describes how Intelligent Realities are used to aid worker cognition and performance in "Intelligent Realities for Workers Using Augmented Reality, Virtual Reality and Beyond."
- In "Keeping Ahead of the Curve with Custom ASICs," Edel Griffith, Darren Hobbs and Sohrab Modi describe how custom ASICs are now economically viable for smaller volumes; as well as their impact on the Industrial Internet of Things, especially in regards to their edge processing.
- With "Improving Reliability and Security of Global Cold Chain Logistics for Pharmaceutical Assets," Dr. Dave Stanton and Dr. Madhusudan Pai explore the current state of environmentally-sensitive supply chain (otherwise known as cold-chain) logistics for pharmaceuticals.
- Dan Isaacs, Jillian Goldberg, Dionis Teshler and Tal Nisan describe the use of Programmable Logic for automotive security in "Automotive Security through New Communication Lockdown Utilizing Programmable Logic Solutions ."

This edition also includes an article on the IIC Smart Factory Web Testbed based on an interview by Joseph Fontaine with Dr. Kym Watson. The March edition concludes with important updates on IIC activities in "What's New at the IIC."

We sincerely hope you enjoy this edition of the Journal of Innovation. Please let us know if there are any topics you would like to see covered in the future. The next edition will focus on Artificial Intelligence. Stay tuned!

Best Regards,

Edy Liongosari
Chief Research Scientist
Accenture Labs
Thought Leadership Task Group Co-Chair
Industrial Internet Consortium

Mark Crawford
Standards Strategist
SAP
Thought Leadership Task Group Co-Chair
Industrial Internet Consortium

Dear Reader:

The Industrial Internet Consortium, now incorporating OpenFog, is pleased to publish this tenth edition of the *IIC Journal of Innovation*.

This collaborative effort is the sum of very many parts and we would like to take this opportunity to recognize and thank the team of editors and peer reviewers who lent their time and expertise to the editorial process of enhancing each of the articles contained in this edition with their unique perspectives and wisdom.

Sincere thanks goes to this edition's editors and peer reviewers:

Mr. Mark Crawford, Director Standards Strategy, SAP Strategic IP Initiatives
Mr. Edy Liongosari, Chief Research Scientist, Accenture Labs
Ms. Enas Ashraf, Technology and Business Solutions Manager, Advancys ESC
Dr. Vincent Bemmel, Industry Technology Lead, Corlina
Mr. Bassam Zarkout, Executive Vice President, IGnPower
Mr. Sudhanshu Mittal, Director, Industry 4.0, NASSCOM
Mr. Abhik Chaudhuri, Domain Consultant - Design & Architecture CoE, Tata Consultancy Services
Mr. Jijun MA, Director of Industrial Internet, Wanxiang Group
Ms. Zhao Huidan, CEO, Jiangsu Sino Logistics Networking Technology Co., Ltd.
Mr. Gavin Green, VP Strategic Solutions, XMPro
Ms. Cheryl Rocheleau, Sr. Marketing Manager, Industrial Internet Consortium
Mr. Andrew Hazerjian, Marketing Specialist, Industrial Internet Consortium

As a global, not-for-profit, public-private partnership of over 200 member organizations, we encourage every organization across all industries to get involved and be proactive in shaping the Industrial Internet. We welcome your feedback and participation.

Many thanks,

Kathy Walsh
Vice President of Marketing
Industrial Internet Consortium

Contents

Intelligent Realities For Workers Using Augmented Reality, Virtual Reality and Beyond

Authors:
Michael D. Thomas
Senior Systems Architect
SAS
Michael.Thomas@sas.com

INTRODUCTION

Through all the industrial revolutions, tools and machines have been central to workers' realities. But it is only recently that large portions of a worker's reality could be digitized with IoT devices and approaches. In 2015, Henning Kagermann, former CEO of SAP AG, argued that this "digitization—the continuing convergence of the real and the virtual worlds will be the main driver of innovation and change in all sectors of our economy." [1] This simple act of creating digital streams produces information that can be expressed in many different ways, on many different types of materials, and in many different systems.[2] This article argues that modern reality presentation technologies are compelling mediums for the expression of digital IoT streams.

Such reality presentation technologies include the eXtended Reality (XR) family of technologies [3] -- Augmented Reality (AR), Mixed Reality (MR) and Virtual Reality (VR) – as well as more mature and accepted technologies such as smart phones, tablets, and PC flat screens. When combined with IoT, analytics and artificial intelligence, applications can be created that can aid workers by making their realities more intelligent.

An intelligent reality is defined here as a technologically enhanced reality that aids human cognitive performance and judgement. As compared to the base reality, an intelligent reality can have much greater dimensionality, reduced occlusion, transcendence of distance, better guidance and improved communication with other actors. This definition deliberately does not exclude non-physical realities in domains such as finance and cybersecurity, but the focus of this article is on intelligent realities based on physical realities and fed by IoT.

Consider a technician looking at a machine while wearing an AR Head Mounted Display (HMD) can see both the service history and prediction of future failures. This gives the worker a view on the fourth dimension of time, both backwards and forwards. Instead of having to take the machine apart, the worker can see an IoT driven mixed reality rendering projected on the outside casing. By just glancing away from the machine, he can see a virtual rendering of the operations of the same type of machine at a distant location. Then, he can interface with both artificial and human remote experts about next steps, which could include the expert

[1]H Kagermann, "Change Through Digitization-Value Creation in the Age of Industry 4.0," Management of Permanent Change, p 23, 2015. Available: https://www.researchgate.net/publication/284761944_Change_Through_Digitization-Value_Creation_in_the_Age_of_Industry_40

[2] J. Brennen, D. Kreiss, "Digitization," The International Encyclopedia of Communication Theory and Philosophy, p 557, 2016 .

[3] C. Fink, "War of AR/VR?MR/XR Words," Forbes, Oct 2017. Availablehttps://www.forbes.com/sites/charliefink/2017/10/20/war-of-arvrmrxr-words/#2d75b71e8d07

driving virtual overlays of his view. As he decides on next steps, he can communicate with appropriate management systems through that same HMD without having to pull out a phone or laptop. As a wearable computer, the HMD brings distant resources in to the worker's operational reality.

An intelligent reality may be proximate to a worker, like a machine on a factory floor. Or that factory might be half way around the world and understood by the user through 3D modeling of the factory. AR or VR headsets may be involved, but do not have to be – smart phone screens or flat screens on a desktop may be a better option. The worker may be mobile and use an AR head mounted display or smart phone, or the worker may be stationed in a command center at the company headquarters. They may be observing a reality in real time, or they may be performing data-driven review of an event that occurred in the past. In all cases, though, the context dominates – both visually and in the design of the presentation.

Intelligent reality can be achieved today with off-the-shelf technologies spanning IoT, analytics, XR technologies, and more traditional user interface technologies. This new paradigm is introduced here to help decision makers and architects navigate the expansive terrain of technologies that can enable intelligent realities for workers. First, the XR space is overviewed along with more traditional mobile and desktop flat screens. This leads to the consideration of intelligent reality architecture and the development of intelligent reality applications. From there, specific use cases are proposed that exercise combinations of reality presentation technologies, IoT and AI.

The Reality-Virtuality Continuum and Modern eXtended Reality

In 1995, Paul Milgram et al published a paper "A Taxonomy of Mixed Reality Visual Displays", which introduced the Reality-Virtuality Continuum.[4] This paper remains useful for discussing the current state of XR as well as considering the role of mobile and stationary flat screens. Figure 1 illustrates the continuum between purely physical reality and purely virtual.

[4] P. Milgram et al., "Augmented Reality: A class of displays on the reality-virtuality continuum," Proceedings Volume 2351, Telemanipulator and Telepresence Technologies, Dec 1995. Available: http://etclab.mie.utoronto.ca/publication/1994/Milgram_Takemura_SPIE1994.pdf

Figure 1: Reality-Virtuality Continuum (From Wikipedia)

From left to right, the user moves from a normal view of physical surroundings to a completely digital view. In between the extremes is Mixed Reality (MR) -- the mixing of the physical reality with one or more digital realities. MR assumes an AR device that is capable of stereoscopic rendering of dynamic 3D scenes on top of a physical view of the world.

On the far right, a virtual environment is completely digital, but not necessarily completely immersive. The authors include both the completely immersive experience of a VR Head Mounted Display (HMD) as well as large flat screens not worn by the user. Both VR HMDs and virtual environments rendered on flat screens can provide a user with a dynamic, real-time 3D rendering of a remote or abstract 3D reality.

Just as the authors did not limit virtual environment presentation to HMDs, their definition of AR does not exclude mobile flat screens. In 1995, they lacked the terms "smart phone" and "tablet," but they described "monitor based (non-immersive) video displays – i.e. 'window-on-the-world' (WoW) displays – upon which computer generated images are electronically or digitally overlaid."

The Modern Reality-Virtuality Landscape

Figure 2 illustrates the Reality-Virtuality Continuum with commercially available products. The lower quadrants are traditional flat screens while the upper quadrants contain the newer and less established HMDs.

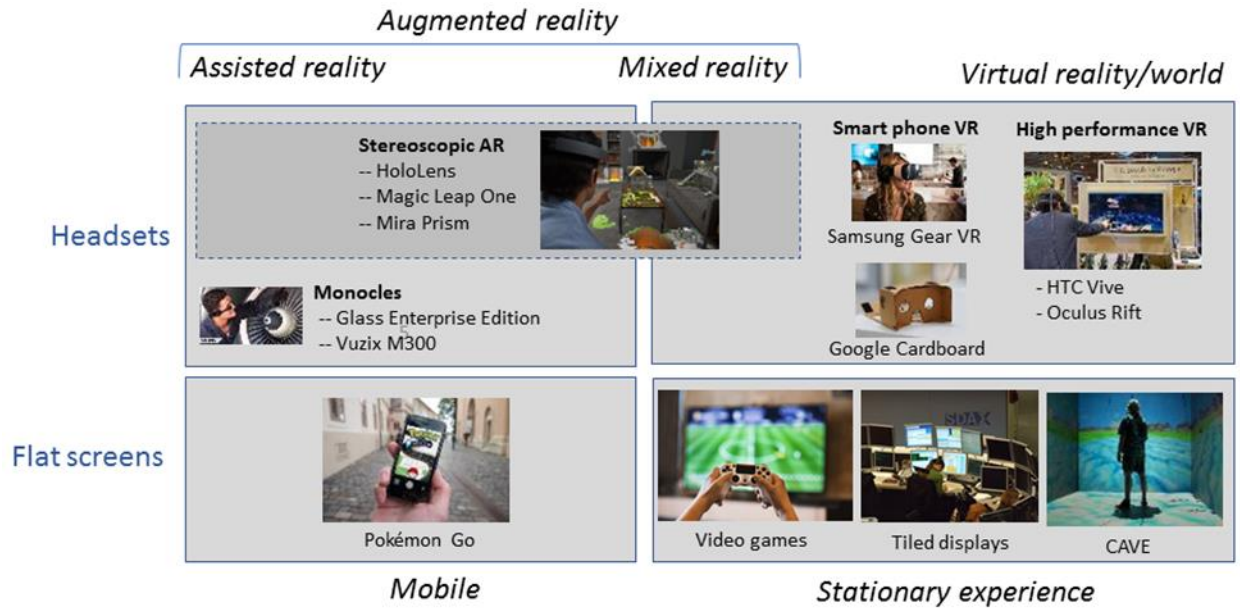


Figure 2: Devices on the Reality-Virtuality Continuum[5]

The quadrants represent different use cases and approaches:

- Lower left: smart phone and tablet AR. Smart phones and tablets are so pervasive that the incremental hardware cost is often zero. But they are not heads-up and hands-free.
- Lower right: flat screen virtual worlds (or VR for the flat screen). This includes virtual worlds on flat screens, tiled displays and Computer Assisted Virtual Environments (CAVEs).
- Upper right: virtual reality for HMDs. The market supports different VR devices with different resolution, field-of-view, and processing power. Tethered HMDs connected to powerful PCs, such as Oculus™, HTC™ and PiMax™ products, are the most capable. Less capable but also less expensive and sometimes more convenient are smart phone approaches such as Google® Cardboard™ and Samsung® Gear VR™.
- Upper left: AR HMDs. With AR, the design fragmentation is the greatest. There are three basic design categories:
  - Stereoscopic headsets. Larger and compatible with prescription glasses. Microsoft® HoloLens™ is a headset. Mira Prism™ is a headset that utilizes a smartphone.
  - Stereoscopic smart glasses. Smaller but users may need to procure prescription lenses for

[5] Attribution for embedded images, starting clockwise from upper left: Vuzix, Microsoft Sweden, Nan Palmero, Google, Jean-Pierre Dalbéra, Dave Pape, Wikimedia user Dontworry, EVG photos, pixabay.com

the device. Magic Leap One exhibits the smart glasses form factor.

- Monocle devices. A small screen in front of one eye. They tend to be compatible with prescription glasses. These devices focus on "assisted reality" – the display of flat content such as charts, videos, and text to the side of a person's view.[6]

As illustrated with the dotted box around stereoscopic AR, these devices can handle assisted reality use cases and can satisfy some use cases in the VR space that don't require full immersion. One example of the latter is examining a virtual 3D model of a building at arm's length.

Efficacy of XR in Commercial Settings

Most any new technology both generates hype and begs questions of its usefulness. In the case of XR, there have been several encouraging studies about their efficacy:

- In its November-December 2017 issue, Harvard Business Review published an article "Why Every Organization Needs an Augmented Reality Strategy."[7] The article also discussed commercial use of VR. In their research, the authors found various positive outcomes, including:
  - DHL® and Intel© saw AR related warehouse picking productivity gains of 25% and 29% respectively, with Intel seeing error rates falling to near zero.
  - Newport News Shipbuilding® used AR and reduced inspection time by 96% because the final design could be superimposed on a ship.
  - Lee Company®, which sells and services building systems, calculated a return of $20 on every dollar it has invested in AR.
- An AR experiment by Dr. Steven J. Henderson and Dr. Steven Feiner of Columbia University in 2009 found that a "prototype AR application allowed a population of mechanics to locate individual tasks in a maintenance sequence more quickly than when using an improved version of currently employed methods."[8]
- In the paper "Virtual Reality and Augmented Reality as a Training Tool

[6] S. Crucius, "What Is Assisted Reality and How Does It Relate to Augmented Reality?" Wearable Technologies, June 2018. [online] Available: https://www.wearable-technologies.com/2018/06/what-is-assisted-reality-and-how-does-it-relate-to-augmented-reality/

[7] M. Porter and J. Heppelmann, "Why Every Organization Needs an Augmented Reality Strategy," Harvard Business Review, Nov-Dec 2017. Available: https://hbr.org/2017/11/a-managers-guide-to-augmented-reality

[8] S. Henderson and S. Feiner, "Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret," 2009 8th IEEE International Symposium on Mixed and Augmented Reality, Oct 2014. Available: https://ieeexplore.ieee.org/document/5336486

for Assembly Tasks" from The School of Manufacturing and Mechanical Engineering at The University of Birmingham, the authors investigated if AR and VR offered potential for training of manual skills.[9] They compared AR and VR training methods to the use of conventional 2D engineering drawings and found that AR and VR approaches resulted in significantly reduced task completion times.

- In the 2015 paper "Augmented Reality as a Tool for Production and Quality Monitoring," the authors tested use of an AR system rendering information from Computer Aided Quality (CAQ) software and compared it to scenarios using only CAQ software and using no software.[10] AR integrated with CAQ was found to be the fastest approach.

INTELLIGENT REALITY ARCHITECTURE

An architecture for an intelligent reality should be centered on aiding a worker's cognition and performance. For workers in an IoT-enabled reality, the cornerstone of an intelligent reality architecture is the integration and sense-making of raw IoT data. This is discussed first in this section. With the data and analytical foundations in place, an architectural view of the reality presentation technologies is presented for making the best tactical last-mile UI decisions for rendering to the workers.

IoT Data Pipeline

For real time sense making of an IoT asset, a streaming analytics engine is necessary.[11] A streaming analytics engine, like SAS® Event Stream Processing, analyzes data streams in motion as the atomic events of the stream pass by. In addition to applying analytical methods, it can also provide inferences derived from machine learning models as well as contribute to the training of such models.

In addition to immediate presentation, analyzed data streams can also be transferred to data stores for further analysis and later presentation. While the big data problems related to IoT described by Belli et al in "A Scalable Big Stream Cloud Architecture for the Internet of Things" need

[9] A. Boud et al., "Virtual Reality and Augmented Reality as a Training Tool for Assembly Tasks," 1999 IEEE International Conference on Information Visualization, Jul 1999. Available: https://pdfs.semanticscholar.org/a563/afc2156eb7285dc67c1c5be7dd3787f0db04.pdf

[10] D. Segovia et al., "Augmented Reality as a Tool for Production and Quality Monitoring," Procedia Computer Science 75:291-300, Dec 2015. Available: https://core.ac.uk/download/pdf/81959814.pdf

[11] B. Klenz, "How to Use Streaming Analytics to Create a Real-Time Digital Twin," SAS Global Forum 2018, Mar 2018. Available: https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2018/2004-2018.pdf

to be addressed[12], they are not unique to reality applications. Reality applications can be networked and fit in with a traditional query-and-reporting client-server architecture. As observed in "Immersive Analytics: Building Virtual Data Worlds for Collaborative Decision Support", traditional 2D data visualization can work across the reality-virtuality continuum.[13]

General Model of Device and Reality Interaction

On the UI front, intelligent reality has been enabled by wearable and portable computers, including XR HMDs and smart phones, and high-performance graphics that can faithfully render realities. While HMDs represent an important shift in computing, they are still wearables that may not be acceptable to many users and use cases. The following architectural view attempts to de-emphasize the importance of HMDs in reality intelligence architecture.

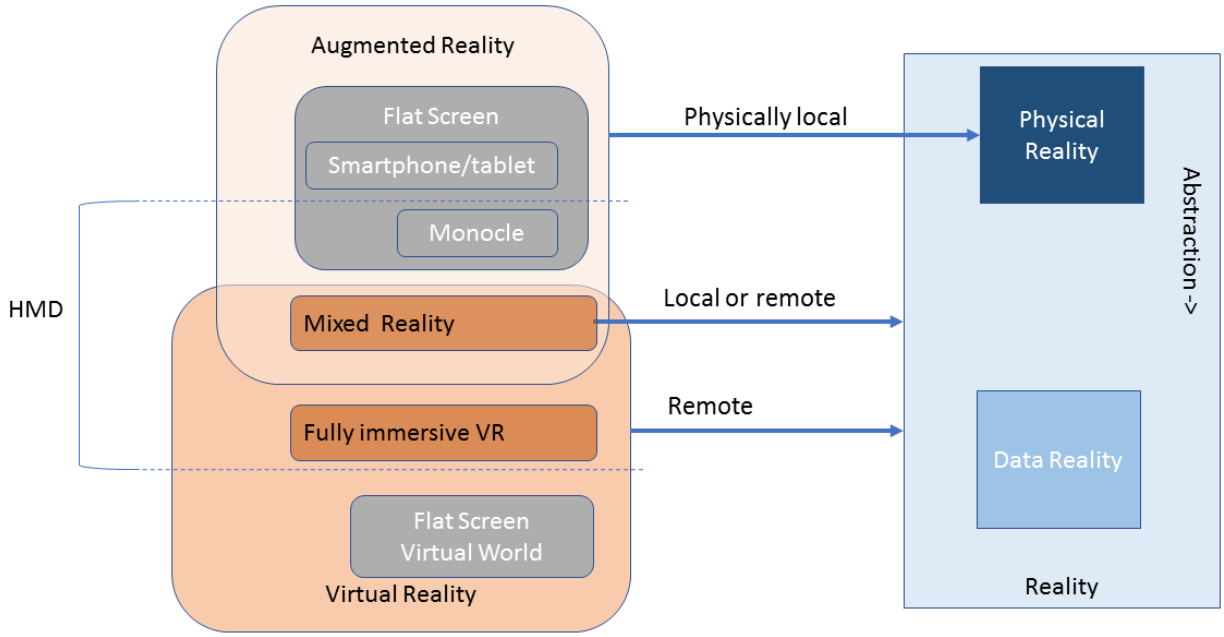


Figure 3: Device and reality interaction view

[12] L Belli et al., "A Scalable Big Stream Cloud Architecture for the Internet of Things," International Journal of Systems and Service-Oriented Engineering, 5(4), 26-53, October-December 2015. http://www.tlc.unipr.it/ferrari/Publications/Journals/BeCiDaFeMeMoPi_IJSSOE15.pdf

[13] R. Hackathorn and T. Margolis, "Immersive Analytics: Building Virtual Data Worlds for Collaborative Decision Support," IEEE VR2016 Workshop, March 2016. Available: http://www.immersiveanalytics.com/wp-content/uploads/2016/05/VR2016-IA-HackathornMargolis-20160322.pdf

Starting on the right, in Figure 3, is the general concept of reality drawn to include both physical and abstract realities. A machine is a physical reality, while the supply chain that created that machine is an abstract reality derived from data – a data reality. At the most abstract, data realities can be completely de-coupled from any physical reality. For example, large volumes of live streaming data from a commodities market could be used to form a data reality which a user could explore in VR.

The left side combines the three main concepts of augmented reality, mixed reality and virtual reality. Due to the see-through nature of AR devices, proximate physical reality is always part of an AR experience. But an implementor who may choose to use a mixed reality device to satisfy a use case, has nothing to do with the proximate physical reality – for example, rendering a 3D model of a supply chain independently. They may make this choice because users prefer mixed reality over virtual reality because being aware of physical surroundings is more comfortable. Thus, mixed reality could be effectively remote or local.

VR tends to mean a fully immersive experience with an HMD. But a VR asset could be rendered in MR or on a flat screen. This article has not taken on the trouble of constantly restating that a virtualization of reality can be rendered on an AR stereoscopic HMD, a flat screen, or a fully immersive VR HMD. Unless specifically qualified, the term VR takes the meaning of a virtualization of a reality and does not assume the target device.

The Base Software Layers Across the Reality-Virtuality Continuum

Reality applications of any type, including games as well as industrial applications, rest atop lower level software layers that have emerged to solve the different problems described here:

Gaming Engines for Dynamic and Interactive 3D Models

Many developers use gaming engines like Unity and Unreal to develop these models as well as output the executable application.[14] These development tools can output applications across many platforms, including AR and VR HMDs, smart phones, tablets and PCs, which can communicate over networks with servers and other applications. Amazon Sumerian is a web-based alternative that gives developers an easier but more limited alternative to gaming engines. [15]

The gaming engines can import 3D models from other sources, including tools such as Autodesk® Maya™ that are focused on original 3D content creation by artists as well

[14] W. Wise, "How to pick the right authoring tools for VR and AR," O'Reilly On Our Radar, October 2017. Available: https://www.oreilly.com/ideas/how-to-pick-the-right-authoring-tools-for-vr-and-ar

[15] R. Marvin, "Inside Sumerian, Amazon's Big Bet on Augmented and Virtual Reality," PC Magazine, April 2018. Available: https://www.pcmag.com/feature/360323/inside-sumerian-amazon-s-big-bet-on-augmented-and-virtual-re

as tools that can import existing CAD drawings. In addition, AI can create 3D content. Booz Allen® has demonstrated the use of generative adversarial networks to greatly reduce the time and expense of content creation for simulation.[16]

World Knowledge with Computer Vision

In Milgram's paper cited earlier, the authors identified the dimension of Extent of World Knowledge, with the end points of an unmodelled world and a completely modelled world. For AR applications, an application needs to at least partially model the world, which the authors divided in to knowing what an object in the view is and where it is in the view.

For enterprise reality applications, the identity of an object matters. For example, a fleet manager standing in front of a bus needs to know exactly what bus it is, not just that it is a bus. Object identity can be achieved by reading bar codes, QR codes, text or other uniquely identifying marks.

Once an asset is identified, natural feature tracking and object recognition can be employed to recognize component parts. For example, the fleet manager can first glance at the license plate of the bus and then see the correct data-driven overlays for that bus as they move around because computer vision is identifying parts of the bus, such as the tires, engine, etc.

Mixing realities

An additional type of SDK is focused on the mixing of digital and physical realities. To properly place a digital reality in to physical reality, the surfaces in the physical world and lighting need to be understood. ARCore™ from Apple® and ARKit™ from Google solve these problems for iOS and Android, respectively.[17]

Intelligent Reality and AI

The various technologies under the AI umbrella are quickly becoming just additional tools in the developer toolbox. This is certainly true in the XR space. For example, when the original Microsoft HoloLens released in 2016, it came with both voice recognition and computer vision available to application developers. [18]

But there is a lot more room for AI in the intelligent reality space than these baked-in features. The following sections look at several areas where AI can provide value to workers' perceptions of their realities.

Overcoming XR UI Limitations

Just as the transition from desktop to smart phone apps required new approaches, reality apps introduce their own UI

[16] N. Mehta et al., "The Role of AI in a VR World," GPU Technology Conference, October 2018. Available: http://on-demand.gputechconf.com/gtcdc/2018/pdf/dc8209-the-role-of-ai-in-a-vr-world-presented%20by-booz-allen-hamilton.pdf

[17] M. Giles, "AR still doesn't have a killer app, but Google's ARCore is here to help," MIT Technology Review, Feb 2018. Available: https://www.technologyreview.com/s/610336/ar-still-doesnt-have-a-killer-app-but-googles-arcore-is-here-to-help/

[18] D. Bohn and T. Warren, "Up close with the HoloLens, Microsoft's most intriguing product in years," The Verge, January, 2015 [online]. Available: https://www.theverge.com/2015/1/21/7868251/microsoft-hololens-hologram-hands-on-experience

constraints. Both VR and AR reality analytics apps must deal with the basic problem of putting context first. If users are going to gain value from having their analytics in context, then the analytics cannot overly obscure the context. In VR, that means that a 3D model of a factory should be visually dominant if it is to properly contextualize a chart about some aspect of the factory's operations.

As UI space is at a premium, it becomes important to use that space wisely. The challenge is to give the user the best information for their role at that point of time and for their current location. AI can help solve that problem. Rather than forcing the worker in to a data exploration UI paradigm which would require many selection actions, AI can make content selections on behalf of the worker.

Artificial Remote Experts

In the popular remote expert use case for AR[19], the remote expert could be human or artificial. For example, a field technician wears an AR HMD and a human remote expert can see what the technician is seeing through the head-mounted camera. The remote expert could also access equipment history and metrics.

An artificial expert could also carry this burden or work in concert with a human expert. The AI chatbot practices[20] seen at call centers can be brought to bear. Just as chatbots replace first level call center representatives, they can alleviate remote experts from first level work. Then, a single remote expert can cover more junior workers and focus on tougher problems.

Digital Twin Overlay

The Industrial Internet Consortium defines a digital twin as "a digital representation of an entity, including attributes and behaviors, sufficient to meet the requirements of a set of use cases." [21] It is not only data about a physical asset, like its service history. A good digital twin takes the information about the design, production and operational life of the asset and virtualizes it in to a digital asset that can be tested and modified in ways that you would never treat an operating physical asset. Instead of a single expensive crash test of a car, you could perform millions of crashes virtually. Rather than a couple of turns around a test track, a car could be virtually driven for millions of miles across multiple tests with different service histories. Such tests could then be used to feed machine learning neural nets which are then queried when servicing the real asset.

[19] E. Hadar et al., "Hybrid remote expert - an emerging pattern of industrial remote support," CAiSE Forum, 29th International Conference on Advanced Information System Engineering, Essen, Germany, June 2017. Available: http://ceur-ws.org/Vol-1848/CAiSE2017_Forum_Paper5.pdf

[20] K. Nimavat and T. Champaneria, "Chatbots: An Overview. Types Architecture, Tools and Future Possibilities," International Journal for Scientific Research & Development, October 2017 https://www.researchgate.net/publication/320307269_Chatbots_An_overview_Types_Architecture_Tools_and_Future_Possibilities

[21] Q1 Digital Twin Interoperability TG Meeting Minutes, Feb 2019, Available: https://workspace.iiconsortium.org/higherlogic/ws/groups/interop-tg/download/25418/latest

With an intelligent reality application, the digital twin can be overlaid directly on the physical twin. When a bus rolls in to the garage, a fleet manager can view important output from the bus's digital twin as an AR overlay. A simple example is showing an alert because the bus is overdue an oil change. But the real power of the digital twin would come from more nuanced cases that aren't simple violations of established single-dimensional rules. Perhaps the bus is within the accepted ranges across several aspects of maintenance, but the digital twin sees that a combination of near violations greatly increases the risk of a mission-critical failure. The AR device can communicate to the manager who can then take appropriate action.

Feeding AI from XR Devices

When a factory deploys a thousand AR HMDs to workers, they are also deploying at least a thousand head mounted cameras. Those cameras are well-positioned to provide a rich set of video content. Such content can be piped through computer vision and then on to machine learning and other analytical models. In addition to video from the cameras, HMDs can transmit precise information about the position and orientation of the head of the wearer.

For manufacturing, an AR-enabled workplace could generate machine learning models that are trained based on head position, gaze, placement of components in the workspace, and quality outcomes. Once trained, such a model could detect small movements and practices that lead to poor quality outcomes and suggest better practices immediately through the HMD. Such learnings can be deployed back in to workers' intelligent realities.

While VR doesn't offer the same connection to the physical world, a VR HMD can also communicate position and orientation of the workers' heads. Eye tracking is also making it in to XR products, including HTC Vive Pro Eye™ and Microsoft HoloLens2. Such information can be used to improve the simulation as well as strengthen the understanding of how humans would react in the physical analogue of the virtual environment.

EXAMPLE USE CASES FOR INTELLIGENT REALITY

This section considers three use cases for reality analytics along with architectural solutions to their problems.

Augmented Reality Chess Coach

In this first use case, the work is developing Science, Technology, Engineering, and Mathematics (STEM) skills in young children, and the workers are parent volunteers that want to share the STEM benefits of chess with elementary schools.

Chess is known to aid the development of STEM skills for students as young as elementary school and elementary school chess clubs can provide a venue for youth

chess play.[22] It should be possible to use computer vision to interpret and analyze a chess position on a physical board. This capability could be packaged in an app that aids the parent volunteer in both ensuring legal chess play and providing chess coaching and knowledge.

But the economics of youth chess clubs are daunting. Chess equipment is cheap, with a set that will last twenty years costing cents per player per year. While there are sensor-laden boards that can stream moves across the Internet in the IoT style where the "things" are the chess pieces, such as the DGT smart board, they are much too expensive for a typical club. Computer vision is the better economic choice over a sensor approach.

Either a smart phone or an AR head set is a reasonable choice to host the app. But headsets are both expensive and new to most potential parent volunteers. Smart phones, on the other hand, are already in the pockets of most parents.

The computer vision problem breaks in to three parts – finding the board in the image, creating a 3D coordinate space that finds all 64 squares, recognizing the pieces in legal play on the board, and then correctly placing the pieces on the squares. Then the position can be stated in the standardized Forsyth-Edwards Notation (FEN) and passed to a chess analysis engine. The engine can then check if the position is legal, checkmate, draw or stalemate and communicate that to the volunteer. It can also analyze the position and provide coaching info. The architecture is illustrated in Figure 4.

[22] M. Thomas, "Scholastic Chess Clubs: 10 Reasons Why," SAS Voices, Aug 2014 [online]. Available: https://blogs.sas.com/content/sascom/2014/08/29/scholastic-chess-clubs-10-reasons-why/

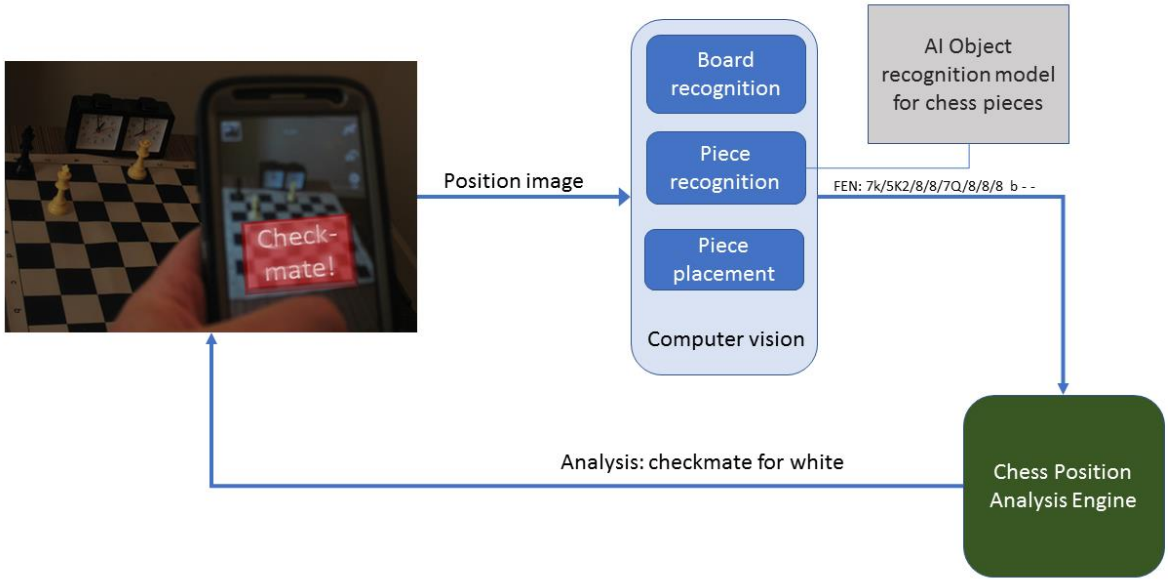


Figure 4: Architecture for Augmented Reality Chess Coach

While the output is still analytical in nature, traditional charts would not be prevalent. If a pair of third graders raise their hands to ask if they have arrived at stalemate, then the single word 'stalemate' or 'checkmate' is all the output the parent volunteer needs. Even coaching tips, such as "count the number of attackers and defenders on the e4 square" are best expressed textually rather than visually. It is not always the case that reality analytics is visual.

The same architecture could be adapted for other usages by replacing the chess position analysis engine with another type of analytical engine. This engine could be built with machine learning or other techniques.

Smart Facility Maintenance with a Digital Twin and AI

As sensor-laden facilities become smarter, digital twins can be developed to aid in the care of smart facilities. Ideally, a digital twin for a facility would include operational models of machine behavior provided by manufacturers. When manufacturers deploy machines that "phone home" with their operational data they can build a strong digital twin with machine learning based on many cases of production usage. Then, an individual facility can compose a digital twin by combining those digital twins into a composite digital twin for the entire facility. The facility digital twin would also absorb building plans, such as CAD drawings, and encapsulate all of it in to a gaming engine project that reaches out to the included models. The gaming engine app front-ends

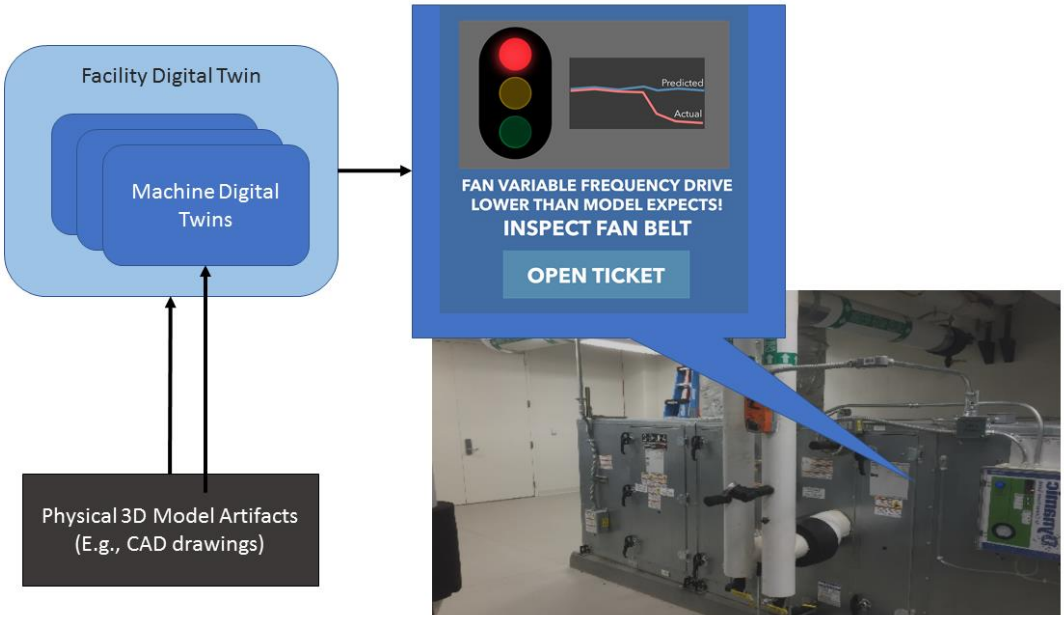


Figure 5: XR and Digital Twins

the digital twin for human consumption over AR or VR.

For performing the maintenance work, the digital twin provides a powerful twist on the typical remote expert use case, as discussed earlier. An AI expert could either completely replace a human expert, or, more likely, significantly aid the human expert.

For the field technician, a handheld computer such as a smart phone or tablet with an AR app remains a valid choice, but an HMD device is a much more plausible fit than in the previous example. For many field technicians, being heads-up and hands-free is advantageous enough to justify the cost of an AR HMD. An HVAC technician, for example, can enter a machine room and fix an air handler without ever consulting a manual or occupying their hands with a tablet or smart phone and without ever possessing a check list beforehand. Before leaving, the HMD could even direct them to address several other issues so that they would not have to return to the room for a while.

Smart City Emergency Response

Cities are getting smart, thanks to IoT, but all the technology increases the attack surface available to wrongdoers. Traffic accidents can be caused by manipulating signals or even directly hacking vehicles. Building systems, including locks and ventilation, could also be coerced into malignant service.

When responding to IoT based cyberattacks and other emergencies, a city can employ both AR and VR in conjunction with streaming analytics, as illustrated in Figure 6.

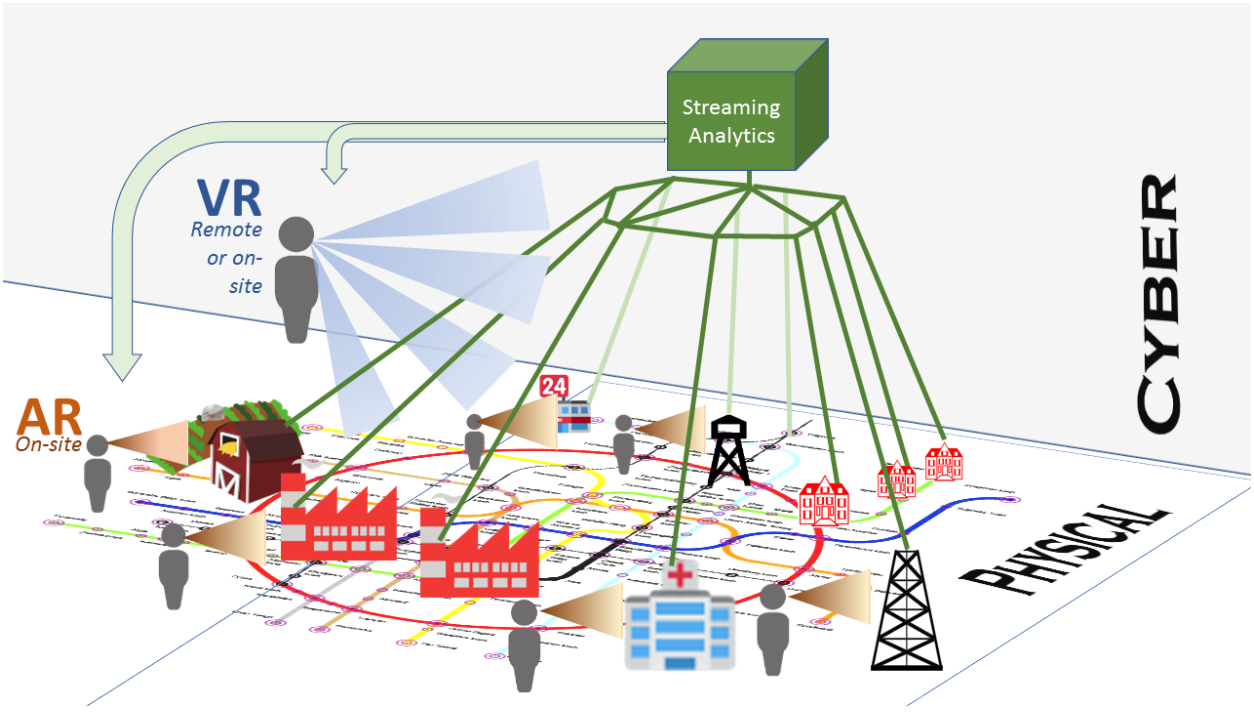


Figure 6: Smart city emergency response with XR and streaming analytics

In an instrumented and digital twinned city, both AR and VR techniques can be useful. At street level, first responders and other field workers can use AR to understand the city's infrastructure and emerging conditions. Either smart phone or a HMD could work, though the heads-up hands-free capabilities of HMDs would be very appealing in this case. Imagine a fireman going into a building that has hazardous chemicals and active gas lines. Fumbling for a smart phone with heavy gloves is far from ideal. But having a rugged HMD as part of the outfit could do much to keep the fireman safe and effective. Still, smart phone and even tablet usages of AR could be useful for those that are not so directly hands-on and in the literal line of fire. For example, a city worker who isn't a front line professional could be deployed in a crisis response and be effective after a quick app download to the phone. That would be far easier and less costly than having excess AR HMDs available.

Remote analysts can utilize VR techniques both to understand a city-wide crisis holistically as well as assist workers on the ground at a specific site. While a first responder is always anchored to a physical location and can only acquire different perspectives at the speed of available locomotion, a remote analyst can easily teleport about a virtualization of the same scene and see through buildings. They can also quickly slice their time between different sites that may be different parts of the same crisis. A VR HMD is a good option for this work, but analysts could also use

desktop flat screens towards the same purpose.

Tying together the field and remote workers is streaming analytics and AI. Both field and remote devices can be fed in real-time by the same engine over standard network protocols. Just as two different users of the same web site can have a consistent view of live data, a proximate first responder and a remote analyst can share the same view. Through standard telephone voice technology, they can stay in sync and communicate around the same data. The streaming analytics engine can tie in to machine learning inference just as in the previous example.

CONCLUSIONS

This article introduced and explored the notion of intelligent reality -- an underlying conceptual transition towards the contextualization of IoT analytics by utilizing gaming engines, AI and computer vision software. The combination of XR, AI and IoT should lead to greater penetration of data science into the world of work. When these technologies operationalize analytics, cost savings can be realized across an enterprise.

- Return to the beginning of this article

- Return to the Table of Contents

Outcomes, Insights and Best Practices from IIC Testbeds: Smart Factory Web Testbed

Interviewee:

Dr. Kym Watson
Principal Scientist
Deputy Head of Department Information Management and Production Control
Fraunhofer IOSB
Fraunhofer IOSB is a member of the "Fraunhofer Center for Machine Learning"
kym.watson@iosb.fraunhofer.de

Interviewer:

Joseph Fontaine
VP Testbed Programs
Industrial Internet Consortium
fontaine@iiconsortium.org

INTRODUCTION

In order to extend the usefulness of the published Testbeds in the Testbed Program of the Industrial Internet Consortium (IIC), the Testbed Working Group has developed an initiative to interview the contributors of selected testbeds to showcase more insights about the testbed, including the lessons learned through the testbed development process. This initiative enables the IIC to share more insights and inspire more members to engage in the Testbed Program.

This article highlights the Smart Factory Web Testbed. The information and insights described in the following article were captured through an interview conducted by Mr. Joseph Fontaine, Vice President of Testbed Programs at IIC, with Dr. Kym Watson, Principal Scientist and Deputy Head of Department Information Management and Production Control at Fraunhofer IOSB. Kym is an active member in the IIC where he has been serving as co-lead of the Smart Factory Web Testbed and is a key contributor to the Testbed Working Group. Kym co-chairs the IIC Distributed Data and Interoperability Management Task Group. In May 2018, Kym was recognized by his peers and bestowed the IIC Testbed Award for his leadership and contribution to the Smart Factory Web Testbed. His nomination indicated the importance of improving manufacturing order fulfillment and cited Kym's technical expertise, support and advancement of the smart manufacturing activities within the IIC.

SMART FACTORY WEB TESTBED – FROM CONCEPT TO REALITY

The Smart Factory Web Testbed aims to set up a web-based platform to allow factories to offer production capabilities and share resources to improve order fulfillment in a much more flexible way than is currently possible with available technology. It seeks to provide the technical basis for new business models, especially for small lot sizes, with flexible assignment of production resources across factory locations. This testbed is designed, in particular, to be a step towards establishing a marketplace for manufacturing where one can look for factories with specific capabilities and assets to meet production requirements. Factories offering those capabilities can then register to be located and participate in the marketplace.

This requires up-to-date information about the capabilities and status of assets in the factory. The characteristics of the products—availability, quality, price and so on—provides a basis for possible negotiation between competing offers.

For this application to work, international standards such as OPC Unified Architecture (OPC UA) and AutomationML are needed to link factories into the Smart Factory Web in order to provide information about the factories in a standardized way. This innovation enables production facilities to offer their services in a global market business and adapt their production in a very efficient way. The Smart Factory Web Testbed enables cross-site usage scenarios with secure Plug & Work functions and data

analytics. It reduces Information Technology (IT) system integration and installation costs, allowing for faster engineering and ramp-up time of components, machines, plants and IT systems—improving upon the utilization of equipment, as well. The core functionality is to describe the capabilities of factory assets in a standardized way, to find assets with the necessary capabilities and to access status data about these assets so that they may be included in the overall order management.

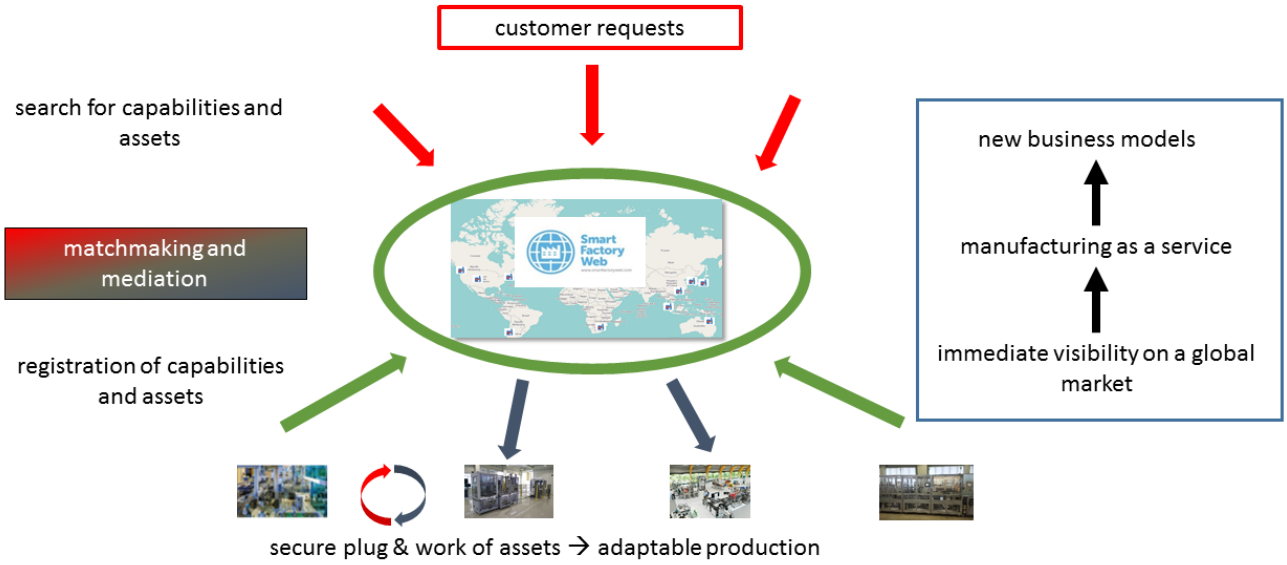


Figure 1: Smart Factory Web as a Marketplace for Manufacturing

The Testbed is directed mainly towards small-lot size environments rather than large manufacturers because companies working with larger line orders usually have their own supply chain management system and do not need to be as flexible and responsive due to the size of the orders. For smaller scale production, there are many more examples of where a moderate or smaller number of a particular part is to be produced, and machine capabilities need to be configured for this particular order.

To accomplish its goal, there are several areas of experimentation in the Smart Factory Web Testbed, including the engineering of automation systems for Plug & Work assets in a factory, flexible engineering, configuration of factory integration into Smart Factory Web and the Microsoft© Azure® platform, and the description of assets in AutomationML.

The Testbed's primary use cases involve manufacturers who seek to find a factory to produce certain parts. The manufacturer accesses the Smart Factory Web to find a factory with the right capabilities, and a potential target factory is identified. After negotiating with the target factory about delivery route, schedules, price and so on, an order can be placed. The target factory may need to adapt its production to meet the requested product specifications, and it

wants to do this as efficiently as possible. Once the production order is finished, the factory provides the finished or partial product to the original manufacturer or to another element in the overall supply chain.

This usage scenario, *Order Driven Adaptive Production*, is a combination of the application scenarios "order controlled production" and "adaptable factory" as defined by Plattform Industrie 4.0 (PI4.0)[1]. In further detail, this scenario is split into the following sub-scenarios:

Sub-Scenario 1.1 Publish: Registration of Smart Factories

Realized in Phase 1: "Geospatial Mapping and Factory Information" with the help of AutomationML to describe factory capabilities and assets.

Sub-scenario 1.2 Find: Discovering Smart Factories

Realized in Phase 1: "Geospatial Mapping and Factory Information" to find smart factories registered in the Smart Factory Web with the desired capabilities best matching the order requirements.

Sub-scenario 1.3 Order: Management and Execution of Orders in Smart Factory Web

The workflows to broker, orchestrate and process production orders in the Smart Factory Web constitute this sub-scenario, but they are not part of experimentation in this testbed. A proof-of-concept implementation in the Smart Factory Web Testbed will handle the ordering workflows and modeling of supply chains. The proposed IIC testbed "Negotiation Automation Platform" led by NEC© will extend the concepts of the Smart Factory Web and take up this sub-scenario.

Sub-Scenario 1.4 Adapt: Adapting the Factory Production

Realized in Phase 2: "Plug & Work" to flexibly and efficiently adapt a production facility to meet order requirements.

Sub-Scenario 1.5 Bind: Smart Factory Web Asset Connectivity and Monitoring

Realized in Phase 3: "Data & Service Integration" to provide current information on product and asset status (including availability of free capacity) for exploitation in the Smart Factory Web, especially to support the discovery process and linking of supply chains through secure data exchange. The Smart Factory Web information model will be updated dynamically.

Sub-Scenario 1.6 Collaborate: Collaborative Engineering

To be realized in Phase 4: "Collaboration" to enhance the efficient adaptation of factory production with shared engineering workflows and software Plug & Work.

[11] https://www.plattform-i40.de/I40/Redaktion/EN/Downloads/Publikation/aspects-of-the-research-roadmap.pdf?__blob=publicationFile&v=10

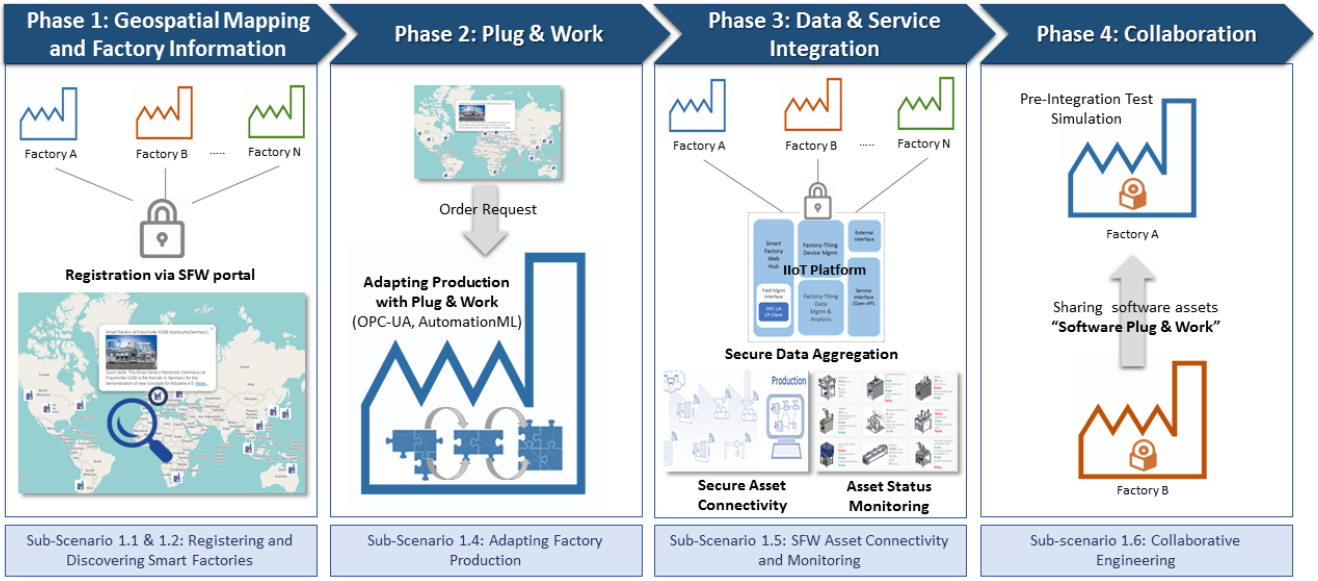


Figure 2: Phases of the Smart Factory Web Testbed

There are three primary technologies involved in the testbed. The first is the OPC UA, used to implement data communication between factories in the Smart Factory Web. Second, the standard AutomationML is used to describe the necessary information models—the semantics of the data transport from the factory to the Smart Factory Web. The other fundamental technology is the Smart Factory Web portal, a web-based information management system and application development environment which provides full support for access rights, work flows and ontology-based information models.

The primary experimentation for the testbed is working out an effective way of describing assets and capabilities and developing very efficient ways of achieving the overall software engineering where a new asset can be introduced. An asset can be described in terms of its capabilities but also in terms of its information model as interfaces. That asset must then be integrated into the information flow of a factory, the Smart Factory Web, and potentially cloud platforms such as Microsoft Azure. The testbed's core challenge lies in the software engineering processes, in an effort to make a factory adaptable. Other considerations include the electrical and mechanical interchangeability of a new device.

The testbed is deployed in model factories located in Karlsruhe and Lemgo, Germany and Ansan and Pangyo, South Korea. The model factories in Germany are operated by Fraunhofer IOSB and those in South Korea by the Korea Electronics Technology Institute (KETI). The two factories in Karlsruhe and Pangyo deal with handling, filling and transport. Both factories involve filling small bottles with either pellets or fluid, transporting these bottles with a small conveyor belt, and emptying the bottles—

with a few quality inspections. The Karlsruhe factory will look at implementing the (PI4.0) Asset Administration Shell for a number of assets in the next few months to validate the concepts of PI4.0. The model factory in Lemgo in northern Germany also involves handling and filling but on a larger scale including assembly within a versatile production facility.

Ansan's model factory is a large facility with real production equipment to accomplish tasks including the visual inspection of pistons from a local vehicle manufacturer's factory. The model factory in Ansan features a fully implemented digital twin of the manufacturer's production. There is a real production line where various parts are transported and inspected. A floor simulation model of the robot motions of the conveyor belt show how the engineering process is actually conducted. Therefore, if a change to the line would be needed, it can be done in the simulation environment (the digital twin) before going live, which would otherwise be a high risk. The Smart Factory Web Testbed strives to work closely with manufacturing and automation companies and eventually transfer its technology into real productive environments.

To date, the main deliverables of the testbed are documents describing the key concepts, standards application and implementation architecture of the Smart Factory Web. These concepts can then be adapted and adopted for use by a company. Another planned output of the testbed is the experience of how to describe asset capabilities, efficiently integrating assets into an overall software architecture. Additionally, the testbed is driving standards by providing feedback to the relevant standards bodies—OPC UA, AutomationML, and also standards work within the IIC and PI4.0. While other organizations are working in the area of asset administration, the Smart Factory Web Testbed strives to play a forerunner role in this area by tackling the whole combination of technologies involved.

TESTBED PLANNING

The IIC ecosystem has played a significant role in the planning of the testbed. Regular presentations of the Smart Factory Web Testbed and resulting discussions with IIC members at quarterly meetings and special IIC events were important mechanisms which allowed for continuous discussions and constructive feedback about the Testbed's purpose and function. Participating in the IIC Member Pavilion at events such as IoT Solutions World Congress in Barcelona and Hannover Messe has led to high visibility of Testbed activities and a better understanding of the requirements and potential applications.

In establishing alliances for various extensions to the Smart Factory Web Testbed, the IIC ecosystem played an instrumental role. The IIC's collaboration with PI4.0 is enabling the realization of the I4.0 component concept for selected assets in the Smart Factory Web Testbed. An I4.0 component comprises an Asset Administration Shell (a digital representation) and the respective asset. Working with IIC member, Microsoft, the integration of factories into the Microsoft Azure platform led to visualizing factory

process data. A new IIC testbed for the brokering of production and logistic services was proposed in conjunction with IIC member NEC's, the Negotiation Automation Platform. An alliance with IIC liaison, International Data Spaces Association (IDSA), brought the implementation of an IDS connector for trustworthy data exchange between factories and the Smart Factory Web portal. Furthermore, the IIC ecosystem fostered collaboration between PI4.0 and the IIC, facilitated the international dissemination of the benefits of standards in a testbed, and promoted work on the description of assets of IIC members.

There have also been benefits for the companies operating the model factories—Fraunhofer IOSB and KETI. Both organizations perform applied research and development for industry. Through the Smart Factory Web Testbed, they aim to improve and better market their own offerings in the field of IIoT and automation. In addition, the Smart Factory Web Testbed is a showcase for products and technologies of participating companies, enhancing their market opportunities. The network of companies taking part will form an 'innovation community' supported by KETI and Fraunhofer IOSB to identify and fill technology gaps by linking the knowledge and requirements of users, companies and research organizations. KETI and Fraunhofer IOSB advise companies on technology assessment and development of technology roadmaps. Furthermore, the Smart Factory Web Testbed has been integrated into training programs offered by Fraunhofer IOSB and KETI on industrial automation and security, including the application of OPC UA and AutomationML standards.

In choosing partners, it was important that the prospective organization was a leading innovator in IIoT in the manufacturing domain and a strong promoter of open standards. In addition, expertise with the standards used in the Testbed was required to participate.

IIC INTERACTIONS

The 3-tier architecture of the Industrial Internet Reference Architecture (IIRA) was applied in two places within the testbed: 1) in each model factory and 2) in the Smart Factory Web with gateways to the factories in the edge tier. The testbed has also adopted the general terminology used in the IIRA, a crucial step to facilitate clear messaging to the rest of the industry.

The activities of the Smart Factory Web Testbed have contributed to several aspects of the IIC. The Smart Factory Web Testbed is a candidate vehicle for an IIC-PI4.0 collaboration aiming to trial the PI4.0 specification *Details of the Asset Administration Shell* which defines how data exchange shall happen between Industrie 4.0 components based upon international standards. In addition, Fraunhofer IOSB is using testbed results and its own experience to contribute to the whitepaper "Digital Twin and Asset Administration Shell, Concepts and Application", of the IIC-PI4.0 Joint Task Group. The IIC DDIM TG (Distributed Data Interoperability and Mananagement Task Group) is working on a whitepaper to be published in 2019 dealing with IoT information models for semantic

interoperability and the characteristics of these models with the aim of proposing a meta-model. The information models and standards used in the Smart Factory Web Testbed have been contributions to the DDIM TG work. NEC submitted a research and development proposal related to the Smart Factory Web Testbed for the Japanese government which has been accepted. As part of this large national project, NEC has proposed the IIC testbed Negotiation Automation Platform which extends the concepts of the Smart Factory Web Testbed platform and includes information models to describe assets and supply chains as well as AI methods for negotiation. The work of Fraunhofer IOSB will be carried out within the Fraunhofer Cluster of Excellence "Cognitive Internet Technologies".

Standards

The Smart Factory Web Testbed employs a plethora of noteworthy standards. When possible adaptions to a standard are identified, the testbed reports to the relevant standards body. This report may involve submitting a change request or undertaking an accommodating process, depending on the standards organization. Regarding Open Source projects for example, the Open Source communities can process comments submitted and incorporate changes into the latest releases of the software. The Smart Factory Web Testbed supports standards with Open Source Development as a way of trialing a standard and receiving practical feedback about the specification.

IEC 62541 standard OPC UA is used for the data transfer between automation devices within a factory, between different factories and between the factory and the Smart Factory Web Testbed. The standards work in OPC UA is supported by the Open Source project open62541 where Fraunhofer IOSB has made major contributions, see https://open62541.org . KETI will also be contributing to open62541 in 2019. In addition, Fraunhofer IOSB has developed the Fraunhofer Open Source SensorThings API Server (FROST). Both open62541 and FROST are deployed in the testbed, and these Open Source projects contribute to the maturity and onward development of the respective standards.

IEC 62714 standard AutomationML is used to describe the semantics of the data, which data will be integrated into the Smart Factory Web, and how that data is going to be visualized. The Smart Factory Web Testbed uses AutomationML to provide the basis for the automatic generation of OPC UA servers, following the standard Companion Specification OPC UA for AutomationML. Experience gained in the Smart Factory Web Testbed is fed back into the onward development of the Companion Specification.

One recent activity involved implementing OPC UA over an OPC UA publisher subscriber (pub sub) and running it over a Time Sensitive Network (TSN). The OPC UA pub sub is a relatively new aspect of OPC UA, and this implementation executed in conjunction with an IIC member helped to mature the specification of this OPC UA pub sub.

Certain areas of relevant standardization are not yet fleshed out in the industry but are needed to fulfill the overall use cases, such

as the area of geospatial data, e.g., to consider environmental aspects as part of a smart factory. The Open Geospatial Consortium has standards in this area, but they are not yet fully integrated into the standards typically used in the manufacturing automation domains, i.e., OPC UA and AutomationML. Another gap lies in the information models available for IIoT. While there is progress in the companion standards being worked on for OPC UA, the development process is ongoing.

A standard of Open Geospatial Consortium called SensorThings API is becoming popular in the IIoT domain: The Smart Factory Web Testbed uses this standard to easily integrate additional sources of data, particularly sensor data, into a Factory Web. The oneM2M standard is being used by KETI, though it is not central for the overall Smart Factory Web concept. This standard can be added into the Smart Factory Web if devices within factories support oneM2m. There are standards used from the PI4.0 area which describe reference architectures for service architectures.

More work is also needed in the specification of the Asset Administration Shell from PI4.0, a digital representation of an asset by which modular and interoperable digital twins may be built according to the Industrie 4.0 concepts. The Smart Factory Web Testbed hopes to provide support to this standardization work. Additionally, semantic descriptions of asset capabilities are an important aspect of standardization still needed in the industry, but there is further work to be done in this area.

TESTBED RESULTS

There are four phases in the Smart Factory Web Testbed:

- Phase 1: Geospatial Mapping and Factory Information
- Phase 2: Plug & Work
- Phase 3: Data & Service Integration
- Phase 4: Collaboration

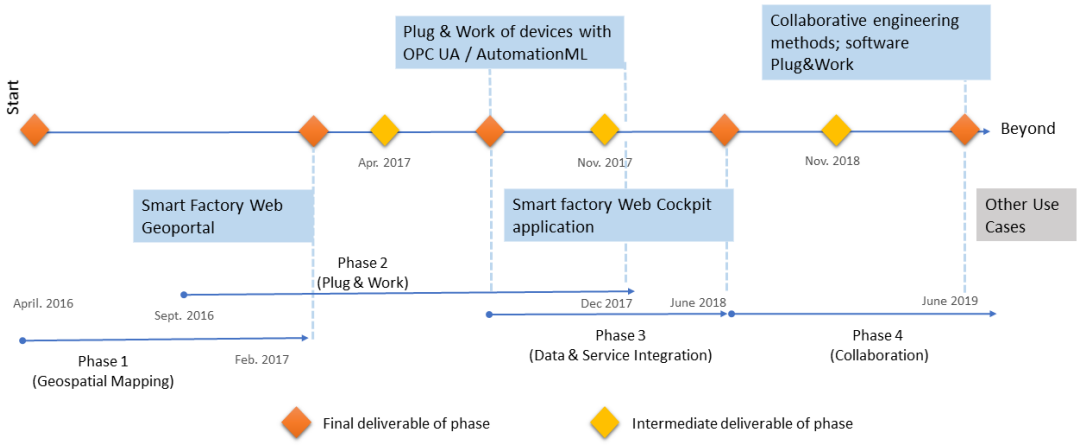


Figure 3: Timeline of the Smart Factory Web Testbed

The Testbed's architecture and experiences gained in the testbed over all phases will be documented in a technical design report to be published as a whitepaper in 2019. The Testbed team plans to extend the report to describe the work being done in the Digital Twin/PI4.0 Component Testbed, a project under the IIC-PI4.0 Joint Task Group and on the IDS connector.

Functionally, the first three phases have been completed up through the Data & Service Integration. Phase 4 involves the collaborative software engineering of these systems. There will be more work on the overall system architecture to include new developments with the Asset Administration Shell, as well as extensions of the Smart Factory Web Testbed to support the Negotiation Automation Platform from NEC. Though this work has started, the specifications are still a work in progress. Currently, the collaborative software engineering phase is ongoing—intensifying the work on the Asset Administration Shell, on the extension of the Smart Factory Web platform for other testbeds, and for the work with IDSA.

The technical report highlights the description of assets in AutomationML, covering:

- Their capabilities based on an ontology (to discover and integrate them as resources in a factory or supply chain),
- The definition of data to be sent to Smart Factory Web and Microsoft Azure through OPC UA or SensorThings API utilizing the automatic generation of OPC UA aggregation and FROST servers and
- The visualization of asset data in Smart Factory Web and Microsoft Azure.

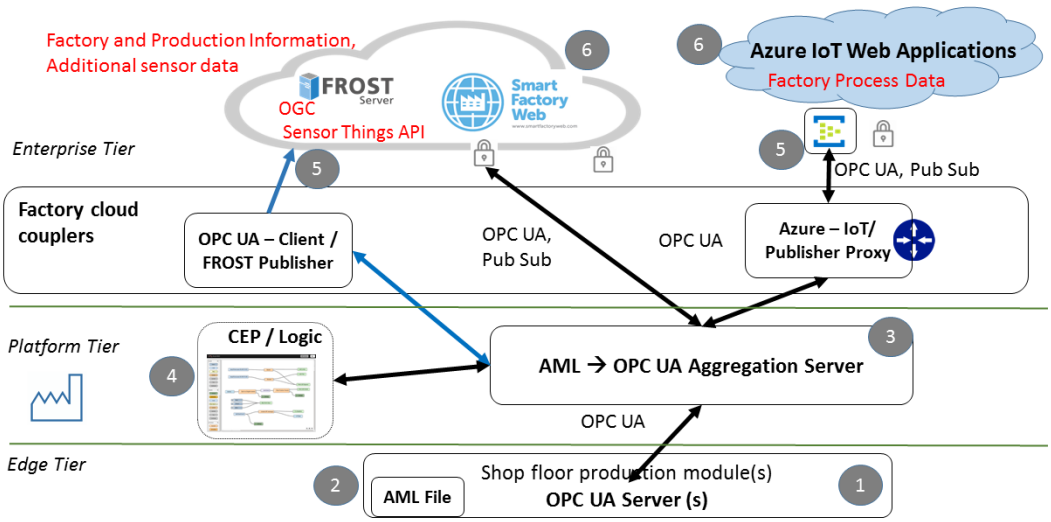


Figure 4: 3 Tier Architecture for Factory Integration. Abbreviations AML: AutomationML, CEP: Complex Event Processing, OGC: Open Geospatial Consortium, FROST: Fraunhofer Open Source SensorThings API Server

The results of Phase 3 are also summarized in the paper *Cloud-based Plug and Work Architecture of IIC Testbed Smart Factory Web* from the 2019 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (EFTA)[2].

Because the focus for Phase 4 is on collaborating to achieve the necessary software engineering to integrate factories together, the engineers of the various factories and assets in the factories are needed to provide data and semantics of their assets in a way that can be integrated into a cloud—Smart Factory Web or Azure.

There has been a notable level of interest in the Smart Factory Web Testbed coming from the industry, resulting in several types of customer engagement. Fraunhofer is currently working to form advanced, leading-edge models and move them into the industry. To enable this entrance into the field, the Smart Factory Web Testbed has had ongoing discussions with industrial companies to transfer research and development results from the experimental environment. This would entail setting up a type of Smart Factory Web for the production environment.

In addition, Fraunhofer is transferring general knowledge and training as part of its mission, and the Testbed has already conducted a number of training exercises on OPC UA and AutomationML for the industry. The Testbed has also transferred this knowledge to KETI, who are now conducting similar training sessions for Korean companies. Another example of customer engagement is consultancy work on how to design factories of the future and how to set them up to include new emerging technologies.

This area represents a challenge because there are so many new technologies arising, and it is difficult for anyone to assess whether these technologies will have a real impact and can be relied upon for the next fifteen years. In addition, the testbed must be able to transfer these technologies to client applications, help set up the necessary software environments and concepts, and take a multitude of steps to implement the Smart Factory Web or aspects of the Smart Factory Web in the clients' own workflows. It is crucial to increase the level of understanding and skills about certain technologies—trust in those technologies needs to be established so that there is a sufficient level of proven experimentation and best practices on how to apply the technologies. This level of trust is necessary before using these technologies in critical manufacturing applications where large production costs, and employee well-being, is at stake.

One of the major lessons learned from the Testbed is that open interfaces based on standards are essential to realizing a system architecture that can be adapted to changing requirements and technologies with a

[22] Heymann, S.; Stojanovic, L.; Watson, K.; Nam, S.; Song, B.; Gschossmann, H.; Schriegel, S.: "Cloud-based Plug and Work architecture of IIC Testbed Smart Factory Web". Proceedings of 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), Torino, Italy, September 4th to 7th, 2018

reasonable effort. The new version of the technical design report will contain best practices and how to set up the overall system architecture. It will be a blueprint comprised of advice on how to accomplish this integration in a sustainable way.

EXPERIENCE

The Smart Factory Web Testbed derives different forms of business value from participating in the IIC Testbed Program. The testbed has been able to procure new projects in the IIoT domain based on the experiences gained, as well as the marketing support given by the IIC. Visibility and the number of clients in major IIC regions—Europe, North America and Asia—have noticeably grown. From the perspective of the IIC member companies, it is hoped that there will be value for new clients to be able to apply some of the key technologies from the testbed more efficiently and with a higher degree of confidence.

The Smart Factory Web Testbed would offer three pieces of advice to other testbeds and companies considering an IIoT implementation:

1) Follow open standards as far as possible—this is a prerequisite to the second piece of advice.
2) Develop a sustainable, robust and flexible implementation architecture where one can make adaptations and demonstrate new technologies as easily as possible.
3) Ensure that there are sufficient accompanying projects to maintain synergy, funding and stakeholder commitment—this will bring the testbed from concept to reality and help maintain it over a period of several years.

CLOSING

Having been through this testbed process and coming to the end of Phase 4, the Smart Factory Web Testbed team finds that they did not encounter many major surprises in the technical area, but were surprised by their findings in the area of marketing. The level of interest in Smart Factory Web for various application scenarios involving cross-facility collaboration is much more prolific than originally expected. There are many different ideas and opportunities to transport these ideas to different applications, especially where some form of cross-organization or cross-facility collaboration is needed.

The Smart Factory Web Testbed embraces the spirit of why the IIC offers its testbed program. The level of effort put into the Testbed correlates with the high level of output and discovery, and the Testbed continues to be a model example of innovation in the IIoT domain.

- Return to the beginning of this article

- Return to the Table of Contents

Keeping Ahead of the Curve with Custom ASICs

Authors:
Edel Griffith
ASIC Technical Marketing Manager
Adesto Technologies
edel.griffith@adestotech.com

Darren Hobbs
Vice President, ASIC Sales & Marketing
Adesto Technologies
Darren.hobbs@adestotech.com

Sohrab Modi
Chief Strategy Officer
Adesto Technologies
sohrab.modi@adestotech.com

INTRODUCTION

Disruption in electronics has always been considered as an output from the consumer market. However, for a change, a recent candidate for disruptive technology is very much the Internet of Things (IoT) in industrial markets. Since the German strategic initiative of Industry 4.0 was announced to the public at the Hannover Messe Industrie fair in 2011, the Industrial IoT (IIoT) has spread around the world and is disrupting industry in all territories. With the IoT, communication can seamlessly occur between cyber physical systems and humans in real-time and via the Internet of Services[1], making possible the vision of smart factories, where a virtual copy of the physical world can be created, and decisions decentralized.

When the concept of IoT was first proposed, it was envisaged that all the data measured could be transmitted to the cloud, stored there and then the data retrieved whenever it was needed. With the daily numbers of over 2.5 quintillion bytes of data [2] being created and growing year by year, the concept of a purely cloud-based computing being the only solution is raising questions. Is moving all data to the cloud always a good idea? Indeed, is it always necessary? Can it be a bad idea? What happens if there is latency in communications? What is the cost associated with transmitting, gathering and storing of all this data? Does this mean an end to your IoT adoption? Is it really necessary to achieve your business goals?

While disruption is sometimes seen as being negative and bringing fear of the unknown, early adopters do in fact reap the benefits. These include the ability to be more proactive rather than reactive, having better control over inventory and facility management, being able to optimize logistics and having improved safety. To achieve these benefits however, cost, size, performance and other optimizations must be more flexible and responsive to end customer need. Doing the same thing the same way will no longer cut it. With the IIoT, demands are greater and more and more it becomes apparent that standard commercial off-the-shelf chips are not always the answer to developing each system as each provider has unique requirements. Awareness is growing that custom silicon is a compelling solution to reap the benefits from the disruptive forces of IIoT.

Custom chips or Application Specific Integrated Circuits (ASICs) are every device maker's dream, offering a single piece of silicon packaged in a single chip that is highly integrated, optimized and efficient, designed specifically for your product requirements. But over the years, ASICs have had bad press. They were seen by many as expensive and a luxury only of companies that were shipping millions of units a year and likely focused on the consumer markets.

[1] https://conceptsystemsinc.com/the-internet-of-services-in-industrie-4-0/

[22] Forbes. How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read https://bit.ly/2TTLHNZ

Custom ASICs were not considered possible for lower-volume industrial markets. In addition, the semiconductor industry has been laser-focused on supplying standard parts to the homogenous, large-volume consumer markets (the smart-phone being the pinnacle). Today, with the slow-down in the consumer segment, the semiconductor industry is now heavily energized to seek out new growth segments.

In this paper, we will show that custom silicon is no longer the expensive solution unique to high-volume products. With the increasing changes in technology, which will be addressed later in the paper, custom silicon is now economically viable for smaller volumes (50k+ units/year). And with the changes that the Industrial Internet of Things is bringing, we will show how custom ASICs are keeping suppliers ahead of the curve by enabling their edge processing applications through integration, power budgeting and cost reduction. With integration, it is possible to incorporate, into a single piece of silicon, in a single packaged chip, all the circuitry needed to sense, calibrate, control and communicate.

INDUSTRY CHANGES

The level of changes in Industry is at an unprecedented level. Industry 4.0 is driving phenomenal growth in hardware requirements. The number of sensors being deployed is growing year on year. In the foundries the wafer capacity[3] is increasing and silicon revenue continues to grow year on year. Data released by SEMI shows this growth in 200mm wafer size capacity in the foundries. 200mm are typically the wafer sizes for technology nodes greater than 0.18μm.

[3] https://m.eet.com/media/1309524/200mmforecastSEMI-min.png

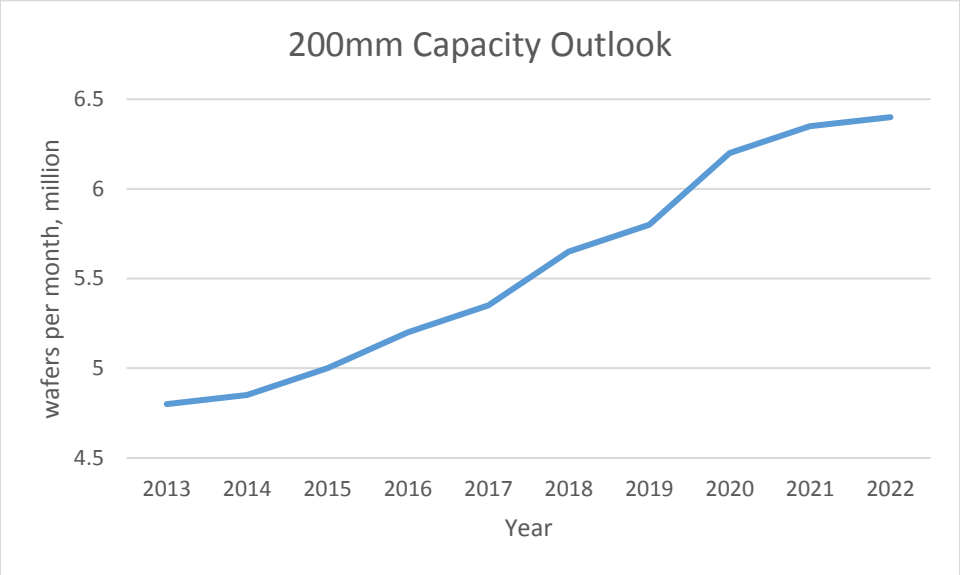


Figure 1: SEMI

Market

Market reports from Frost & Sullivan[4] are predicting that the revenue from the global sensor market in IIoT alone will grow to $11.23 billion by 2021 with a CAGR of 16.8%. While the analyst firm IDC[5] is projecting the worldwide spend on IoT to surpass the $1 trillion mark in 2020 with manufacturing being one of the industries that is expected to spend the most.

As IoT adoption increases, there is also growth in the level of automation and the number of sensors being monitored. Changes in the Industrial market are occurring as the move from being reactive to being predictive occurs. This has a knock-on effect on increasing the levels of monitoring taking place and this drives an increase in the number of components in a system performing the processing and communication demands. Accurate monitoring of these sensors is vital and there also needs to be a seamless flow of correct information in and out of systems, which is driving an increase in the proliferation of specialized types of integrated chips and circuits, such as custom ASICs to meet these needs.

Chip Technology

In 1965, Gordon Moore highlighted how the number of transistors per square inch of integrated circuit had doubled every year since their invention. This observation became known as Moore's Law and it has

[4] Frost & Sullivan, Analysis of Sensors in the Global Internet of Industrial Things Market, 2015

[5] IDC, Worldwide Semi-Annual Internet of Things Spending Guide
www.businesswire.com/news/home/20171207005963/en/IDC-Forecasts-Worldwide-Spending-Internet-Things-Reach

largely held true over the years with it being pushed to every two years from about 1975[6]. Technology nodes or process nodes are a reference to a semiconductor manufacturing process and the design rules that define the process. Each process node will often enable different circuit architectures. The smaller the technology or process node, the smaller the transistors available and therefore the more integration is possible[7]. The major drivers in pushing the decrease in size of the process nodes available in the foundries has been attributed to the consumer market space. As the demand for more features with smaller, less power-hungry and faster circuitry continues to grow, the requirement for more transistors on chip continues to increase. To meet these requirements, the transistors need to reduce in size thereby forcing the process nodes to reduce in size.

According to data published by International Business Strategies and reported by ExtremeTech[8], the cost of advanced design goes from $28.5M at 65 nanometer (nm) up to $542.2M at 5nm.

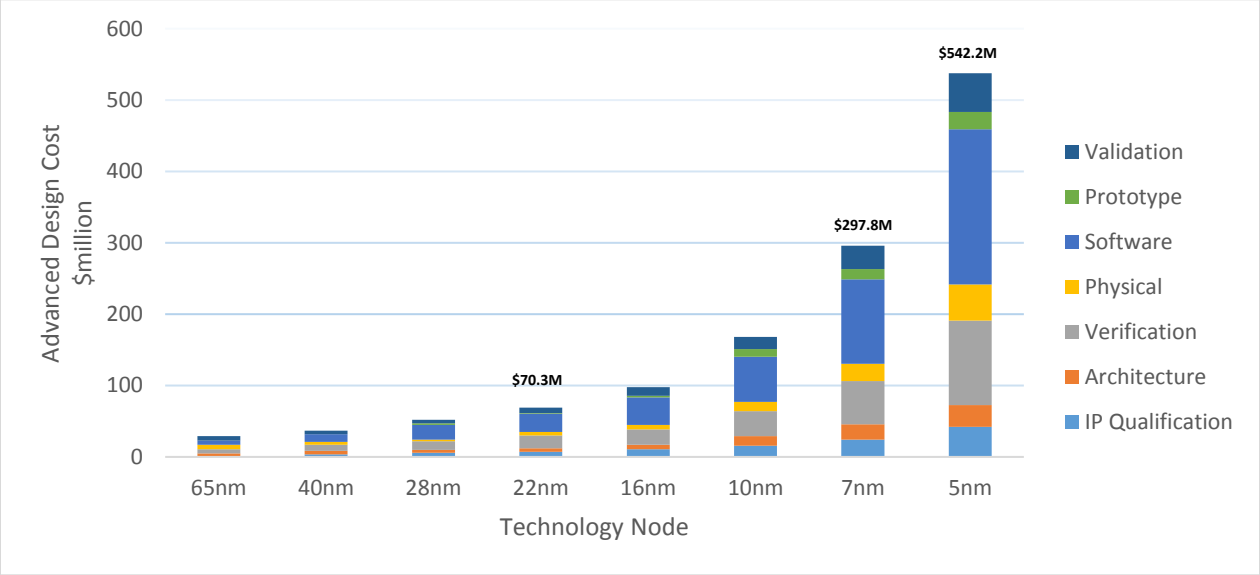


Figure 2: IBS data on the cost of Advanced Design

[6] Progress in Digital Integrated Electronics, IEEE Technical Digest 1975

[7] https://en.wikichip.org/wiki/technology_node

[8] J.Hruska, As Chip Design Costs Skyrocket, 3nm Process Node Is in Jeopardy
https://www.extremetech.com/computing/272096-3nm-process-node

Reports by SEMI[9] show that the foundry market is projected to grow to $97.5 billion by 2025. The high-volume markets will continue to push for technological improvements and be part of the bleeding edge nodes, thus freeing up the demand on more mature nodes. When we think about the bleeding edge technologies of <20nm, we think of digital circuits and how optimization at these nodes tends to be centered around the digital functionality. That is not to say that there are not analog circuits at these nodes. However, it should be noted, that analog circuits do not always directly benefit from the effects of scaling like digital circuits do. And in actual fact, reducing the size of analog components does not improve always performance, and having them in close proximity to noisy digital circuits can cause interference in the analog performance[10]. Estimates of process node usage by SEMI show that over 50% of the demand will still be for processes greater than 20nm and also that production is on the rise at older but cost-effective technology nodes[11]. So, what does this mean? The drive for more content for automotive, industrial, IoT and mobile applications are driving this need for 200mm wafer demand and production in older technologies. It is great news for developers of products for the Industrial IoT. Advanced analog and digital circuitry to sense and measure from the large number of sensors can readily be developed on these mature nodes. These mature nodes have been well tested in the past and therefore are fully de-risked. With the foundries wanting to maximize fabrication (fab) utilization across all process nodes, now more than ever these fully depreciated fabs are offering costs that are much lower than the bleeding edge processes. Fab utilization is the total percentage of the plant and equipment that is in production use in any given time frame. Foundries want to maximize this number, as the machines are running regardless of the volume throughput and therefore they want to get the most usage out of them. The higher the utilization rate, the more volume that has been processed and therefore the more stable and reliable the processing becomes.

[9] Dr. Handel Jones, Semiconductor Industry from 2015 to 2025, International Business Strategies (IBS) www.semi.org/en/node/57416

[10] https://semiengineering.com/mixed-signal-issues-worse-at-10-7nm

[11] https://www.eetimes.com/author.asp?section_id=36&doc_id=1334312

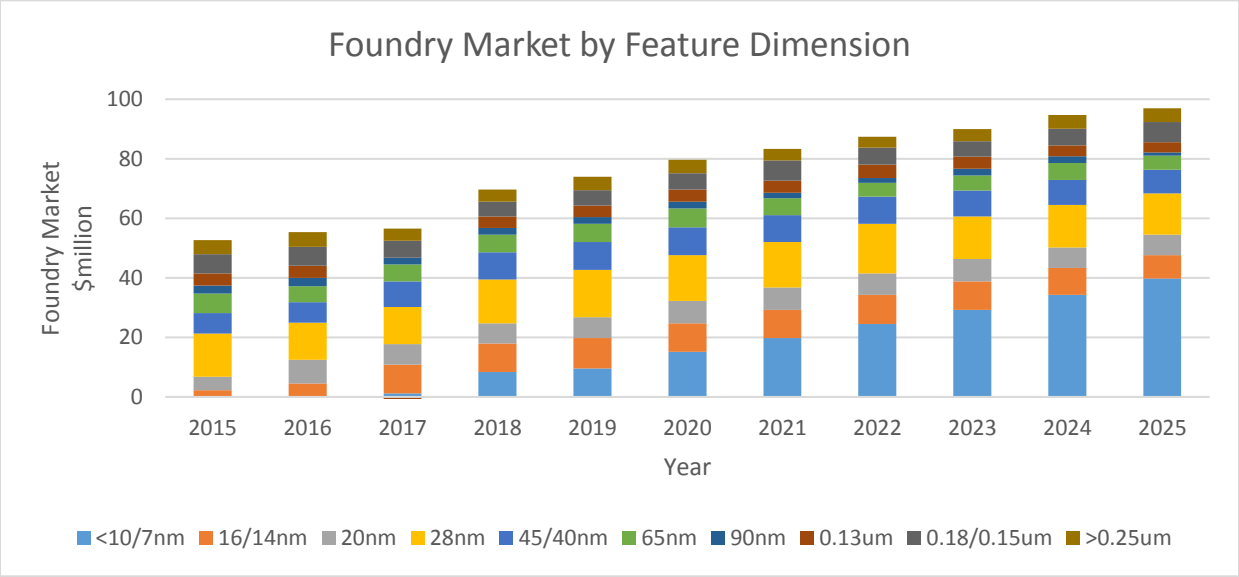


Figure 3: Foundry Market by Feature Dimension

As a result, there are excellent opportunities to leverage these technology nodes for the generation of custom ASICs for the IIoT and to do so at much lower costs than would have been thought previously.

EDGE COMPUTING

Earlier we discussed whether there is always a need to send all data to the cloud to achieve the desired business goals. In order to perform the level of monitoring needed, we need to install all the required sensors – and as already discussed, the numbers of sensors being deployed is growing at a large rate. In some cases, these sensors are placed in difficult to reach locations and running off battery power. Additionally, some remote locations mean communication constraints, so it is not always possible to communicate with the cloud when needed. So, while the ideal is to sense, measure and process all the data immediately and make intelligent decisions, the reality is that it is not always possible.

In real-time control, system processing can be performed at the source. This means that the data can be analyzed straight away, and decisions made quickly. A decision can then be made as to what data needs to be stored and what data must be sent to the cloud versus the data that is no longer required. This concept has become known as edge processing or edge computing. Edge processing involves bringing the processing power closer to the sensor edge where the physical world and the origins of digital data meet. Edge processing means response time is shortened and only the data that is truly required is transmitted. This helps to unburden the network and cut communication costs.

Edge processing is not without its difficulties, however. By bringing the processing tasks right to the sensor edge, there can quite often be space constraints. Also, as we

mentioned previously, the sensors can run off battery power and therefore power budgets available for processing circuitry can be demanding. Implementing a solution with discrete components while meeting demanding power budgets can be difficult.

CUSTOM SOLUTIONS ENABLING THE EDGE

The arrival of the Industrial Internet of Things has placed many demands on technology. We need to be able to monitor the data from large numbers of sensors which can be in space-constrained locations, working on battery power and with communications latency issues. Ideally, the sensors used would produce an electric signal that is directly proportional to the physical quantity that is being measured and therefore would allow a linear transfer function. However, this is not the case and the ideal sensor does not exist[12]. Therefore, we need to be able to monitor the sensor with better-than-ever accuracy, allow for these inherent non-idealities of sensors, react to the information received, and perform key functions based off that information. And then we need to be able to quickly store the data required in the cloud to be accessed whenever it is needed.

Integration

Custom silicon was historically considered the luxury of high-volume shipments. However, since advanced semiconductor nodes track consumer high-volume segments, the more mature process nodes have opened up for lower-volume products. This means that custom ASICs are now possible for many companies who would previously have found such designs out of their budgets.

Even what seems like a moderately simple printed circuit board can contain hundreds of components. Add to this the overhead associated with specifying, purchasing and testing, the time and cost can be considerable in choosing to go the discrete component path. And this is all before you consider risks of obsolescence and security of your intellectual property.

With a custom ASIC, you can integrate all your analog and digital circuitry onto one single piece of silicon. Added to that you can include a microprocessor or microcontroller, memory (Flash and SRAM), various interfaces and wired or wireless communication protocols.

Integration can give major surface area size savings – for example, a custom off-the-shelf 12-bit Digital-to-Analog Converter (DAC) discrete component may have a physical area of $10mm^2$. The same DAC with equivalent performance integrated into a custom chip occupies just $0.1mm^2$. Later in this article we will share case studies that show how customers have achieved 80-90% area savings on their systems by using custom silicon versus discrete component solutions.

[12] https://bit.ly/2RMudWh

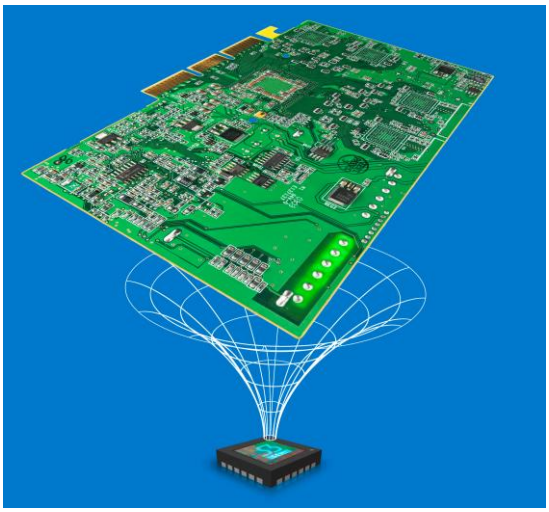


Figure 4: Visual representation of benefits of custom integration

Bill of Material Cost Reduction

Custom silicon not only leads to large savings in area, but also significant savings on your bill of material (BOM) costs. In particular, for the Industrial Internet of Things where the lifetime of products can be 10 years or more, the return on investment using custom silicon can be considerable. Also, with integration, you are moving from having all the costs associated with sourcing, storing and physical placement of large numbers of components, down to one custom chip which has incorporated all this functionality.

Protecting Intellectual Property

It is relatively trivial for a competitor to reverse-engineer a circuit board consisting of off-the-shelf semiconductor components. However, it is exceptionally difficult, if not inconclusive to reverse engineer an integrated circuit. Integration can protect the hard-earned intellectual property from been copied.

CASE STUDIES

Industrial

An industrial company in the oil and gas market developed their existing solution using commercial off-the-shelf discrete components. They were shipping in the region of 60k units per year. The cost of the solution was extreme and did not offer the flexibility the company needed if they were to be able to offer their end customer the advantages they desired – namely maximizing production while minimizing operating costs and being compatible with linear and rotary valves and actuators. High reliability was also a key requirement due to potential hostile working environment of the product. The company determined that they could benefit from a custom solution with:

- The ability to allow for portfolio tiering and expansion
- Multiple sensor interfaces (pressure, temperature, diagnostics)
- Integrated smart control loop
- Accurate value positioning
- Multiple communications protocols
- Integrated Arm processor and PIC controller
- Intrinsically safe by design
- Power efficient

As well as meeting the needs of their end customer, of being able to maximize production while minimizing operating costs and being compatible with different valves and actuators, this major supplier of equipment for the oil and gas market wanted to incorporate product tiering into the ASIC, something that was not

economically possible previously when using discrete components. Central to the design discussions were the sensing and measurement needs, the control, programmability, and connectivity needs and finally the security needs for the final solution. A solution using TSMCs 0.18um CMOS process was delivered featuring:

- Analog front end (AFE) comprising 14-bit SAR Analog-to-Digital Converter (ADC), 12-bit control DAC, power switches, analog multiplexers and operational amplifiers
- Multiple industrial communications interfaces including FOUNDATION Fieldbus and Highway Addressable Remote Transducer (HART)
- Arm Cortex-M4 CPU core
- PIC microcontroller
- Flash and SRAM memories
- Multiple peripheral interfaces including SPI, UART, I^2C and Parallel

The resulting solution which contained all the above functionality achieved the following results for the end customer:

Results
Bill of materials cost reduction of 85%
Substantial reduction in the physical footprint with a custom ASIC in a 19mm x 19mm package
Power efficiency, meeting the low-power budget supplied from the 4-20mA control loop
Design for portfolio tiering
Investment breakeven 42k units

Examination of the results also show that this customer broke even on the ASIC investment after shipping over 42k units and will have estimated cost savings of over $21m over the accumulated cost of the lifetime of the project.

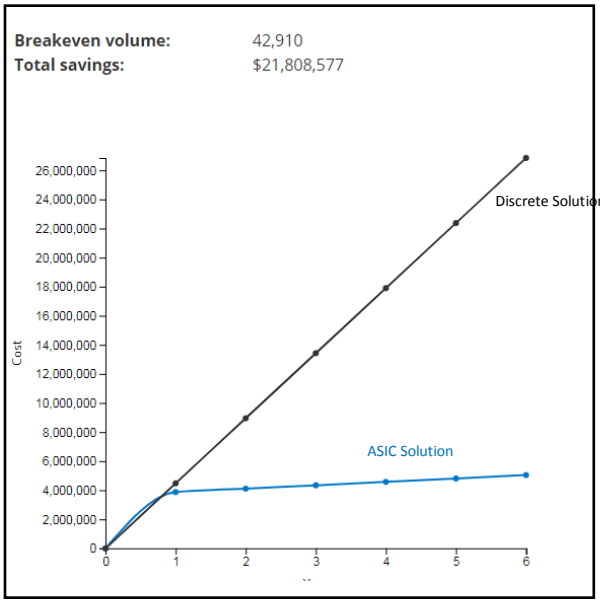


Figure 5: Comparison of ASIC solution vs Discrete Solution total savings including breakeven volume

In another case study, a company was developing Machine-to-Machine (M2M) technology in the area of mobile satellite services. They were at a run rate of 100k units per year. This is a very niche area of wireless communications, providing two-way voice and data communications for mobile assets in remote locations. Reducing the system size and being able to guarantee signal quality and enhanced connectivity while keeping costs low are key concerns. While the company had already moved from a fully discrete solution by using an Application Specific Standard Product (ASSP) on their board, the product was still too generic for their requirements, and was therefore not optimized for the performance

they required. It was inefficient and costly. As a result, they were in a "make do" situation and being challenged by their end customer on performance and price.

By going the custom ASIC route, the company was able to specify exactly the requirements they wanted, with key criteria including:

- Smaller physical footprint
- Improved performance
- Lower power consumption
- Decreased cost

The custom ASIC was developed in 12 months on a 0.18um RF CMOS process from TSMC and included:

- Integrated transmitter and receiver blocks (L-band receiver to support multiple modulation schemes)
- On-chip calibration functionality (RC Time constant, IP2, Image Rejection and IQ Gain/Phase)
- Embedded algorithms for DC offset correction

The resulting solution which contained all the above functionality achieved the following results for the end customer:

Results
55% reduction in the physical size of the product
Improved signal integrity and reliability
57% reduction in power consumption
80% reduction in the bill of material costs
Investment breakeven 175k units

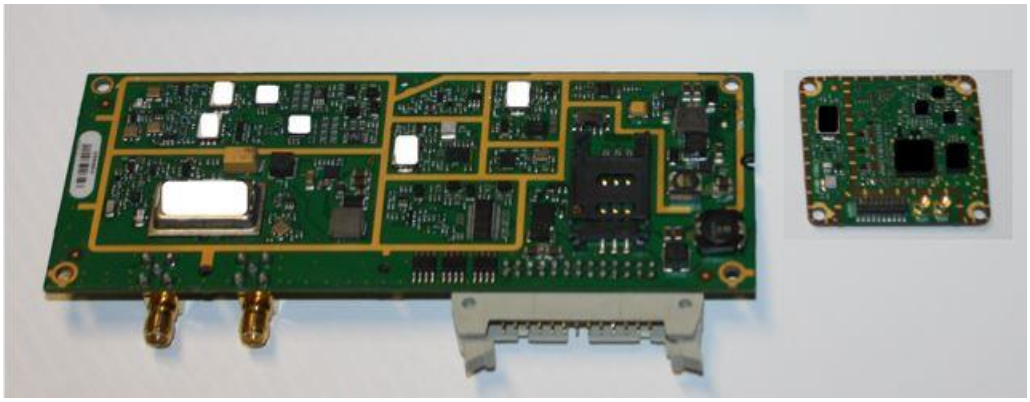


Figure 6: Before and after of product size comparison (resultant board is on the right)

CONCLUSION

While the Industrial Internet of Things offers many opportunities to enhance the efficiency and control of our processes and systems, and increase our business potential, there are many challenges. In this new era, where the cyber world is meeting the physical world, developing edge computing systems using the same tried and tested methods of the past with commercial off-the-shelf electronic components is no longer a viable approach. Considering the

heterogeneous requirements at the industrial edge, no one size fits all. Recent advancements in technology mean we can now leverage custom ASICs to provide optimized performance, lowest power, and smallest area at viable economies for the IIoT.

- Return to the beginning of this article
- Return to the Table of Contents

Improving Reliability and Security of Global Cold Chain Logistics for Pharmaceutical Assets

Authors:
Dr. Dave Stanton
Director of IoT Partner Engineering
Wipro
dave.stanton@wipro.com

Dr. Madhusudan Pai
Director of Global IoT Partner Engineering
Wipro
madhusudan.pai@wipro.com

OVERVIEW

Global cold chains require multiple organizations to collaborate, which results in a heterogenous system of hardware, software and assets. Cold chains in the pharmaceutical industry are particularly challenging, owing to the stringent requirements associated with the movement of life-saving vaccines, cultures and medications. This article explores the current state of cold chain logistics for pharmaceuticals and identifies weaknesses in the lifecycle of a cold chain. The article also discusses how IoT can play an enabling role, and concludes with a summary of prevention and risk mitigation strategies.

INTRODUCTION

A cold chain is a variation of a supply chain, whereby the assets that must be moved have additional requirements of being kept refrigerated or in some other manner under control of environmental parameters, such as light and humidity. Exposure of the assets to heat or humidity, even briefly, can cause diminished efficacy or complete waste.

While it is logically simple to refrigerate or otherwise store an asset in an environmentally controlled container, great complexity and risk is introduced when assets must move between multiple parties through a holistic supply chain.

In this article, we will describe the current state of cold chain logistics following the use case of pharmaceuticals, where a known chain of custody and assurance of quality has enormous implications for the efficacy of vaccines and other medical assets.

CHARACTERISTICS OF COLD CHAIN LOGISTICS

We can think of the cold chain as being the environmentally controlled lifecycle of a single good or asset and cold chain logistics as "a systemic project to ensure the quality and performance of goods in the production, storage, transportation, sales, and all aspects" of the product life cycle leading up to consumption of the product.[1] The single asset either has a continuous or broken cold chain. The overall machinations to store, handoff and transport the assets are the cold chain logistics. The importance of the following characteristics may vary across different types of assets, but they still must be considered at least in some manner to determine the efficacy of a cold chain logistics system.

Environmental Sensitivity

Every asset is sensitive to extreme environmental conditions. Excess heat spoils vaccines. Excess humidity fuels mold growth on fabrics. Excess light can destroy certain chemicals and films. There are many other examples of destructive environmental

[1] Wang, H., Lan, Y. and Kong, F. (2018). Research on Development Model and Strategy of Agricultural Products Cold Chain Logistics in Jilin Province. *IOP Conference Series: Materials Science and Engineering*, 452, p.022033.

factors such as vibration, gas composition, pressure and radiation. Most assets can be kept in standard atmospheric (i.e., ambient) conditions for a modest amount of time without suffering damage, but we cannot assume or rely on these kinds of gaps in the cold chain.

While we often need to control environmental parameters while manufacturing, storing and transporting assets to ensure the efficacy of the asset, it is unreasonable to spend far beyond the value of an asset. For food supply, general guidance is that logistics costs should not exceed 50% of the food cost.[2]

This guideline is harder to apply to pharmaceuticals, as the costs may be known but the societal value becomes difficult to quantify. An ineffective vaccine might lead to a human death. Even worse, a patient that we think is inoculated but truly is still at risk can inadvertently become an infection vector to other patients. To add complexity, different vaccines have varying sensitivities to heat and light.[3] There is not a single, perfect design for cold chain logistics to support all vaccines, so dynamic sensors and controls are critical.

Many Handlers

Many parties handle a vaccine during its lifecycle. First, the vaccine must be synthesized in a lab. Next, the vaccine is collected for transport. Then, transportation may occur over long distances using multiple modes of transportation such as by land, sea or air. Finally, local delivery brings the vaccine to the person that will receive the pharmaceutical.

Across these stages, there is ambiguity regarding who did what, for how long, and to what effect. The advent and adoption of cloud computing has made it easier for each handler to write logs to databases and keep all interested parties advised of the status of the cold chain.[4]

[2] Ji, G. and Guo, R. (2009). Research on the security of cold-chain logistics. *2009 6th International Conference on Service Systems and Service Management*.

[3] WHO. (2014). *Temperature Sensitivity of Vaccines*. [online] Available at: https://www.who.int/immunization/programmes_systems/supply_chain/resources/VaccineStability_EN.pdf [Accessed 9 Jan. 2019].

[4] Li, X., Wang, Y. and Chen, X. (2011). Cold chain logistics system based on cloud computing. *Concurrency and Computation: Practice and Experience*, 24(17), pp.2138-2150.

Stage Name	Handler Types	Enabling IoT Technologies
Manufacturing	Additive manufacturer Discrete manufacturer	Digitial twins Heavy edge compute Machine vision
Collecting	Land trucking Railway	LTE-M, 5G gateways Medium edge compute
International ports	Airplanes Ocean freighters	Satellite connectivity Heavy edge storage
Regional Hubs	Warehousing	Autonomous vehicles Track and trace
Freight	Land trucking Railway	LTE-M, 5G gateways Medium edge compute
City Hubs	Warehousing	Autonomous vehicles Track and trace
Local Delivery	Small trucks Cars	LTE devices

Figure 1: Stages of a Global Cold Chain

Evidence of Control

While sensors and IoT devices simplify the collection and logging of environmental parameters during the cold chain, there are still major concerns regarding evidence of control by a given party within the cold chain. Log files need to be immutable and unforgeable to be regarded as true evidence of control.

Blockchain and other distributed-ledger technologies have shown great promise in recent years to provide a distributed, immutable log shared by all parties participating in a cold chain logistics system.[5] Additionally, there needs to be a trust that participants acting within the system are properly authenticated and authorized. The Industrial Internet Consortium's "Industrial Internet Security Framework" [6] provides

[5] Plaga, S., Wiedermann, N., Anton, S., Tatschner, S., Schotten, H. and Newe, T. (2019). Securing future decentralised industrial IoT infrastructures: Challenges and free open source solutions. *Future Generation Computer Systems*, 93, pp.596-608.

[6] Industrial Internet Consortium. (2016). *Industrial Internet of Things Volume G4: Security Framework*. [online] Available at: https://www.iiconsortium.org/pdf/IIC_PUB_G4_V1.00_PB-3.pdf [Accessed 9 Jan. 2019].

guidance regarding the need to have a hardware root of trust built into the silicon used in the processing units of internet-connected devices.

To further complicate matters, it is nearly impossible to maintain an always-on internet connection throughout a global cold chain. Great advances have been made in cellular- and satellite-based communication networks but we must still accept a small amount of network downtown either during access-point handoffs or failovers.

Heterogenous Operating Conditions

There have been many attempts at creating frameworks to describe the utility of cold chain logistics systems, particularly for food. As there is typically a different handler for each stage of the cold chain (e.g., warehousing versus local delivery) and form factors between stages. This heterogeneity of inputs and outputs for different stages means we cannot assume a single container and connectivity type can be used to truly observe a cold chain from start to finish.

This fact of heterogenous operating conditions has led to optimization focused on a single stage of the cold chain.[7] While we should strive to improve what we can, it is a folly to overly optimize one stage (e.g., regional distribution centers) while neglecting another stage (e.g., local delivery).

RISKS ASSOCIATED WITH GLOBAL COLD CHAINS

A vaccine must make a long journey from fabrication in a lab to administration in the field. Along the way, there are many opportunities for both observed and unobserved damage.

Handoff

As damage can occur at any point in the cold chain, we cannot defer inspection solely toward the end of the chain. Typically there is some level of inspection at the handoff between stages in a cold chain. The inspections, for practicality, typically involve a sample of the assets instead of the entire population. For additive manufacturing (i.e., liquids and gases), a sample can be intrinsically trusted. For discrete manufacturing (i.e., solids), the likelihood of tampering and forgery increases drastically.

A palette with hundreds of assets will most likely be inspected by taking a sample from the outer layer of the palette. This ease of access reduces the complexity of the inspection but also increases the temptation for replacement, forgery, or other types of tampering. Vaccine vial monitors (VVMs) can be used to provide evidence of heat damage[8], but damaged vials can be hidden within larger lots midway through the cold chain. Also, VVMs cannot detect damage

[7] Accorsi, R., Cholette, S., Manzini, R. and Tufano, A. (2018). A hierarchical data architecture for sustainable food supply chain management and planning. *Journal of Cleaner Production*, 203, pp.1039-1054.

[8] Vaccination: rattling the supply chain. (2011). *Bulletin of the World Health Organization*, 89(5), pp.324-325.

from excessive freezing or vibration. In the end, the damaged vial will be found, but the responsible party will be unknown.

Authenticity

As the value of assets increases, the temptation for forgery and dilution increases. Within food supply logistics, saffron stands as a good example of risk and mitigation. Minimally invasive inspection methods were created that correlate spectroscopy and purity.[9] Additionally, the specific process for inspection was standardized as ISO 3632. Similar strategies should be adopted to create standardized inspection for common pharmaceutical asset types that require a cold chain.

Standards alone do not provide a guarantee of authenticity and purity. Tamper resistant IoT devices can be used for the inspection themselves, with the results written through machine-to-machine (M2M) communication to blockchain or other distributed database technology.

Root Cause for Damage

Each handoff between different parties within the cold chain creates a risk for false negatives, whereby one party takes over custody of the asset with the assumption that the asset is undamaged. These false negatives further delay the identification of the root cause. Data must be recorded continuously for us to have true confidence in the status of the cold chain.

As with inspection devices, tamper-resistant IoT devices can monitor environmental parameters of the cold chain and log data using unforgeable M2M communication. The ultimate goal for root-cause detection is to predict failure, repair equipment, improve processes and remediate the cold chain before assets have been damaged.[10]

Recall and Replacement

All participants in a supply chain agree to some manner of agreement regarding replacement of assets or repayment for damages. Identifying the responsible party often becomes a lengthy and messy legal exercise. For precaution, each party should be recalling any tainted or possibly tainted assets. Without clear identification of the root cause of damage, recalls will be broader and costlier than necessary and, undoubtedly, damaged vaccines will be administered to patients before the issue is identified.

IoT provides a great hope for faster identification of damaged pharmaceuticals, clear traceability to lot numbers, and hence much more focused and effective recalls. While vaccines are more generally administered, there are other medically related assets that are extremely costly to replace. An organ en route for a transplant

[9] Khilare, V., Tiknaik, A., Prakash, B., Ughade, B., Korhale, G., Nalage, D., Ahmed, N., Khedkar, C. and Khedkar, G. (2019). Multiple tests on saffron find new adulterant materials and reveal that Ist grade saffron is rare in the market. *Food Chemistry*, 272, pp.635-642.

[10] Anderson, R., Lloyd, J. and Newland, S. (2012). Software for national level vaccine cold chain management. *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development - ICTD '12*.

can be replaced, but the cascading delays can cause irreversible damage to patients and severe loses in utilization of medical personnel, equipment and facilities.

Diminished or Unknown Efficacy

Assets managed via a cold chain may either be completely unusable after sustaining damage or the asset may have some alternate usefulness. For many vaccines, excessive heat reduces the efficacy but does not destroy them completely. The damaged vaccines can still be administered to patients within certain age ranges and yield expected efficacy for those patients. However, the same damaged vaccines would not be effective if administered to patients outside of this age range.

The ability of IoT to provide more granular monitoring provides more resilience to the cold chain by identifying the lots or individual assets that have been minorly damaged yet can still continue within the cold chain to deliver expected results without compromising other assets or patients.

Energy Consumption

The energy costs for maintaining a cold chain can be significant. [11] Without detailed sensing and asset tracking, the participants in the cold chain will either have to use far more energy than is necessary, or they will be taking on unacceptable risk for spoilage. The survival of assets is paramount, but we must also strive for reducing energy usage whenever possible to meet green standards needed to be ethical participants in sustainable supply chains.

SUGGESTED IMPROVEMENTS TO GLOBAL COLD CHAINS

While the maintenance of a global cold chain is challenging, new technologies afford ways to improve the reliability and immediacy of monitoring goods. Additionally, distributed ledgers allow for increased trust by all participants in the cold chain. Combining immediacy and trustworthiness opens new possibilities for how we can improve the reliability of cold chains while reducing risks and costs.

Trusted Handoff Spaces

Varied lot sizes, transport modalities and regionalized data laws makes it unlikely that we will ever be able to have a single, standard way to move goods. As such, there will be a continued need to hand off a shipment from one party to another. During this handoff, we should provide redundant environmental isolation, such as a refrigerated crate being transferred within a refrigerated loading dock. These double-sealed locations can be designated as trusted places to take actions that would otherwise be deemed a breaking of the cold chain. For example, a tamper-resistant seal on a container owned by one party might need to be broken to shift goods to another container. Additional documentation and

[11] Chen, S. and Lan, H. (2016). The cold chain logistics enterprise's green level evaluation. *2016 International Conference on Logistics, Informatics and Service Sciences (LISS)*.

logging will need to be done in these trusted spaces.

Real-Time Tamper Detection

The majority of time in a good's lifecycle is spent between trusted spaces, notably during transportation. During this time, the custodian of the goods should provide real-time monitoring of the good for both environmental parameters but also for tamper attempts. This can be done through a combination of pressure, light and vibration. Sensor readings should be logged frequently and to a distributed ledger to prevent retroactive manipulation of data.

Distributed Ledger for Trustworthiness

The great hope of blockchains and distributed ledger technologies is to provide an immutable log of transactions. These transactions do not need to be financial but instead can be sensor readings proving the maintainance of the cold chain. Depending on the privacy implications of the goods, it may be necessary to use a private distributed ledger versus a public distributed ledger. The benefits of private distributed ledgers are beyond the scope of this article, but in short, there are energy/computation benefits of using private distributed ledgers for logistics when all parties within the cold chain are assumed to be known to each other.

CONCLUSIONS

IoT provides new ways by which we can monitor pharmaceutical assets such as vaccines in an unbroken cold chain from creation to administration. Further, new capabilities in data acquisition and edge-computing analytics capabilities now allow us to predict failures and remediate equipment before assets have been damaged.

Heterogeneous operating environments will continue to pose challenges, but advances in global cellular connectivity and distributed ledger technologies provide logical solutions to having real-time, tamper-proof asset monitoring for pharmaceuticals moving through cold chain logistics systems.

- Return to the beginning of this article

- Return to the Table of Contents

Automotive Security through New Communication Lockdown Utilizing Programmable Logic Solutions

Authors:

Dan Isaacs
Director Automotive Business Unit
Xilinx
dani@xilinx.com

Jillian Goldberg
VP Marketing
GuardKnox
jillian.goldberg@guardknox.com

Dionis Teshler
CTO and Co-Founder
GuardKnox
dionis@guardknox.com

Tal Nisan
Software Architect
GuardKnox
tal.nisan@guardknox.com

INTRODUCTION

In today's connected world, ensuring a vehicle's security must be addressed through a comprehensive understanding of all networked communication channels internal and external to the vehicle. This article presents an innovative methodology, Communication Lockdown, including Network Orchestration, from development to production, as implemented in a centralized communication gateway based on Xilinx's configurable Zynq® SoC programmable technology. The methodology, advantages and differentiating values of the technologies involved have been described.

PROBLEM STATEMENT

As vehicles drive towards autonomy, they multiply in complexity, becoming far more connected. Today's vehicles are highly sophisticated local area networks on wheels that controls numerous complex systems via hundreds of micro-processors, up to 150 ECUs [1] (automotive computers), and numerous sensors, interconnected by a high-speed, high-availability internal communications network.

The automotive industry is moving towards a service-oriented vehicle, where the passengers (or drivers) and their needs are the focal point rather than the vehicle itself. This concept is focused on the ability to continuously and securely change vehicle capabilities, instantaneously, on-demand and over-the-air (OTA) from future OEM app stores.

Automotive cyber security for modern connected and autonomous vehicles require a solution that is:

1) Cyber-secure: In today's vehicles, safety and security are inherently the same. Modern vehicles host hundreds of sensors and ECUs powered by more than 100 million lines [2] of software code. Cameras and sensing devices stream gigabytes of data in real time. A typical vehicle might also host several different types of local area networks such as CAN bus, Ethernet, and LIN. Manufacturers source hardware and software from different suppliers. No single player controls, or is familiar with, all of the possible attack vectors within any vehicle. As such, vehicles constitute a massive attack surface that could be used to exploit sensitive data, financial information and much more.

[1] Techopedia. "Your Car, Your Computer: ECUs and the Controller Area Network" https://www.techopedia.com/your-car-your-computer-ecus-and-the-controller-area-network/2/32218

[2] MIT Technology Review. "Many Cars Have a Hundred Million Lines of Code" https://www.technologyreview.com/s/508231/many-cars-have-a-hundred-million-lines-of-code/

2) Flexible and Scalable: Current and next generation automotive architecture will be based on high speed communication and high-performance computing. This will require the handling and securing of multi gigabit data channels and the running of dozens of applications per ECU. This task cannot be achieved with current architectures based on micro-controllers which are not optimized or flexible to cover evolving requirements. Cybersecurity solutions require both hardware and software flexibility and scalability to provide ample processing resources and provision for future software extensions/additional services. Having extra computing power and storage management from the onset will not require costly and resource-intensive changes to vehicular hardware architecture as the connected and autonomous industry develops and matures.

3) Interoperable: Mission and non-mission critical operating systems and applications need to run simultaneously on one ECU without interference. Additionally, a compartmentalization is needed to ensure that if one application should be compromised, all others will be unaffected.

4) Service-Oriented: The automotive industry is moving towards OEM future app stores for vehicle customization, requiring a multi-platform and multi-service approach with the ability to host multiple operating systems and services with secure separation between all resources, applications and operating systems.

5) Personalized: In-vehicle safety is critical for the automotive industry as additional levels of connectivity, especially for vehicular personalization, are added. Cybersecurity needs to serve as the foundational platform for added connectivity, services, and customization to create new markets and added revenue streams for OEMs.

Car manufacturers and Tier 1 suppliers have traditionally been turning to IT experts for in vehicle solutions. Unfortunately, IT cyber solutions may not be applicable to the automotive industry as vehicles are moving platforms with a finite set of messages compared to static computers with an infinite set of messages. In addition, cybersecurity within vehicles is an extension of safety, therefore the reliability of automotive cybersecurity solutions must be as close to 100% as possible. Furthermore, there is no room for false positives in such systems.

INNOVATION ADDRESSING AUTOMOTIVE CYBERSECURITY CHALLENGES

Overview - Expertise / Capabilities

GuardKnox Cyber Technologies Ltd[3]. is an automotive cybersecurity company which provides secured ECU's, domain controllers,

[3] www.guardknox.com

and gateways to the automotive industry with a hardware and software cybersecurity solution to protect vehicles including passenger cars, commercial vehicles, mass transportation and more. GuardKnox solutions are based on the secure network orchestrator or SNO™ technology[4] that can be centralized or localized. SNO features include in-vehicle network defence, high-performance gateway, application host domain computer, advanced body control and security for all external communication coming into the vehicle and from the vehicle externally.

Applications include:

- EV Charging ECU - a secure in-vehicle charging platform for V2G communications and
- Secured High-Performance ECU's - high-performance, scalable, and secured ECU's due to manufacturers' specification.

Applications include EV Charging ECU- providing a secure in-vehicle charging platform for V2G communications and Secured High-Performance ECU's- design high-performance, scalable and secured ECU's.

The GuardKnox platform also hosts service-oriented architecture, or SOA [5, 6] allows unified communication as well as access control and service level partitioning to ensure further levels of connectivity, customization, and additional revenue streams across the automotive value chain. These systems can also serve as platforms within the vehicle for secured data processing and storage.

The Communication Lockdown approach to cybersecurity is uniquely differentiated as it is not a learning or IT-based solution. Computer industry solutions such as Intrusion Detection/Intrusion Prevention Systems (IDS/IPS) and/or firewalls are seemingly obvious choices to protect computers-on-wheels, but they are ill-equipped to meet the cybersecurity needs of the automotive industry as they:

- Cannot offer a real-time/immediate response to new threats or morphing malware

[4] https://www.guardknox.com/products/

[5] GuardKnox Secure Hardware Architecture Patent "Hardware Components Configured for Secure Physical Separation of Communication Networks in A Vehicle and Methods of Use Thereof" http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fnetahtml%2FPTO%2Fsearch-bool.html&r=4&f=G&l=50&co1=AND&d=PTXT&s1=guardknox&OS=guardknox&RS=guardknox

[6] GuardKnox Services Oriented Architecture (SOA) Patent "Specially Programmed Computing Systems with Associated Devices Configured to Implement Centralized Services Based on Services Oriented Architecture and Methods of Use Thereof": http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnetahtml%2FPTO%2Fsrchnum.htm&r=1&f=G&l=50&s1=10,055,260.PN.&OS=PN/10,055,260&RS=PN/10,055,260

- Require time to analyse specific new threats and to develop and deploy software updates
- Do not offer support for an external Security Operations Center (SOC) to continuously monitor vehicles and communicate with owners and drivers

These challenges are overcome through a patented cybersecurity "lockdown" approach that is successfully used by Israel's F-35I and F-16I fighter jets and missile defence systems. By enforcing a formally verified and deterministic configuration of communication among the various networks of the vehicle, the Communication Lockdown methodology eliminates all known and unknown cybersecurity risks by approving or discarding all inbound and internal vehicle communications in real-time.

Table 1: Different Approaches to Automotive Cybersecurity

Capability	Communication Lockdown	Firewall	IDS/IPS	Anti-Virus
Security Mechanism	- Formally verifiable state machine - Agnostic to attacks - Certifiable (safety and security) - Approved configuration lockdown	- Static ruleset firewall - Needs updating as new attacks materialize	- Heuristic detection of attacks (anomalies) - Reliability can't be proven	- Local Anti-Virus - Signature updates required
Defense Capability	- All vehicle networks - Prevention on bit level	- Several car networks	- No prevention	- 1 ECU
Reliability	- 99.99999% - with deterministic mathematical model That can be verified, tested, and certified - Zero false positives	- Can be tested by automotive standards but can't be qualified	- 98% Detection rate - Up to 5% False positive rate	- Reliability can't be proven
Maintenance	- No cloud connectivity required - No on-going updates required	- Requires cloud connectivity and regular updates	- Requires online cloud connectivity and	- Updates for every change of the ECU

Capability	Communication Lockdown	Firewall	IDS/IPS	Anti-Virus
			continuous updates	
Physical Separation	- Hardware, software and firmware level separation between networks	- None	- None	- None
Integration	- Minimal integration - Transparent to other ECUs	- Requires integration into 3rd party (Tier1) ECU	- Requires integration into multiple 3rd party (Tier1) ECUs	- Requires integration into ECU and development environment
Scalability	- Full services and application secure hosting platform - Full support for virtualization and service oriented environment	- Fixed functionality - Requires integration into each environment	- Fixed functionality - Requires integration into each environment	- Need to recompile and re-certify the ECU
Cost Effective Hardware	- No need to modify vehicular hardware architecture for additional software extensions/applications	- None	- None	- None
Compliance to Standards	- Safety: ISO 26262 - Security: Common Criteria (ISO 15408)	- None	- None	- None
Physical Security	- Tamper proof: erases information upon tamper attempt	- None	- None	- None

Capability	Communication Lockdown	Firewall	IDS/IPS	Anti-Virus
Fit to Automotive value chain	- Full fit to tiered hardware value chain, no integration	- Requires software integration	- Extensive integration	- Extensive integration

Target Areas for the Communication Lockdown based technology include:

1) Automotive OEMs: where the systems would be provided directly to the OEM, where connected vehicles need security, without compromising the safety and integrity of the vehicle. Specific product implemented depends on the OEM's concern:
   a) Preserving the holistic security for the vehicle and all of its components.
   b) Dedicated ECUs or a single interface, such as telematics or infotainment.

The technology can be implemented during production or retrofitted after production or in the aftermarket.

2) Tier 1 Suppliers: Tier 1 suppliers are responsible for building the in-vehicle ECUs. The ECUs may possess many cybersecurity vulnerabilities due to the number of networks and other ECUs communicating within a vehicle. By implementing Communication Lockdown technology into their ECUs, these vulnerabilities can be addressed. This applies to aftermarket products that need protection and seamless integration as well.

3) Telematics Providers - Fleet Tracking: Telematics and in particular trucking telematics is considered one of the biggest growth industries. Telematics and fleet management solutions enable commercial trucking OEMs and large fleets to monitor and better understand their usage through location providing services. These behave the same as a car ECU but are referred to as a TGU, or Telematics Gateway Unit within a commercial vehicle. A Communication Lockdown-based solution protects the TGU and ensures both secure functioning of the vehicle from any external cyber threats as well as securing the data obtained from the device.

Solution

Application Illustrating Innovation – "Communication Lockdown"

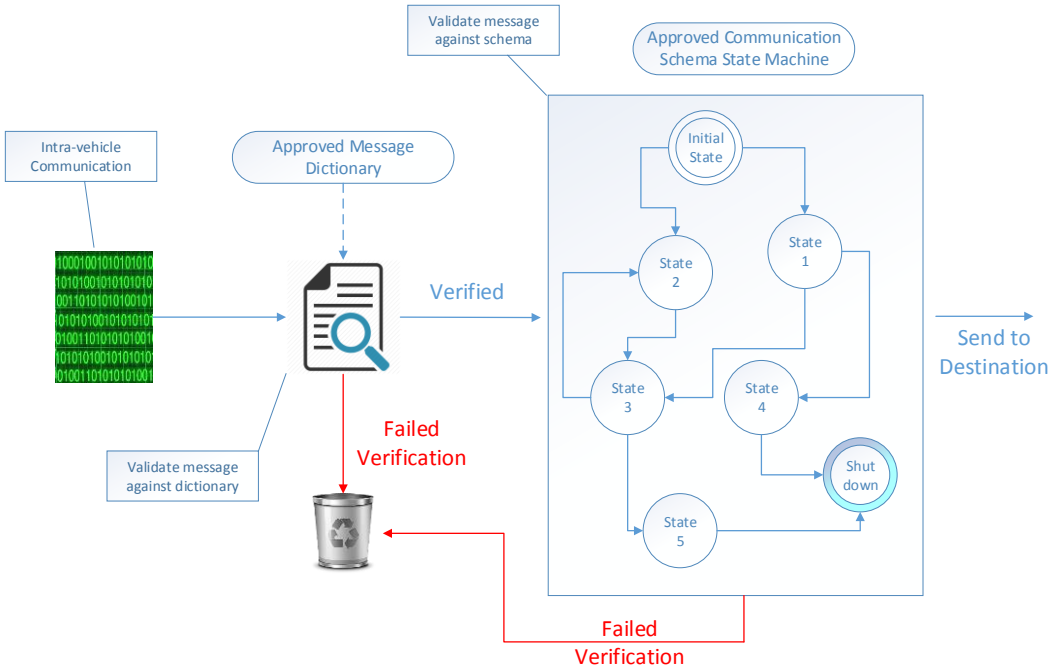


Figure 1: Communication Lockdown with State Machine

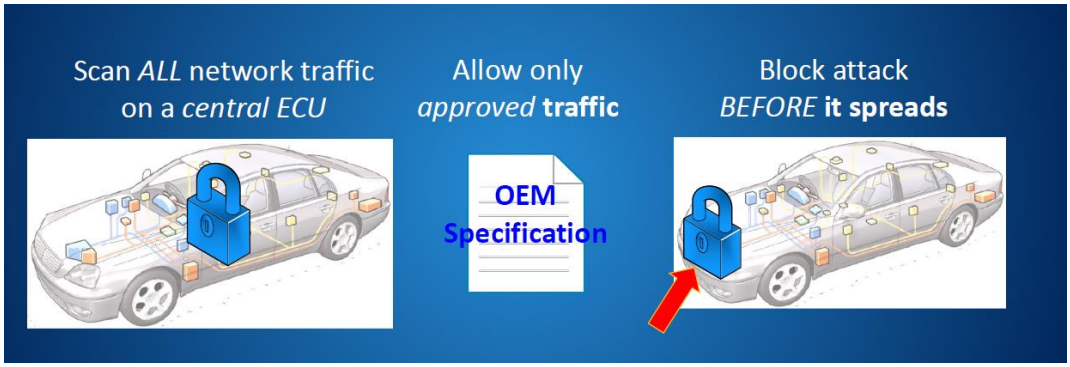


Figure 2: Communication Lockdown

Description of Methodology "Communication Lockdown" with Network Orchestration

The Communication Lockdown methodology delivers an innovative approach to automotive security, permitting authorized communication while being impervious to any type of inappropriate transmission, including all cyber-attacks. This includes but is not limited to DoS attacks.

The methodology detects and prevents the injection and the spread of malicious

messages between the various ECUs used to control the vehicle. All incoming messages are inspected, and only approved/legal messages can continue to their destination. Since communication lockdown looks at the approved frequency and the size of the messages, this effectively limits the case of bus overload. This is furthermore achieved even more effectively in hardware using field-programmable gate array (FPGA) logic since it is able to deal with a higher bandwidth communication than solely reaching it in software. All cyberattack attempts—in which illegal or improper messages are discarded—can be logged and reported over a wireless vehicular communication channel to a remote OEM SOC for further technical and statistical analysis, including fleet information, geographies and trends.

Intrinsic to the Communication Lockdown methodology is the ability to use the OEM Technical Specifications, specifically the communication matrix, where the bus message database and the functional specifications are used, to create a communication schema that models the proper behavior of all vehicular data.

The Communication Lockdown methodology is agnostic to attacks since it does not look for them. Instead it only models the "correct" behavior. In this approach of not looking for attacks from a defense methodology standpoint, you do not care about the incoming attacks since they are not being looked for. In Communication Lockdown the communication is efficiently modeled and verified to comply with the vehicle specification. This enables full autonomy after installation and operates deterministically without the need for frequent software or firmware updates—unlike Intrusion Detection/Intrusion Prevention Systems (IDS/IPS) or firewalls.

Three Layers of Communication Security

The effectiveness of the Communication Lockdown methodology is based on the patented ability to inspect and verify messages on multiple layers. This ensures that if an external message from the vehicle's ecosystem is compromised, the internal vehicle network remains fully protected from the spread or propagation of malicious code.

All incoming messages are inspected on three layers:

Routing Layer

- The origin and destination of each incoming message (type) is checked by the Communication Lockdown™ schema to ensure that they are permissible or "legal". For example, messages from the infotainment subsystem to the powertrain components (steering, brakes, etc.) are prohibited and would therefore be discarded.

Content Layer

- The content of each incoming message is checked down to the bit level for compliance with the permissible format as defined in the OEM's Technical Specifications. Messages that do not conform to the defined format are dropped.

Contextual Layer

- The content of each incoming message is checked for legitimacy in the specific functional state of the vehicle, subsystem, ECU, etc. Messages from specific origins to specific destinations are permitted or discarded depending on the contextual/functional state of the vehicle. For example, messages received from the OBD-II maintenance connector during the vehicle movement on the road (speed > X Kmph) will be discarded.

Communication Lockdown Methodology Unique Benefits:

DETERMINSTIC

The Communication Lockdown approach is a fully deterministic cyber security methodology. The closed-system approach is not to look for attacks, but rather to ensure that the vehicle continues to function in the way it was designed.

UPDATEABLE

Using automatic tools to create layered protection, a fully deterministic, yet updateable mathematical model that can be formally verified is generated.

FORMALLY VERIFIED

On three different layers, down to the bit level. Additionally, open fields are also 'locked down' to ensure stringent security.

FINITE STATE-MACHINE

This model includes a state machine, which enforces predetermined states, with a dedicated ruleset generation tool. Only allowed communications, as detailed by OEM technical specifications and bus network communication matrices, are approved.

STAND ALONE SOLUTION

There is no need for cloud connectivity nor for ongoing updates. No malware can sneak in and corrupt the safety requirements of the vehicle. The Communication Lockdown methodology delivers the requirements of the Safety Critical Subsystem of the connected car.

SECURED CLOUD CONNECTIVITY

The Communication Lockdown model behaves as a secured landing point within the vehicle for cloud connectivity which enables secured OTA and data transfer, among other things. The mechanism supports mutual authentication and encryption between the backend, the cloud and the vehicle, therefore enabling secured cloud connectivity when needed as opposed to resource-intensive and vulnerable continuous connectivity.

CAN BE INTEGRATED WITH ANY SOC

Supports any SOC to monitor, log and report any and all activities.

Software - Service Oriented Architecture (SOA)

SOA partitioning [7] utilizes the patented architecture allowing for unified communication as well as access control and service level partitioning. Using a separation kernel allows for abstraction and concealment of communications across the platform, this allows for simplified and transparent interface to service providers. Service providers include but not limited to:

- Another process
- Different partition
- Reside on a different operating system
- Reside on a different processor

Furthermore, CORBA brokers may be used in order to standardize service access across the platform.

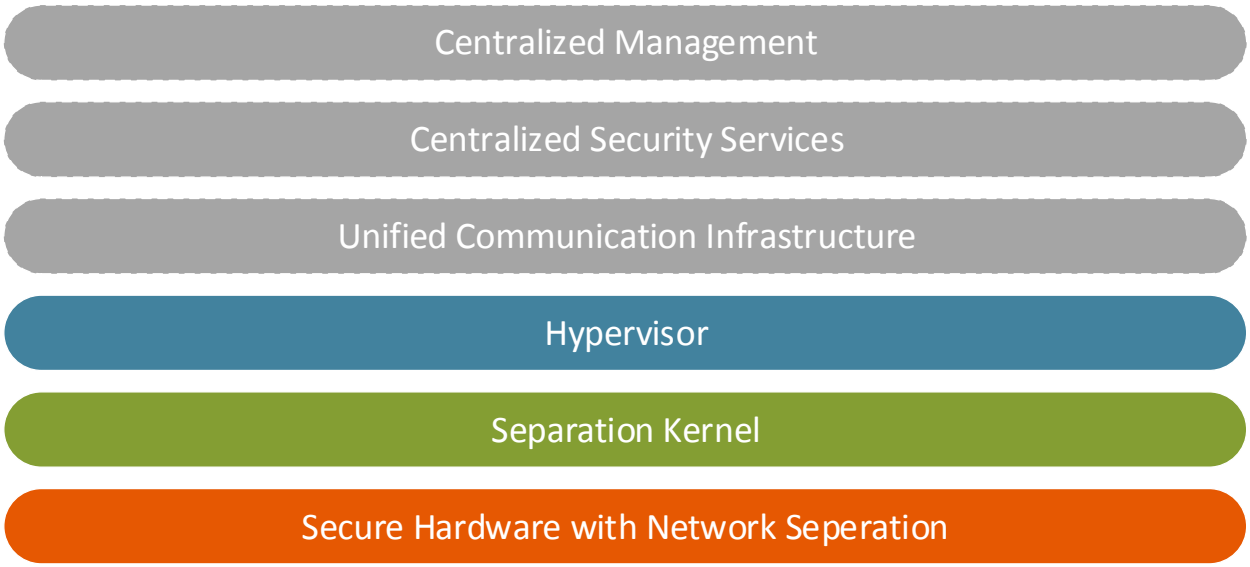


Figure 3: Services Oriented Architecture

[7] GuardKnox Distributed SOA Patent "Specially Programmed Computing Systems with Associated Devices Configured to Implement Centralized Services ECU Based on Services Oriented Architecture and Methods of Use Thereof "http://appft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fnetahtml%2FPTO%2Fsearch-bool.html&r=3&f=G&l=50&co1=AND&d=PG01&s1=GuardKnox&OS=GuardKnox&RS=GuardKnox"

SOA Unique Benefits[8]:

REUSABLE CYBERSECURITY SERVICES

Including firewalls, remote server management, cryptography or the Communication Lockdown framework

INCREASED CONNECTIVITY

Hosting services and downloadable applications for customization

NEW REVENUE STREAMS

Supporting an app store [9] for downloadable personalized apps and features

SCALABILITY

Flexible hardware architecture for future unforeseen needs and data requirements

INTEROPERABILITY

Ability to host and communicate with all operating systems, whether mission critical or not and containing the failure of a single app/service so that others are unaffected.

Hardware – physical separation of safety critical networks in distributed environment

At the core of the hardware architecture is the physical separation of critical networks by isolating the communication interfaces. In order to pass data to one another, the communication interfaces have to go through a security mechanism. The platform can ensure data paths are enforced by physical means and not only by traditional software permissions. Custom IP cores can also be used and placed along those paths to further boost security assurance.

Distributed Systems: The patent on distributed systems ensures that multiple units (SNO's) within a vehicle work together in a cohesive manner in which they are not independent entities. This therefore enables multiple lockdown devices to operate together in a vehicle (e.g., internal and external ones). Since each device is only seeing a part of the network traffic, they can cooperate and exchange metadata about the traffic they see and approved/blocked. Thus, making the overall model more accurate.

[8] GuardKnox Services Oriented Architecture (SOA)https://www.guardknox.com/services-oriented-architecture-automotive-services/

[9] Goldberg, Jillian (2017,11). Turning Drivers to Subscribers. https://blog.guardknox.com/connected-vehicle-vulnerabilities-turning-drivers-to-subscribers

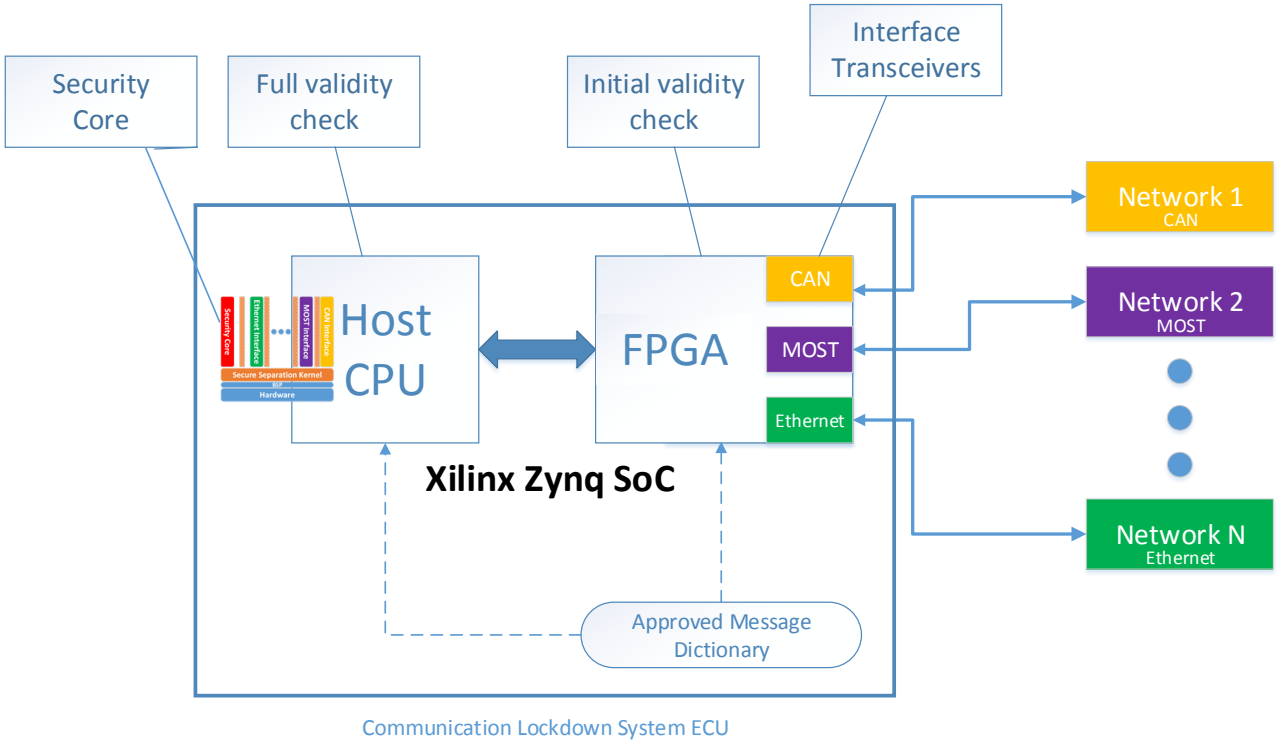


Figure 4: Communication Lockdown System ECU

Secure Updates

Secure delivery, authentication of sender and verification of data integrity, bug fixes, improved user experience as well as new functionality can be introduced safely and securely to address current and future needs, increasing the value of the product to the customer, while reducing development and integration costs.

The firmware image (both hardware and software) can be encrypted and signed at the vendor's site and delivered to devices securely at all times and by any means.

Xilinx Zynq SoC decrypts and authenticates the firmware image prior or during first boot ensuring no unauthorized hardware configuration or software can be loaded and executed on the device.

Role of Data Analytics and Machine Learning

Utilizing FPGA has other performance related benefits as well, advance Artificial Intelligence (AI) and analytics can be implemented in hardware.

Preprocessing can be done on the endpoint (in this case the G platform SNO™) in order to save bandwidth by offloading and distilling data to only what is necessary for the cloud application.

The platform also enables secure end-to-end connection to a cloud infrastructure. Afterwards, the cloud can be used for use cases such:

- Predictive Maintenance
- Fleet health monitoring

Xilinx Programmable Technology

Flexibility and scalability are key advantages that programmable technology provides. These solutions support a wide variety of industry standard interfaces for interoperability with other devices, including virtually any type and combination of interfaces through use of the programmable fabric and configurable IO. In the context of security and flexibility, security accelerators can be implemented in the programmable logic. Cryptography can be managed with keys embedded in hardware (also creating secure memory from FPGA) and further enhanced using the integrated Physical Unclonable Function (PUF) technology in the Zynq MPSoC family of devices [10] .From an isolation point of view, true hardware separation is utilized – where the communication interfaces can be passed through security mechanism(s), such as watchdogs, isolation of data and control paths and other mechanisms in order to pass data to one another.

Additionally, a certifiable methodology for isolation of separate areas on a single device can be achieved through use of Isolation Design Flow (IDF) and Vivado® Isolation Verifier (VIV) / Isolation Verification Tools (IVT). Designs placed into these regions are physically isolated. The areas can be changed at any time without impacting other isolated regions.

System responsibility can be distributed between the processing system (i.e., software) and the programmable logic (i.e., hardware). Unique to programmable technology, both the software and the hardware can be reconfigured, either in total or partially, utilizing the reconfigurable nature of the device. This essentially provides new functionality and updates to existing functionality via OTA SW and OTA Silicon, including systems already deployed in-field.

Application	Standard
Automotive	ISO 26262
Industrial and Medical	IEC 61508, IEC 62061 and IEC 13849
Aerospace & Defense	DO-254/DO178b

Table 2: Functional Safety Standards

Functional Safety Standards Supported

Security and Functional Safety should be designed in from the start.

[10] Physical Unclonable Function (PUF) technology in the Zynq MPSoC family of devices: https://scholar.uwindsor.ca/cgi/viewcontent.cgi?article=8596&context=etd

Valuable System Features

The Zynq MP SoC architecture (Figure 5) includes a high bandwidth interconnect between the PS, PL allows for tight integration and flexible partitioning between both hardware and software.

Optimization of algorithms running on the processing system can be achieved by offloading and accelerating the algorithms using the inherent parallelism of the programmable logic fabric while providing the processing power scalability and flexibility essential for these programmable platforms across the Zynq SOC device family.

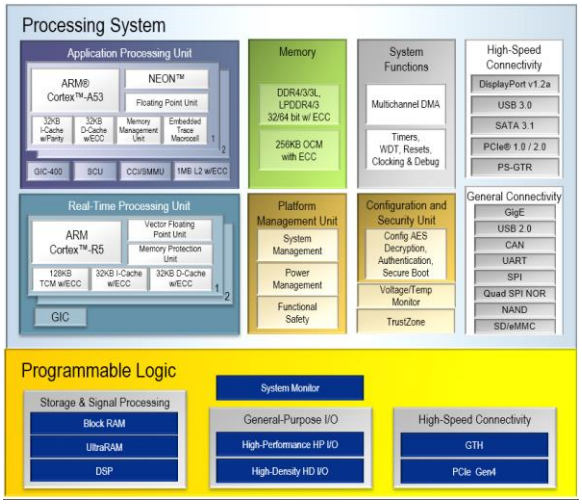


Figure 5: Zynq US+ MPSoC Partitioning: Processing System and Programable Logic

Scalability can be applied where required.

Extending this compute capability, is the focus on Machine Learning and AI Engines[11]. Optimized resource and power efficient neural networks can be implemented in the Programmable Logic to augment compute in Automotive solutions[12].

Future Direction

The Communication Lockdown methodology as described in this article extends beyond the automotive industry and is applicable to any kind of closed system. This includes but is not limited to

- Industry 4.0
- Smart-grid applications
- Critical infrastructure
- Medical devices.

In addition, GuardKnox is an active member of the *ISO/SAE joint working group, ISO-TC22-SC32-WG11.* This working group's mission is to create a new standard, ISO-SAE 21434: Road vehicles – Cybersecurity Engineering[13].

[11] Xilinx AI Engines and Their Applications https://www.xilinx.com/support/documentation/white_papers/wp506-ai-engine.pdf

[12] Power efficient neural networks Augmenting compute in Automotive solutions: https://www.itu.int/en/journal/001/Documents/itu2017-2.pdf

[13] Goldberg, Jillian (2018, 04) Setting the Standard for Automotive Cybersecurity https://blog.guardknox.com/setting-standard-automotive-cyber-security

- Return to the beginning of this article
- Return to the Table of Contents

Driving Membership through SymbIOTic Connections

Quantifying return on investment (ROI) in terms of a membership organization or association isn't easy. For members of the Industrial Internet Consortium (IIC), now incorporating OpenFog, ROI lies in active participation, the overall experience and a common thread that rises to the surface—"connection." The industrial internet is all about connecting industries to the internet and bringing together the organizations and technologies necessary to accelerate the growth of the industrial internet. Membership empowers this ecosystem to grow.

Woven throughout every working group and task group are activities designed to connect across groups – vertical industry-focused groups to testbeds, liaison organizations to verticals, technology to business strategies, architecture to use cases, technology challenges to special interest groups and end-users to solutions. Forward progress is embodied through delivering technical insights, actionable intelligence and solutions to real-world problems. In order to do this, the IIC relies on initiatives such as those described below.

Combining Forces with OpenFog

On January 31, 2019, the two largest and most influential international consortia in Industrial IoT, fog and edge computing joined forces to work together under the IIC umbrella to drive the momentum of the industrial internet, including the development and promotion of industry guidance and best practices for fog and edge computing. The combined organization offers greater influence to members, more clarity to the market, and a lower-risk path to the future for end users. We will be the center of gravity for the future of Industrial IoT systems across industry verticals.

End User Engagement

In Q3 2018, an End User Leadership Council was formed to share end user insights in implementation challenges and set the vision for the industry. Founding representatives from Boeing, B&R Automation, Bosch Rexroth, Church & Dwight, Deere & Co. and TRUMPF met again in Q1 2019 to continue discussions with IIC leaders around key challenges and opportunities to test IIC-member driven solutions within their own factories.

End User Leadership Council participants may seize opportunities to "test drive a testbed" to apply an IIC testbed to new use cases, showcase their organization as an early adopter and solve an immediate problem with a practical IoT implementation directly on their site. These sorts of

cross-pollination activities being furthered by the End User Leadership Council will benefit end-users and members directly, as well as the industry at large.

Connecting via Industry Events

The IIC's Global Event Series continues to offer world-class knowledge learning opportunities and are often aligned with vertical industries, featuring speakers from around the globe. A Building Intelligent Infrastructures Forum hosted by SAS on February 15, 2019 at their Executive Briefing Center in Cary, NC showcased Intelligent Transport Systems and Smart Cities, with experts and thought leaders from SAS, Cisco, RIoT, Building Clarity, TCS, LHP Engineering Solutions, RTI, Volvo Trucks, NetApp, Itron and the Town of Cary, NC. Presentations and videos from this event may be viewed here. The next Global Event Series will be hosted in Cork, Ireland on May 23, 2019.

In addition to the Global Event Series, IIC members can be found at other participating events such as Hannover Messe, the Intelligent Transport Systems Forum at IoT World Santa Clara and IoT Solutions World Congress in Barcelona.

New Liaisons

The IIC collaborates with global standards development organizations, technology and industry-focused consortia, open source communities, regional organizations and government agencies in a variety of ways including:

- Technical exchange such as harmonization of IIoT architectures and exchange of use cases, best practices, results, technical reports and frameworks
- Joint marketing activities
- Joint workshops
- Joint publications

Since the last Journal of Innovation publication, IIC has added a number of new Liaison Members, bringing the total to well over 40 liaisons.

- The International Data Spaces Association will collaborate with the IIC to improve trustworthiness of IIoT data
- Linaro, an Open-source community with an important role in IIoT market growth will work with IIC to promote the digital economy and preventing industry fragmentation
- GlobalPlatform will work with the IIC to align efforts to maximize interoperability, portability, security and privacy for the industrial Internet.
- The Wi-SUN Alliance, a global association driving the propagation of interoperable wireless solutions for use in smart cities and smart utilities will work with the IIC to maximize interoperability, portability, security and privacy for the industrial Internet.
- A formal relationship with the OPC Foundation is a natural fit for the IIC as the ***OPC Unified Architecture standard*** is widely used in several IIC testbeds.

As an example of testbed progress, the IIC's Track & Trace Testbed has generated requirements for an IIoT standard which would improve the sharing of data from IIoT sensors. These requirements were submitted to the Object Management Group (OMG) who has a nearly 30-year history creating and maintaining IT standards. OMG has begun the process to develop this as a new IIoT standard. To explore the list of our Liaisons and the joint work in process, please refer to the Liaison Working Group activities page.

New Publication

As organizations connect their systems to the internet, they become vulnerable to new threats, and they are rightly concerned with security. Addressing these concerns requires investment, but determining investment focus and amount is a difficult business decision. The Security Working Group published the Security Maturity Model (SMM) Practitioner's Guide, which provides detailed actionable guidance enabling IoT stakeholders to assess and manage the security maturity of IoT systems. Corresponding with the publication of the SMM Practitioner's Guide is an update to the IoT SMM: Description and Intended Use White Paper, which provides an introduction to the concepts and approach of the SMM.

The SMM Practitioner's Guide builds on concepts identified in the groundbreaking IIC Industrial Internet Security Framework published in 2016 and defines levels of security maturity for a company to achieve based on its security goals and objectives, as well as its appetite for risk. Organizations may improve their security state by making continued security assessments and improvements over time, up to their required level. Additional information and resources associated with the SMM Practitioner's Guide publication may be researched here. Please also join us for a Webinar, ***"Get a True Sense of Security Maturity"*** presented by the co-authors -- available live, or on demand following the event.

Collaboration Tools

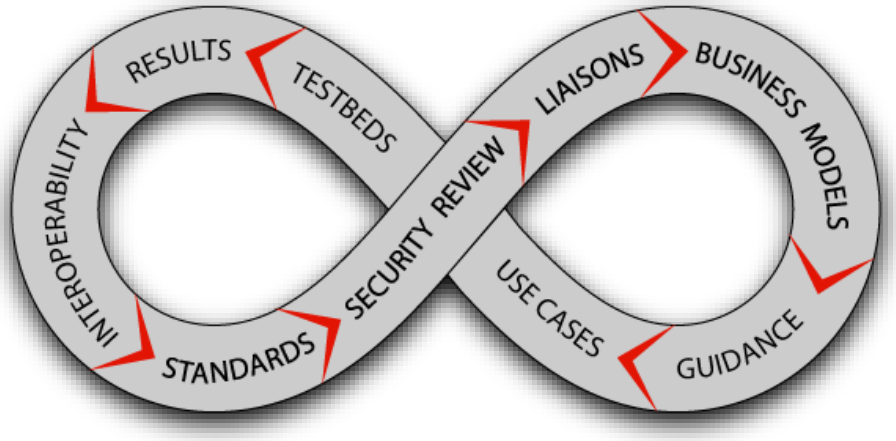
Resource Hub Enhancements

We are committed to providing the best experience for our member organizations, both in providing additional exposure to their products, brands and collaborative bodies of work. In the second half of 2018, we launched the IIC Resource Hub which represents the body of knowledge and activities driven by IIC members converged into a public resource that arms the world with new tools in the IIoT strategic arsenal. The actionable intelligence and interfaces continue to expand and become more robust every quarter as we add new documents, enhance the explorers and provide insights. Be sure to take advantage of all the Resource Hub has to offer.

IIC Connect

The well-established "IIC Connect" program is like "match.com" for member companies. The concept is simple: you enter your profile and a quick description of what might interest you. Then,

everyone looks at the profiles and requests meetings. You accept the ones you like; software does some automated scheduling and you have 20-minute networking sessions. It's fast, easy, efficient and fun!



The Industrial Internet Consortium ecosystem is a continuous flow of inter-related activities

Since connecting internet to Industry is the soul of the next Industrial Revolution, the Industrial Internet Consortium, now incorporating OpenFog, continues to evolve very quickly. All of the above brings us back to the value of connections. When you put this all together, a very "SymbIOTic" ecosystem emerges. From end users, to platform vendors, OEM's and suppliers, systems integrators to consultants and thought leaders – all members play an important role.

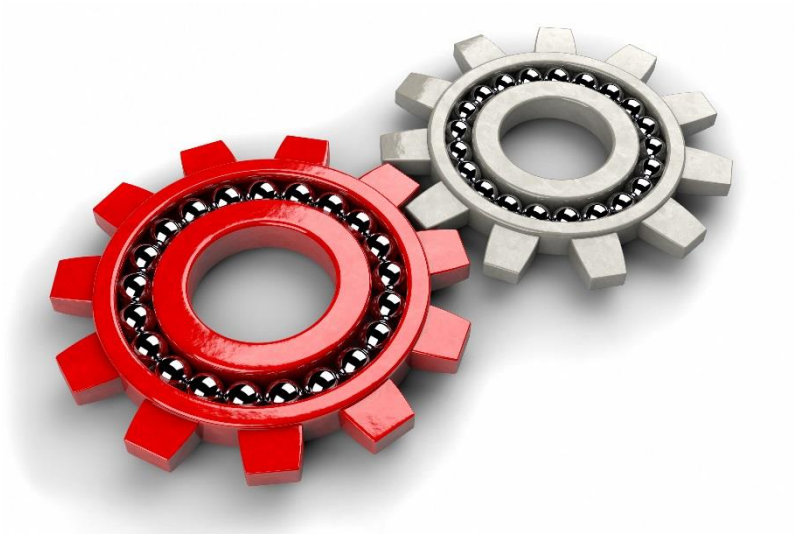
As one member eloquently explained, "Perhaps you don't need cheap sensors, but how about analytics? Middleware? Deep learning? Ad hoc networking? Security? Services? Tools? It's pretty clear: every IIoT company or technology end-user on the planet should be able to leverage the world's largest IIoT ecosystem at the Industrial Internet Consortium. Just understanding the amazing mix is great education on the breadth and depth of the IIoT. Every member has the opportunity to make a difference and grow their business."[1] This could be you! We welcome and encourage new member participation and invite you to join us.

[1] Interested in learning more about how members monetize their IIC Membership in this Webinar.

- Return to the beginning of this article
- Return to the Table of Contents

The views expressed in the *IIC Journal of Innovation* are the contributing authors' views and do not necessarily represent the views of their respective employers nor those of the Industrial Internet Consortium, now incorporating OpenFog.

Eleventh Edition of the IIC Journal of Innovation
Coming in June 2019
Theme: Artificial Intelligence



Things are Coming Together. ™

March 2019