Lecture Notes in Electrical Engineering 670

Sophia Keil
Rainer Lasch
Fabian Lindner
Jacob Lohmer *Editors*

Digital Transformation in Semiconductor Manufacturing

Proceedings of the 1st and 2nd European Advances in Digital Transformation Conference, EADTC 2018, Zittau, Germany and EADTC 2019, Milan, Italy

Springer

Lecture Notes in Electrical Engineering

Volume 670

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Associate Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Executive Editor (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, SCOPUS, MetaPress, Web of Science and Springerlink ****

More information about this series at http://www.springer.com/series/7818

Sophia Keil · Rainer Lasch ·
Fabian Lindner · Jacob Lohmer
Editors

Digital Transformation in Semiconductor Manufacturing

Proceedings of the 1st and 2nd European Advances in Digital Transformation Conference, EADTC 2018, Zittau, Germany and EADTC 2019, Milan, Italy

OPEN ACCES

Springer

Editors
Sophia Keil
Hochschule Zittau/Görlitz -
University of Applied Sciences
Zittau, Germany

Fabian Lindner
Hochschule Zittau/Görlitz -
University of Applied Sciences
Zittau, Germany

Rainer Lasch
Technische Universität Dresden
Dresden, Germany

Jacob Lohmer
Technische Universität Dresden
Dresden, Germany

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-3-030-48601-3 ISBN 978-3-030-48602-0 (eBook)
https://doi.org/10.1007/978-3-030-48602-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This proceedings book covers the 1st and the 2nd European Advances in Digital Transformation Conference (EADTC), which took place in Zittau, Germany, on November 29, 2018, and Milan, Italy, on April 4, 2019. The EADTC is an international and interdisciplinary conference covering research and development in the context of the digital transformation of the European industry and society with a focus on the semiconductor industry, representing an outcome of the European research project "Integrated Development 4.0 (iDev40)." 40 European partners are working on disruptive solutions toward strengthening the electronics components and systems industry in Europe. By developing suitable advancements in digital technology, closely interlinking development processes, logistics, and manufacturing, a significant speedup in time to market is the aim of iDev40. On the initiative of Josef Moser and Sabine Allmayer (Infineon Technologies Austria AG), as well as Sophia Keil (Zittau/Görlitz University of Applied Sciences), the EADTC has been established with the primary objective to create a joint platform for dissemination and discussion of research results for the European ECS industry and related research organizations. This event will foster the dialogue and exchange of ideas between practitioners and academics in the context of digital transformation.

This proceedings book includes the first results of the work in the project with 13 contributions from industry and research. This volume of proceedings is aimed at interested practitioners and researchers in the ECS industry, scientists, and people eager to enhance their knowledge. After the initial call for papers, 16 contributions were submitted to the conference, all of which were subjected to a process of single-blind review by two individual reviewers. Thirteen papers were recommended for publication for the conference proceedings that will be published by Springer in its Lecture Notes in Electrical Engineering (LNEE) series.

Thematically, the papers published in these proceedings will be clustered in four segments: semiconductor supply chain management, semiconductor operations, human factors in semiconductor manufacturing, and smart software and hardware solutions for semiconductor manufacturing. The papers focus on the several specific topics, including virtual factory clusters, real-time planning and control, proof of qualification and reliability, smart training, skill and knowledge management, smart collaboration, smart human–machine interaction, and KPI enhancement as well as human factors.

We want to thank all iDev40 partners that participated in any way in the EADTC conferences in 2018 and 2019 and made it successful events, but especially Sabine Allmayer and Josef Moser for their outstanding engagement to strengthen the industry perspective and to assure the practical relevance of the conferences. We thank all the reviewers for reviewing the papers submitted to the conferences. Besides, we are grateful for the support of the hosts Zittau/Görlitz University of Applied Sciences and Università degli Studi di Milano-Bicocca. Finally, we are grateful to the speakers, participants, referees, organizers, and sponsors of the iDev40 consortium, who made a significant contribution to the success of the conferences. With this in mind, we wish iDev40 and the EADTC all the best and are looking forward to successfully continuing our undertaking in jointly strengthening expertise in digital transformation in Europe.

The project iDev40 has received funding from the ECSEL Joint Undertaking under grant agreement No 783163. The JU receives support from the European Union’s Horizon 2020 research and innovation program. It is co-funded by the consortium members, grants from Austria, Germany, Belgium, Italy, Spain and Romania. It is coordinated by Infineon Technologies Austria AG.

This proceedings book is sponsored by:

ECSEL Joint Undertaking

iDev40

Diese Maßnahme wird mitfinanziert durch Steuermittel auf Grundlage des von den Abgeordneten des Sächsischen Landtags beschlossenen Haushaltes.

SPONSORED BY THE Federal Ministry of Education and Research

MINISTERO DELL' ISTRUZIONE, DELL'UNIVERSITÀ E DELLA RICERCA

FFG Promoting Innovation.

March 2020

Sophia Keil
Rainer Lasch
Fabian Lindner
Jacob Lohmer

Contents

Semiconductor Supply Chain Management

A Holistic Digital Twin Based on Semantic Web Technologies to Accelerate Digitalization

Patrick Moder[1,2](✉), Hans Ehm[1], and Eva Jofer[1,3]

[1] Infineon Technologies AG, Corporate Supply Chain Engineering Innovation (IFAG CSC E IN), Am Campeon 1-15, 85579 Neubiberg, Germany
patrick.moder@infineon.com

[2] Department of Mechanical Engineering, Institute of Automation and Information Systems, Technical University of Munich, Boltzmannstr. 15, 85748 Garching, Germany

[3] Faculty of Electrical Engineering and Computer Science, Technische Hochschule Ingolstadt, Esplanade 10, 85049 Ingolstadt, Germany

Abstract. This proposed research is concerned with the development of a Semantic Web representation of supply chains deploying and manufacturing semiconductors. The so-called digital twin is a prototype for virtualizing the electronic component supplier (ECS) value chain, aiming for a holistic digitalization across the entire product life cycle. With its unique advantages, Semantic Web will contribute to overcoming current Big Data issues and to getting one step closer to smart factories. This includes enhanced smart collaboration between both business and project partners. Furthermore, it is shown in first business use cases how the developed ontology is supporting selected applications to manage data more efficiently and to build a comprehensive knowledge representation.

Keywords: Semantic Web Technologies · Semiconductor Development · Digitalization · Product Life Cycle · Internet of Things · Big Data · Digital Twin · Smart Factory

1 Context and Motivation

The emerging fourth industrial revolution towards both Industry 4.0 and smart networks includes the progressive approach of Internet of Things (IoT) as one main paradigm. One basic principle behind IoT is the connection and remote interaction between (smart) objects, called things. This mainly includes numerous steps of data processing and hence for the initial data generation it requires integrable sensor devices (Andriopoulou et al. 2017). Improvements during the last years let these objects become smaller and enable enhanced collaboration of such embedded devices. In this regard, things include both physical and virtual entities that need distinct identification for efficient interaction and to ensure real time communication (Siozios et al. 2017). It is important to maintain and improve well-defined environments for uninterrupted communication and consequently high performance. Furthermore, a growing number of domains trying to be part of an interconnected world leads to the need for flexible,

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 3–13, 2020.
https://doi.org/10.1007/978-3-030-48602-0_1

scalable and versatile approaches to overcome emerging complexity issues (Mujica et al. 2017). This also holds true for automation approaches that make use of embedded sensors and other smart devices (Rumpl et al. 2002). Furthermore, involving data plays a central role in the future. This is already the case in consumer sectors and is highly emerging in the industrial environment as well. Consequently, it leads to a data boom that entails critical impact on industry due to Big Data issues, namely volume, velocity, veracity and variety, among others (Hofmann 2017).

The digital transformation of business and society offers enormous growth potential for Europe. Digitalization is both an enabler and driver of fundamental disruptive business innovations. European industries can build on their strengths in advanced digital technologies and strong presence in traditional sectors to seize the range of opportunities offered by technologies such as the Internet of Things, Big Data approaches, advanced manufacturing, blockchain technologies and artificial intelligence. Integrated Development 4.0 (iDev40) as a European Union funded project is focusing on digitization of integrated product development processes as well as value chains. Nevertheless, despite large initiatives and associations like Industrial Data Space (Fraunhofer IDS) and Platform Industrie 4.0, digitalization is slowly adopted by European companies. One reason is, among others, that humans understand the digitalization papers generated in the European research community but computers do not. Furthermore, one is observing slow broad-scale adoption in the industrial space whereas digitalization is gaining more attention in other environments. This is among others due to persistent uncertainty about specific business and application cases. Additionally, Big Data issues named above are not yet overcome comprehensively. The holistic digital twin based on Semantic Web technologies is an approach to bridge the gap between the physical and digital world in order to accelerate digitalization. This basically leads to IT systems being enabled to process information from web sites and other data resources in order to recognize relationships and dependencies between pieces of data. Hence, one is able to make implicit knowledge explicit and link data from different data resources effectively. Although this was very successful for search engines (Google, e.g.) and in social networks (Facebook, e.g.), it has not been applied in the industrial space – also known as B2B environment – in a large scale yet. Summarizing, one requires a semantic annotation of the web and existing relationships for digitalization progress in industry environments (Baumgärtel et al. 2018).

2 The Semantic Web

In order to overcome current hurdles described above, we propose the application of Semantic Web technologies. Semantic Web provides a powerful toolset to define and maintain a controlled vocabulary of processes, roles, objects and interactions. The Semantic Web expands on the current World Wide Web (WWW) framework. Linked open data allows for data to be read and interpreted by both humans and machines, consequently better enabling cooperation between computers and humans. While the traditional WWW links information via human-readable documents encoded in HTML, Semantic Web links information on the data level using the Resource Description Framework (RDF). Therefore, it is machine-understandable and -interpretable and hence improves data analysis possibilities as well as knowledge extraction (Dustdar and Falchuk 2006; Mane et al. 2019). First introduced in 2001 by Tim Berners-Lee, the dual intelligibility between humans and machines provides a tremendous opportunity as an enabler for emerging technologies such as blockchain, Big Data analytics and automation. Semantic Web is also a big step towards advancing several of the design principles of digitalization and Industry 4.0.

The various building blocks of Semantic Web are depicted in Fig. 1, showing the Semantic Web Stack that is created by Berners-Lee and builds on standards of the traditional WWW. Semantic Web uses RDF to represent semantic knowledge, thus allowing to model resources and the properties and linkages that are defining them. Data is expressed in the form of triples, containing a subject, predicate and an object. A subject, also known as a resource is the object or thing of interest. The predicate, also called property, relates the subject and the object with an attribute. This is done with either an object property or data property. An object property is used to relate a resource with another resource. Similarly, a data property is used to relate a resource with a piece of literal data. There is a unique identifier (URI) for each resource and property, and the URI refers to the address where the defining ontology is stored. In an ontology, a resource corresponds to a class or sub-class, which holds a set of elements. Furthermore, instances are individual objects belonging to a class. The Resource

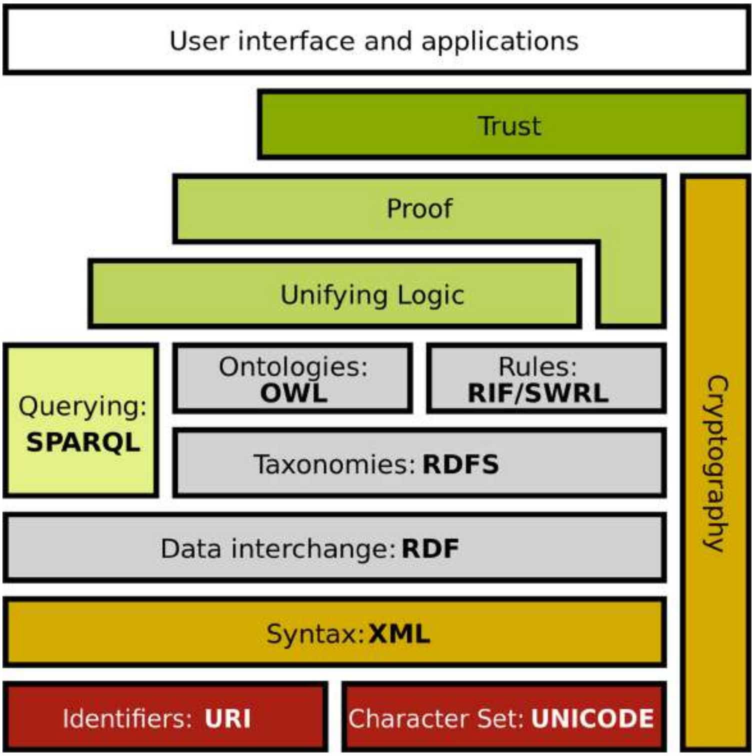

Fig. 1. The semantic web stack, referring to Berners-Lee 2000.

Description Framework Schema (RDFS) goes beyond RDF, by allowing for property and class vocabularies and the generation of hierarchies. A further extension of RDF and RDFS is the Web Ontology Language (OWL). OWL expresses the RDF construct and allows for even more descriptive properties, relationships and class descriptions. OWL is developed and supported as the standard language by the World Wide Web Consortium (W3C) that additionally defines a broad set of standards in the Semantic Web community (Berners-Lee et al. 2001).

Semantic Web uses ontologies and languages such as OWL for class, property and individual definition as well as determining their relationships. When data is represented in an ontology, this facilitates information sharing and collaboration between machines and humans. A common understanding through ontologies helps to reduce potential misunderstanding. By encoding information using an ontology language, the implicit specification of knowledge can be interpreted, extracted and made explicit. Thus, it enables advanced functions being performed on the data. Protégé is one of the primary ontology tools in use. A compatible software like Protégé allows information to be embedded without the need of specific expert knowledge initially (Lacy 2006).

While developers of an ontology can define certain relationships and properties related to classes, reasoners provide the ability to make additional inferences, thereby providing explicit deductions from implicit information. These deductions can be thoroughly explained and reviewed by tracing each step and accompanying inference rules. Another feature of reasoners is the ability to review an ontology for – e.g. logical – inconsistencies. Furthermore, data stored within RDF ontologies can be extracted and manipulated using SPARQL. This is accomplished by information selection based on graph assemblies, thus providing filter strategies based on logical comparisons. Reasoners and SPARQL queries are means that leverage the possibilities of a machine-readable form. Conversely, the WebVOWL tool caters to the human operator by providing an interactive visualization of ontologies. An ontology with all its entities can be graphically represented by implementing the Visual Notation for OWL Ontologies (VOWL). The application is interactive, allowing users to customize the spatial arrangement of classes (i.e. nodes) through the Pick & Pin mode. The customized ontology visualization can be saved and shared in the JavaScript Object Notation (JSON) file format, thus improving human readability of the respective representation (DuCharme 2011; Antoniou 2012).

3 Related Work

Semiconductor industry is in many respects strongly linked to the tremendous advance of digitalization and digitization approaches. First, its products and final goods of other downstream tiers serve as enablers of technology in general. Semiconductors are part of every device that drives digitalization and the technological world of the future. This is of course not limited to the domain of sensors but includes all microtechnology based devices, tools and equipment. This leads to the second connection of semiconductor industry, which holds true for many manufacturing industries and especially for data- and knowledge-intensive members. During almost every step of the product life cycle data is generated or used and hence serves as a potential source for the respective nature of a product in a later stage (Hesse and Schnell 2009). To ensure that all relevant process data is generated, maintained and accessible, it is important to provide a sufficient framework. For fluid and uninterrupted data exchange, an automated and fast responding connection between involved parties is required. With regards to automation, one aims at connecting either all systems that process the requested product or the product itself (Leite et al. 2019). In industrial space, this includes systems involved along the entire product lifecycle, which is to a large extent covered by the supply chain, respectively the supply network in a broader sense. Emerging new technologies like cloud-based automation will play a central role in connecting smart entities throughout the product engineering process in an intelligent way (Mahmoud 2019).

Semantic Web technologies are already part of promising approaches in certain areas like e-commerce or health care and life sciences to name but a few. Nevertheless, a broad-scale application in the industrial space lags behind (Petersen et al. 2017). Although current research for industry applications covers many of the issues that appear crucial in relation to Semantic Web technologies, state of science and technology is considered to be limited to proprietary implementations with regard to industrial environments (Leite et al. 2019). In the area of supply chains in general and Supply Chain Management and Planning more specifically, some ontology-based models have already been introduced in the research community (Ye et al. 2008, Zhai et al. 2008, Scheuermann and Leukel 2014, Kim et al. 2016, Ostrowski et al. 2016, Pal 2017). More recently one focus is, among others, on managing Retail Supply Chains with the help of service-oriented computing (SOA) based on Semantic Web (Pal 2018). Regarding manufacturing environments, research is emerging and facing potential solutions for overcoming hurdles like lack of awareness, successful use cases as well as technological issues. For instance, M2M communication is addressed, however lacking proper automation features (Gyrard et al. 2014).

Furthermore, recent research on Semantic Web technologies in the industrial space with special focus on automation, namely for semantic interoperability, is facing challenges and is to some extent limited to partial automation (Svetashova 2018). Recent approaches for semantic integration of sensors for enhanced automation are either restricted to a certain application and complexity layer (Petersen et al. 2017) or they are present in domains where data structures are less complex than the manufacturing environment and therefore considered more suitable for comprehensive digitization (Gray et al. 2011). Present research publications mostly cover specific use cases, yet a generic approach that is applicable to multiple scenarios is absent. Moreover, both holistic representation and integration of important standards that facilitate a broad scale implementation of Semantic Web technologies appear to be missing in the B2B environment.

4 Approach

By closely interlinking development processes, supply chain and production with semantic technologies, the iDev40 digital twin targets to achieve a disruptive step towards a digitalized product life cycle. The intermediate Semantic Web representation called digital reference serves as a basic concept to build the digital twin on. The digital reference is a result of the Horizon2020 ECSEL joint undertaking Productive4.0, being a complementary program to iDev40. Integrated Development 4.0 leads the digital transformation of singular processes towards an integrated virtual value chain based on this model. Development, planning and manufacturing will benefit from the digital twin concept in terms of digitized virtual processes along the whole product lifecycle, for instance via semantic-based supply chain analytics. Hence, it acts as an enabler to validate AI approaches in a variety of areas, including the prediction and simulation for development lots. It will also support structuring the build-up of learning skills (iDev40 2019).

In our understanding, the developed digital twin is representing all relevant data that is created throughout the entire product life cycle. This includes planning and development steps, production and delivery phases as well as data during the actual use of a product and its recycling, i.e. post-production. Possible applications include virtual testing and simulation, predictive maintenance and production failure analytics; among many others. In order to realize an uninterrupted linkage, one requires a unique virtual description of all involved entities (i.e. products, machines, etc.) and their relations. The need for a machine- and human-interpretable language leads to the promising approach of Semantic Web. Beyond the ontology of the digital twin, the challenge is linking networks with solely implicit connections and further optimizing the result. Moreover, it is of high importance to ensure highest quality with regard to validity and accuracy of the respective ontology. In order to fulfill these requirements, the intention is to use semantic web reasoners. Furthermore, the reuse of existing ontologies already agreed on (by the W3C, e.g.) is intended wherever possible. In this case, mapping and merging of classes and properties is necessary.

By providing the ontology as a formal shell of relations and entities, the representation yet lacks a proper integration of relevant data. This can either be solved by certain domain experts being very knowledgeable about specific topics that are no common knowledge. In this case, Protégé serves as a facilitated user interface for data inclusion. Another approach is accessing existing databases that are present in other formats by linking them to the ontology. This is for instance done by mapping database objects to classes and properties of the ontology or by having an additional external connection (via anchor layers, for instance). However, to ensure that all included data is both relevant and correct, it needs to be proven and accepted.

5 Preliminary Results

We are providing a basic principle of a semiconductor supply chain with its entities, processes and relations, represented by an RDF based ontology. The current version of this digital twin is depicted in Fig. 2 as visualized by WebVOWL. For better human readability, the various domains are visualized by separating them into different areas, called lobes. Furthermore, the single nodes are arranged based on a traditional understanding of organizational structures; using the Pick & Pin mode. Some lobes are referring to already existing standard ontologies that are recommended by W3C (Sensor ontology, for instance). Thus, it ensures a broad consensus of the content quality. Other parts of the digital twin, however, are specifically representing semiconductor manufacturing environment – or more specifically structures, products and processes that are unique for Infineon.

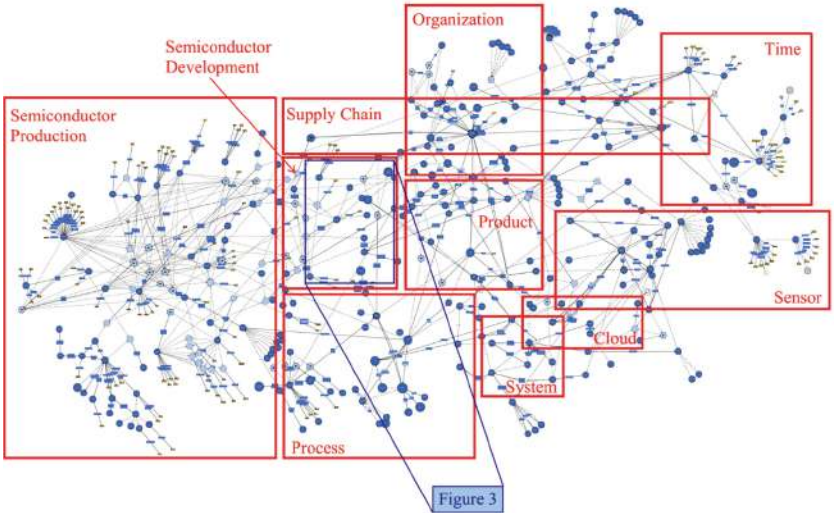

Fig. 2. The current version of the digital twin as applied in iDev40, as of Nov. 2018.

Many different tiers, projects and project partners lead to a variety of small, independent, problem-specific concepts. By representing them as ontologies and matching their entities with the existing digital twin, we are achieving enhanced collaboration due to fast data retrieval within a large amount of expert knowledge. This improves both collaboration between suppliers and customers along the supply chain as well as collaboration between project partners. The technology will help to enable complexity management within electronic component supplier (ECS) value chains that are most likely supply chains employing semiconductors. Additionally, Semantic Web technologies will support supply chains manufacturing semiconductors. Both benefits are specifically important to find solutions for the complex interrelationship between the technology developed and commercial users. Furthermore, Semantic Web technology is contributing to enhanced workplaces and smart collaboration in ECS value chains. A detailed view of the semiconductor development lobe in the current Pick & Pin digital reference version is depicted in Fig. 3. The extract shows the entities with their relationships. Additionally, by clicking on nodes and edges, one is able to gain insight in stored metadata, constraints and comments if existing. The visualization might serve as an explicit user interface for interested parties, having direct access to an expert knowledge base. Furthermore, by relating existing knowledge and reasoning options, it is possible to gain additional insight into related bits of knowledge and roots of the requested information.

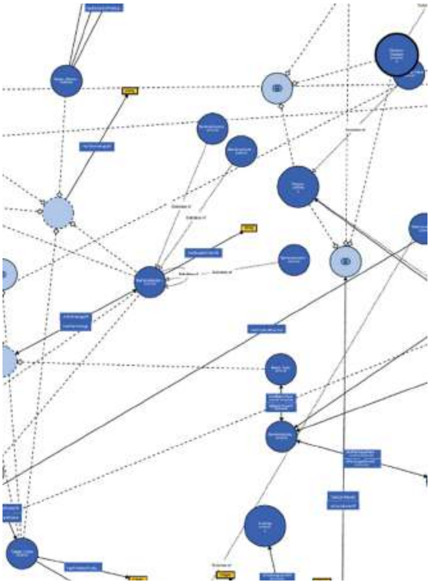

Fig. 3. Magnified area within semiconductor development lobe with detailed view on entities and relations.

Further ontologies have been prepared to support alternative business processes like revenue management (RM), possibly being an enabler for pricing mechanisms based on order lead time (OLT). Results include reducing the bullwhip effect and reaching both high capacity utilization and more deliveries as to customer wishes.

Another concept that is realized with the help of Semantic Web is the open online sales and marketing platform (OOSMP), on which integrated cyber-physical systems (ICPS), product vendors and component suppliers can meet virtually. It is being developed in Productive4.0 and might be extended to the whole product life cycle throughout the iDev40 project. The intention is that product vendors will be enabled to describe their product ideas and concepts with requirements to components, and component suppliers would be able to offer their components with features respectively. The visionary platform system is then intended to find matches between requirements and offers in an ecological, economical and societal balance along the entire life cycle. This is particularly important for a huge, unclear and quickly changing product portfolio of a semiconductor vendor. The Semantic Web and reasoners especially could enable further processing of search requests and linking it to prior experiences or other implicit correlations. Hence, recommendations for similar or new products as well as suggestions for substitutes of a requested product (if not on stock, for instance) would be possible. This may be a large step towards efficiently attracting new customers that are not familiar with the current product structures. Furthermore, development and production plans could be controlled and influenced considering knowledge about potential customers' behavior.

6 Reflections

In this paper we described our approach of a digital twin in the semiconductor environment. The described ontology is part of first use case scenarios, where we make use of major advantages like flexible data handling for large data sets, knowledge extraction and deduction of implicit knowledge. Within our current research projects, we are providing the current status to all interested partners for further improvement and large-scale adoption. Moreover, the current digital twin version is maintained on a regular basis to ensure consistency among its users.

The approach described in this paper leaves room for further improvement. In order to approach a holistic representation of the semiconductor manufacturing domain, some areas are missing to be translated into an ontology. This includes for instance automation standards, security and planning. Moreover, more relevant data is to be included in addition to classes and properties. In a further step, there have to be experts that agree on the quality and logical consistency of the imported data sets. After achieving a high-quality digital twin that consists of a sufficient set of relevant data, it is important to show its benefits in various use cases. This may include both business scenarios and collaboration experiments on a project level. Paying special attention to interconnecting different work packages on a use case level (as in iDev40, for instance), Semantic Web may contribute to an improved project exploitation. With getting an increasing number of experts involved, the digital twin will serve as a comprehensive knowledge representation that allows for fast knowledge extraction and deduction. Consequently, the aim for a large-scale adoption of the presented concept is getting closer. This is intended for applications in both the research community as well as the industrial space.

Further research may focus on adding relevant ontologies to the existing digital twin as well as developing a strategy to validate the added and existing domains. A major issue in this regard is that different parties involved have a diverging understanding of the same entity or the same lobe and their relations. This includes both scientific and semantic dimensions. Yet, by including the advice of domain experts, the comprehensiveness of the concept will improve over time in case there is a broadly admitted standard validation method. Other research topics are concerned with the possibilities of representing processes as graphs and of connecting existing databases efficiently. In the long run, the digital twin shall facilitate human-machine interaction along the entire value chain and product life cycle. In particular, it improves the use of relevant data with regards to speed and efficiency.

References

Andriopoulou, F., Dagiuklas, T., Orphanoudakis, T.: Integrating IoT and Fog computing for healthcare service delivery. In: Keramidas, G., Voros, N., Hübner, M. (eds.) Components and Services for IoT Platforms, vol. 54, pp. 213–232. Springer, Cham (2017)

Antoniou, G.: A Semantic Web Primer, 3rd edn. MIT Press, Cambridge (2012). (Cooperative information systems)

Baumgärtel, H., Ehm, H., Laaouane, S., Gerhardt, J., Kasprzik, A.: Collaboration in supply chains for development of CPS enabled by semantic web technologies. In: 14th International Conference on Modeling and Analysis of Semiconductor Manufacturing (MASM) at Winter Simulation Conference 2018. Gothenburg, Denmark (2018)

Berners-Lee, T.: Semantic Web on XML. XML 2000. W3C. Washington DC, USA, (2000). http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html. Accessed 21 Dec 2018

Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. In: Scientific American Magazine, vol. 284 (2001)

DuCharme, B.: Learning SPARQL: Querying and Updating with SPARQL 1.1. O'Reilly Media, Sebastopol (2011)

Dustdar, S., Falchuk, B.: Semantic web. In: Furht, B. (ed.) Encyclopedia of Multimedia, pp. 803–810. Springer, New York (2006)

Gray, Alasdair J.G., Sadler, J., Kit, O., Kyzirakos, K., Karpathiotakis, M., Calbimonte, J.-P., et al.: A semantic sensor web for environmental decision support applications. Sensors **11**(9), 8855–8887 (2011). https://doi.org/10.3390/s110908855. (Basel, Switzerland)

Gyrard, A., Bonnet, C., Boudaoud, K.: Enrich machine-to-machine data with semantic web technologies for cross-domain applications. In: 2014 IEEE World Forum on Internet of Things (WF-IoT), Seoul, Korea (South), 06-08 March 2014, pp. 559–564. IEEE (2014)

Hesse, S., Schnell, G.: Sensoren für die Prozess- und Fabrikautomation. Funktion - Ausführung - Anwendung. 4., aktualisierte und erw. Aufl. Wiesbaden: Vieweg + Teubner Verlag/ GWV Fachverlage, Wiesbaden (Vieweg Praxis) (2009)

Hofmann, E.: Big data and supply chain decisions: the impact of volume, variety and velocity properties on the bullwhip effect. Int. J. Prod. Res. **55**(17), 5108–5126 (2017). https://doi.org/10.1080/00207543.2015.1061222

iDev40: Project Work. Work Plan (2019). http://www.idev40.eu/project-work/work-plan. Accessed 7 Jan 2019

Kim, M., Wang, S.-T., Ostrowski, D., Rychtyckyj, N., Macneille, P.: Technology outlook: federated ontologies and industrial applications. Int. J. Semant. Comput. **10**(01), 101–120 (2016). https://doi.org/10.1142/S1793351X1650001X

Lacy, L.W.: Interchanging Discrete Event Simulation process Interaction Models Using the Web Ontology Language - OWL. Dissertation. University of Central Florida. College of Engineering and Computer Science (2006)

Leite, J.R.E., Martins, P.S., Ursini, E.L.: Internet of Things: an overview of architecture, models, technologies, protocols and applications. In: Iano, Y., Arthur, R., Saotome, O., Vieira Estrela, V., José Loschi, H. (eds.) Proceedings of the 3rd Brazilian Technology Symposium, vol. 113, pp. 75–85. Springer, Cham (2019)

Mahmoud, M.S.: Architecture for cloud-based industrial automation. In: Yang, X.S., Sherratt, S., Dey, N., Joshi, A. (eds.) Third International Congress on Information and Communication Technology. Advances in Intelligent Systems and Computing, vol. 797, pp. 51–62. Springer, Singapore (2019)

Mane, B., Fall, I., Camara, M.S., Bah, A.: Ontological modeling approach for statistical databases publication in linked open data. In: Yang, X.S., Sherratt, S., Dey, N., Joshi, A. (eds.) Third International Congress on Information and Communication Technology. Advances in Intelligent Systems and Computing, vol. 797, pp. 277–292. Springer, Singapore (2019)

Mujica, G., Portilla, J., Riesgo, T.: Deployment strategies of wireless sensor networks for IoT: challenges, trends, and solutions based on novel tools and HW/SW platforms. In: Keramidas, G., Voros, N., Hübner, M. (eds.) Components and Services for IoT Platforms, vol. 9, pp. 133–154. Springer, Cham (2017)

Ostrowski, D., Rychtyckyj, N., Macneille, P., Kim, M.: Integration of big data using semantic web technologies. In: 2016 IEEE Tenth International Conference on Semantic Computing (ICSC), Laguna Hills, CA, USA, 04–06 February 2016, pp. 382–385. IEEE (2016)

Pal, K.: A semantic web service architecture for supply chain management. Procedia Comput. Sci. **109**, 999–1004 (2017). https://doi.org/10.1016/j.procs.2017.05.442

Pal, K.: Ontology-based web service architecture for retail supply chain management. Procedia Comput. Sci. **130**, 985–990 (2018). https://doi.org/10.1016/j.procs.2018.04.101

Petersen, N., Halilaj, L., Grangel-González, I., Lohmann, S., Lange, C., Auer, S.: Realizing an RDF-based information model for a manufacturing company – a case study. In: d'Amato, C., Fernandez, M., Tamma, V., Lecue, F., Cudré-Mauroux, P., Sequeda, J., et al. (eds.) The Semantic Web – ISWC 2017, vol. 10588, pp. 350–366. Springer, Cham (2017)

Rumpl, W.E., Auinger, F., Dutzler, C., Zoitl, A.: Platforms for scalable flexible automation considering the concepts of IEC 61499. In: Mařík, V., Camarinha-Matos, L.M., Afsarmanesh, H. (eds.) Knowledge and Technology Integration in Production and Services, pp. 237–246. Springer, Boston (2002)

Scheuermann, A., Leukel, J.: Task ontology for supply chain planning – a literature review. Int. J. Comput. Integr. Manuf. **27**(8), 719–732 (2014). https://doi.org/10.1080/0951192X.2013.834472

Siozios, K., Danassis, P., Zompakis, N., Korkas, C., Kosmatopoulos, E., Soudris, D.: Supporting decision making for large-scale IoTs: trading accuracy with computational complexity. In: Keramidas, G., Voros, N., Hübner, M. (eds.) Components and Services for IoT Platforms, vol. 1, pp. 233–250. Springer, Cham (2017)

Svetashova, Y.: Semantic model extensibility in interoperable IoT data marketplaces: methods, tools, automation aspects. In: Kirrane, S., Kagal, L. (eds.) Proceedings of the Doctoral Consortium at ISWC 2018 co-located with 17th International Semantic Web Conference (ISWC 2018), CEUR-WS.org (CEUR Workshop Proceedings, 2181), Monterey, USA, 8th - to - 12th October 2018, pp. 73–82 (2018). http://ceur-ws.org/Vol-2181/paper-09.pdf

Ye, Y., Yang, D., Jiang, Z., Tong, L.: Ontology-based semantic models for supply chain management. Int. J. Adv. Manuf. Technol. **37**(11–12), 1250–1260 (2008). https://doi.org/10.1007/s00170-007-1052-6

Zhai, J., Li, Y., Wang, Q., Lv, M.: Knowledge sharing for supply chain management based on fuzzy ontology on the semantic web. In: 2008 International Symposiums on Information Processing ISIP, Moscow, Russia, 23–25 May 2008, pp. 429–433. IEEE (2008)

Framework for Simulation-Based Decision Making in Semiconductor Value Chains

Lars Mönch[1(✉)], Hans Ehm[2], and Thomas Ponsignon[2]

[1] University of Hagen, Chair of Enterprise-wide Software Systems, Hagen, Germany
lars.moench@fernuni-hagen.de
[2] Infineon Technologies AG, Supply Chain Innovation, Neubiberg, Germany

Abstract. Simulation is an important industrial engineering tool in many semiconductor wafer fabrication facilities. In this paper, we are interested in obtaining an understanding of the capabilities of simulation to support various planning and control functions in semiconductor value chains. This research effort is challenging since it requires knowledge of the different planning and control functions and the related decision-making approaches. It seems also to be difficult to generate supply chain simulation models from scratch or reuse existing simulation models of an appropriate level of detail. In this paper, we discuss the major design issues for a framework to guide decision-makers in semiconductor supply chains towards choosing an appropriate simulation-based decision support.

Keywords: Semiconductor supply chains · Production planning and control · Simulation-based decision making

1 Motivation

The process of manufacturing integrated circuits is one of the most complex technological achievements of the 20th century (Doering and Nishi 2007; Mönch et al. 2013; Mönch et al. 2018). The capital-intensive nature of the semiconductor industry leads to production facilities that run consistently at high utilization levels. The reentrant process flows in the production facilities cause complex competition for scarce capacity. The increasing level of automation reduces the ability to use humans to buffer the equipment. The sheer size of the facilities and value chains involved, the pervasive presence of different kinds of uncertainties and the rapid pace of change result in an environment that is challenging for current planning and control approaches and the related information and decision support systems (Chien et al. 2011).

Different simulation paradigms, such as discrete-event simulation (Fowler et al. 2015), agent-based simulation (Bonabeau 2002; Klügl and Bazzan 2012; Achter et al. 2017), or system dynamics (Sterman 2000) can be used at the different planning and control levels in semiconductor value chains. However, while simulation is often applied for representing the execution level of semiconductor value chains, simulation is only rarely used to make planning decisions (Fowler et al. 2015; Mönch et al. 2018).

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 14–23, 2020.
https://doi.org/10.1007/978-3-030-48602-0_2

There are several reasons for this observation. First of all, building fab- or even value chain-wide simulation models from scratch is very time-consuming and error-prone. Although data availability has improved to a large extent over the last decade, it is still difficult to gather and process data for specific purposes related to building simulation models, for instance, for modeling machine breakdowns in a correct way. Secondly, although simulation appears at a first glance to be a tool that can be easily used, successful simulation projects require experience in modeling and statistics and also deep domain knowledge. Using simulation for making planning and control decisions often requires many simulation replications due to multiple scenarios for what-if analysis or for simulation-based optimization approaches. Therefore, it is likely that a simulation application will create a large computational burden.

There is a need to systematically investigate the improvement potential on the value chain level when simulation is used to support planning and control decisions. Important examples are simulation-based optimization (Forstner and Mönch 2013) and using simulation within what-if analysis scenarios for designing and running semiconductor value chains (Gonçalves et al. 2005; Lin et al. 2018). This requires that simulation models for semiconductor value chains are either fully automatically generated on demand using data in information systems of the value chain or that appropriate simulation testbeds are available.

This paper is organized as follows. The researched problem is described in the next section. This includes an introduction to the notion of frameworks due to (Porter 1991). First results are presented in Sect. 3. Moreover, conclusions are discussed and future research directions are identified in Sect. 4.

2 Description

2.1 Initial Situation and Framework Notion

We strive for a better understanding of planning and control situations where simulation-based decision support leads to improved decision making compared to conventional, non-simulation-based approaches. Our goal is to design a framework that supports decision-makers in semiconductor value chains to choose an appropriate type of simulation-based decision support.

According to (Porter 1991), a framework will be designed with the goal to support a certain point of view on a given class of problems. Frameworks offer a broad, comprehensive problem description together with structuring instruments that are more suited to deal with the complexity in companies than models that are based on various assumptions. Frameworks identify the relevant variables and the questions which a user must answer to come to conclusions tailored for a specific situation (Porter 1991).

2.2 Overall Approach

Based on the framework idea, this research effort is divided into the following steps:

1. In an analysis phase, we identify relevant variables that have an impact on the success of a simulation-based decision support in semiconductor value chains.

2. In a second phase, questions are derived that have to be answered to select an appropriate simulation-based decision support. This leads to the framework.
3. In a third phase, the framework will be tested by means of use cases related to the simulation-based support of decision-making activities in semiconductor value chains.
4. Based on the results of the third phase, repeating the first three phases might be desirable.

The overall approach is depicted in Fig. 1. We can clearly observe the iterative nature of the overall approach.

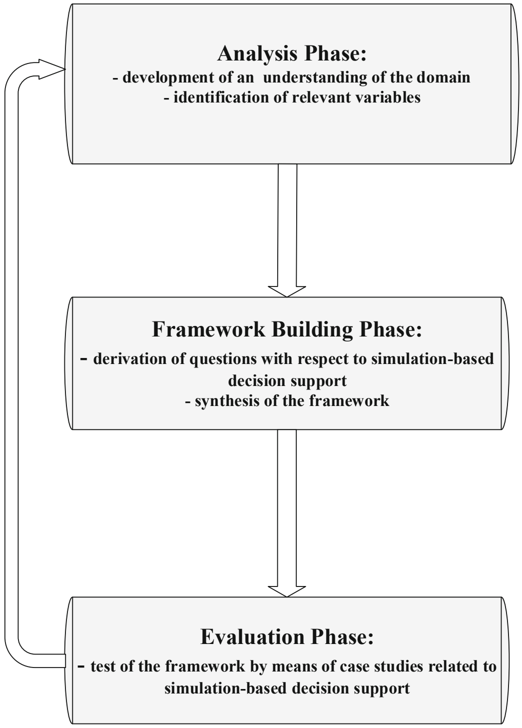

Fig. 1. Iterative phase model of the overall approach

3 Results

3.1 Towards a Framework for Simulation-Based Decision-Making

The main variables are the simulation paradigm and the type of the related simulation models to be applied and the type of the planning tasks. Different simulation model types can be differentiated (Mönch et al. 2018), namely

- agent-based simulation models
- system dynamics models
- detailed discrete-event simulation models
- reduced discrete-event simulation models
- hybrid models that combine different of the aforementioned approaches.

Agent-based simulation models allow for modeling the behavior of a human decision-maker since software agents are software entities that proactively make decisions to fulfill their design goals. Agent-based simulation can be used to study the emergent behavior of the different entities in a system.

System dynamics models allow for capturing the nonlinear behavior of systems based on flows, stocks, feedback loops, and time delays (cf. Sterman 2000). System dynamics models are generally deterministic in nature. They only allow for estimating the average system performance. Moreover, they have an aggregated structure and are based on the continuous time simulation paradigm. Therefore, running system dynamics models results in shorter computing times than performing simulations with discrete-event simulation models for manufacturing systems of the same size.

Discrete-event simulation models allow for a high-fidelity representation of the base system and process of a manufacturing system or a value chain. However, simulation studies based on such models tend to be time-consuming due to the many details included in such simulation models. Various approaches exist to reduce the level of detail, for instance, modeling only bottleneck resources in detail and replacing the operations on non-bottleneck resources by delays (Ewen et al. 2017, van der Zee 2019).

Hybrid approaches, also known as multi-model approaches, are proposed to combine the advantages of the different simulation paradigms and to mitigate the corresponding limitations. For instance, it is possible to use discrete-event simulation to represent the base system and process, while system dynamics is used to represent the planning model (cf. Venkateswaran and Son 2005). However, the management of the different models within a hybrid approach is a nontrivial exercise.

The simulation paradigm is influenced by the following independent variables:

- characteristics of the decision-making processes to be supported
- required level of detail
- data availability in operational and strategic information systems of the value chain
- time available for decision making.

If the decisions are not fully automatically made by algorithms, i.e. if human decision-makers are involved, the agent-based simulation paradigm is appropriate since it is possible to model the behavior of the human decision-makers to some extent. If the required level of detail is low and the time available for decision-making is also low then system dynamics is often more appropriate than other simulation paradigms. This is also true if the data availability is low. In the remaining situations discrete-event simulation is appropriate.

The planning and control tasks to be supported can be derived from the supply chain planning matrix for semiconductor value chains. The planning matrix is shown in Fig. 2.

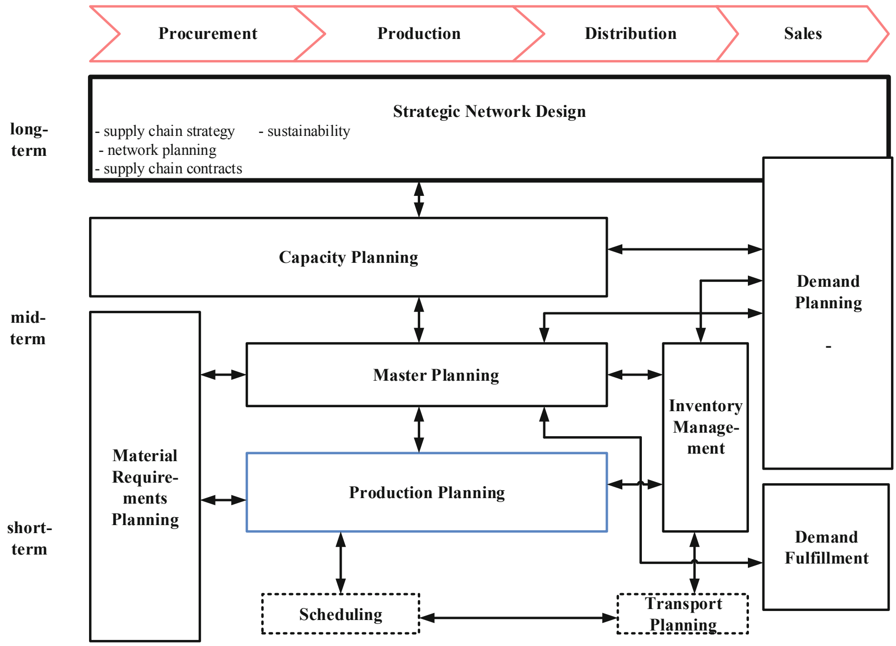

Fig. 2. Supply chain planning matrix for semiconductor value chains (Mönch et al. 2018)

We see that short-term, mid-term, and long-term decisions are supported, i.e., the vertical axis shows the time frame. The horizontal axis of the matrix represents material flow across the business functions. The different planning functions, represented by rectangles, produce decisions that form the input for other planning functions.

The following activities can be supported by simulation:

1. using simulation in a repeated manner within metaheuristic-based optimization approaches to evaluate the objective function (simulation-based optimization)
2. using simulation to estimate parameters of planning models, this leads to a multi-model approach that iterates between an optimization model with prescribed parameter values and a simulation model that estimates parameter values based on the optimization results
3. execution of single plans under uncertainty
4. execution of plans in a rolling horizon setting
5. using simulation to model the planning and control process itself
6. using simulation to model the stochastic behavior of external factors.

The first three approaches can be applied in a static and stochastic setting, while the fourth approach can be used in a dynamic and stochastic setting. Moreover, the first and the second approach can be used to derive plans, for instance, production plans or schedules, while the third and fourth approach requires a plan as an input. Note that the third approach is implicitly embedded into the first and the second approach, but it can also be used in a standalone manner. Moreover, the first and second approach can be combined with the fourth approach when the quality of the derived plans in a dynamic and stochastic setting is of interest.

The fifth approach is useful if several entities are involved in the planning and control process, mainly when human decision makers are involved. In this situation, it might be useful to study the interaction of the different entities. Cleary, a dynamic and stochastic setting is considered. Finally, it might be interesting to model the behavior of external factors such as demand by simulation.

The framework equips each planning function in Fig. 2 with appropriate simulation paradigms and with meaningful simulation activities. Therefore, the decision-maker can ask the following questions:

1. What type of conventional non-simulation based planning or control approach is applied?
2. Why is it likely that a simulation-based approach outperforms the conventional approach?
3. What is the time horizon of the considered planning function?
4. What time is available for decision making?
5. What is the required level of detail when decisions are made?
6. Is the required data for building the required simulation model available in a digitalized and structured form?
7. Which effort is required to build a corresponding simulation model from scratch?
8. Is an automated or semi-automated simulation model generation possible?
9. Which effort is required to maintain and update already existing simulation models?

Some preliminary results and findings of this process are summarized in Table 1.

Demand planning is challenging in semiconductor value chains (cf. Uzsoy et al. 2018). Fully automated planning approaches rarely exist. Therefore, it is interesting to model the planning process itself. The planning capabilities of planners together with the interactions of the different decision-makers can be studied (cf. Hauke et al. 2018). Agent-based simulation is appropriate for this setting.

There are also examples known where the interaction of customers is modeled. An agent-based simulation model to predict the sales of microprocessors in the high-end gaming market, focusing on the decision of a customer to purchase a more powerful computer is proposed by (Adriaansen et al. 2013). A software agent who makes purchasing decisions based on a customer-specific internal logic is used to represent each individual customer.

Production planning and scheduling/dispatching are mainly short-term by nature (see Fig. 2). Therefore, a fairly high level of detail is needed. The time for decision making is typically short, ranging from a few seconds to compute a schedule until an hour for production plans. The data availability is typically given due to existing Manufacturing Execution Systems (MES) and Enterprise Resource Planning (ERP) systems that can be found in most companies. Therefore, it is possible and appropriate to work with discrete-event simulation models. In some situations it might be interesting to study the emergent behavior of lots when they compete for the scarce resources in wafer fabs. In this situation, agent-based modeling and simulation is appropriate. A related example can be found in (Mönch 2006).

Table 1. Application framework

Planning function	Simulation paradigm	Simulation activities
Demand planning	Agent-based simulation	• Using simulation to model the planning process itself • Using simulation to model external factors (i.e., the demand)
Production planning	Discrete-event simulation	• Simulation-based optimization • Using simulation to estimate parameter values of planning models • Execution of single plans under uncertainty • Execution of plans in a rolling horizon setting
Scheduling and dispatching	Discrete-event simulation/ agent-based simulation	• Simulation-based optimization • Using simulation to estimate parameter values of scheduling approaches • Execution of single schedules under uncertainty • Execution of schedules in a rolling horizon setting • Using simulation to model the scheduling and dispatching process itself

3.2 Examples from Production Planning

The framework is exemplified by the production planning function depicted by a blue-rimmed rectangle in Fig. 2. Since production planning activities are short-term by nature, the discrete-simulation paradigm, either based on a detailed or a reduced simulation model, is a reasonable simulation approach as justified in Subsect. 3.1.

Examples for the first four simulation activities are known, for instance, the execution of production plans in a large-scale wafer fab simulation model is used by (Kacar et al. 2013), (Liu et al. 2011) apply simulation-based optimization to compute production plans, iterative simulation and linear programming are applied by (Hung and Leachman 1996), and production planning formulations for a single wafer fab based on linear programming are assessed in a rolling horizon manner by (Ziarnetzky et al. 2018).

Linear programming-based production planning formulations for production and engineering activities are assessed by (Ziarnetzky and Mönch 2016, Ziernetzky et al. 2017) in a rolling horizon setting. Process improvement activities by engineering lots are crucial for semiconductor wafer fabrication facilities to stay competitive in the semiconductor market (Mönch et al. 2013). Production and engineering lots compete for the same equipment with scarce capacity. Two production planning formulations for a simplified semiconductor value chain are discussed. The first formulation assumes reduced available capacity for production, while the second one directly incorporates engineering activities. Additional capacity is considered in this integrated formulation because of learning effects that represent process improvements. The integrated formulation outperforms the conventional formulation in a rolling horizon setting with respect to profit.

4 Discussion/Implications

In this research, we are interested in obtaining a fairly complete understanding of the capabilities of simulation to support various planning and control functions in semiconductor value chains. This research effort is challenging since, on the one hand, it requires knowledge of the different planning and control functions and the related decision-making approaches. On the other hand, it seems to be challenging to generate supply chain simulation models from scratch or reuse existing simulation models of an appropriate level of detail.

While the research described in this paper is an ongoing effort, there are several directions for future research. First of all, it is required to complete the framework by investigating the remaining planning functions in Fig. 2. Moreover, rigorous case studies are required to support the third step of the framework building process. Based on the insights from the case studies, several iterations of the framework building process are possible to improve the framework. The development of a rich simulation testbed for semiconductor value chains is a necessary requirement for performing the case studies.

References

Achter, S., Lorscheid, I., Hauke, J., Meyer, M., Meyer-Riel, D., Ponsignon, T., Sun, C., Ehm, H.: On agent-based modeling in semiconductor supply chain planning. In: Proceedings of the 2017 Winter Simulation Conference, pp. 3507–3518 (2017)

Adriaansen, T., Armbruster, B., Kempf, K., Li, H.: An agent model for the high-end gamers market. Adv. Complex Syst. Multidiscip. J. **16**(7), 1350028 (2013)

Bonabeau, E.: Agent-based modeling: methods and techniques for simulating human systems. Proc. Natl. Acad. Sci. **99**(3), 7280–7287 (2002)

Chien, C., Dauzère-Pérès, S., Ehm, H., Fowler, J., Jiang, Z., Krishnaswamy, S., Mönch, L., Uzsoy, R.: Modeling and analysis of semiconductor manufacturing in a shrinking world: challenges and successes. Eur. J. Ind. Eng. **5**(3), 254–271 (2011)

Doering, R., Nishi, Y. (eds.): Handbook of Semiconductor Manufacturing Technology. CRC Press, Boca Raton (2007)

Ewen, H., Mönch, L., Ehm, H., Ponsignon, T., Fowler, J.W., Forstner, L.: A testbed for simulating semiconductor supply chains. IEEE Trans. Semicond. Manuf. **30**(3), 293–305 (2017)

Forstner, L., Mönch, L.: A heuristic to support make-to-stock, assemble-to-order, and make-to-order decisions in semiconductor supply chains. In: Proceedings of the 2013 Winter Simulation Conference, pp. 3696–3706 (2013)

Fowler, J., Mönch, L., Ponsignon, T.: Discrete-event simulation for semiconductor wafer fabrication facilities: a tutorial. Int. J. Ind. Eng.: Theory Appl. Pract. **22**(5), 661–682 (2015)

Gonçalves, P., Hines, J., Sterman, J.: The impact of endogenous demand on push-pull production systems. Syst. Dyn. Rev. **21**(3), 187–216 (2005)

Hauke, J., Lorscheid, I., Meyer, M.: Individuals and their interactions in demand planning processes: an agent-based, computational testbed. Int. J. Prod. Res. **56**, 4644–4658 (2018)

Hung, Y., Leachman, R.: A production planning methodology for semiconductor manufacturing based on iterative simulation and linear programming calculations. IEEE Trans. Semicond. Manuf. **9**, 257–269 (1996)

Kacar, N., Mönch, L., Uzsoy, R.: Planning wafer starts using nonlinear clearing functions: a large-scale experiment. IEEE Trans. Semicond. Manuf. **26**(4), 602–612 (2013)

Klügl, F., Bazzan, A.L.C.: Agent-based modeling and simulation. AI Mag. (2012)

Lin, Y., Spiegler, V.L.M., Naim, M.M.: Dynamic analysis and design of a semiconductor supply chain: a control engineering approach. Int. J. Prod. Res. **56**(13), 4585–4611 (2018)

Liu, J., Li, C., Yang, F., Wan, H., Uzsoy, R.: Production planning for semiconductor manufacturing via simulation optimization. In: Proceedings of the 2011 Winter Simulation Conference, pp. 3617–3627 (2011)

Mönch, L.: Agentenbasierte Produktionssteuerung komplexer Produktionssysteme. Deutscher Universitätsverlag (DUV), Gabler, Wiesbaden (2006)

Mönch, L., Fowler, J.W., Mason, S.J.: Production planning and control for semiconductor wafer fabrication facilities: modeling, analysis and systems. Springer Operations Research/Computer Science Interfaces, New York, vol. 52 (2013)

Mönch, L., Uzsoy, R., Fowler, J.: A survey of semiconductor supply chain models part I: semiconductor supply chains, strategic network design, and supply chain simulation. Int. J. Prod. Res. **56**(13), 4524–4545 (2018)

Porter, M.E.: Toward a dynamic theory of strategy. Strateg. Manag. J. **12**, 95–117 (1991)

Sterman, J.: Business Dynamics: Systems Thinking and Modeling for a Complex World. McGraw-Hill, Boston (2000)

Uzsoy, R., Fowler, J.W., Mönch, L.: A survey of semiconductor supply chain models part II: demand planning, inventory planning, and capacity planning. Int. J. Prod. Res. **56**(13), 4546–4564 (2018)

van der Zee, D.-J.: Model simplification in manufacturing simulation – review and framework. Comput. Ind. Eng. (2019, in press)

Venkateswaran, J., Son, Y.J.: Hybrid system dynamic - discrete event simulation based architecture for hierarchical production planning. Int. J. Prod. Res. **43**(20), 4397–4429 (2005)

Ziarnetzky, T., Mönch, L.: Incorporating engineering process improvement activities into production planning formulations using a large-scale wafer fab model. Int. J. Prod. Res. **54** (21), 6416–6435 (2016)

Ziarnetzky, T., Mönch, L., Uzsoy, R.: Rolling horizon production planning for wafer fabs with chance constraints and forecast evolution. Int. J. Prod. Res. **56**(18), 6112–6134 (2018)

Ziarnetzky, T., Mönch, L., Ponsignon, T., Ehm, H.: Rolling horizon planning with engineering activities in semiconductor supply chains. Proc. IEEE CASE **2017**, 1024–1025 (2017)

Digital Twin for Plan and Make Using Semantic Web Technologies – Extending the JESSI/SEMATECH MIMAC Standard to the Digital Reference

Patrick Moder[1,2](✉), Hans Ehm[2], and Nour Ramzy[2]

[1] Department of Mechanical Engineering, Institute of Automation and Information Systems AIS, Technical University of Munich, TUM, Boltzmannstr. 15, 85748 Garching, Germany
patrick.moder@infineon.com

[2] Infineon Technologies AG, Corporate Supply Chain Engineering Innovation (IFAG CSC E IN), Am Campeon 1-15, 85579 Neubiberg, Germany

Abstract. This proposed research is concerned with developing an extension of the joint JESSI/SEMATECH standard MIMAC to the Digital Reference. It analyzes the capabilities of a Semantic Web based Digital Twin for the semiconductor supply chain Plan and Make processes. The long-term goal is an accelerated broad-scale digitalization of the entire supply chain. This incorporates benefits for fab simulation, inter-enterprise collaboration, inconsistency identification and knowledge transfer. Industry related use cases show the potential and allow an analysis of current difficulties.

Keywords: SCOR · Supply Chain Management · Manufacturing · Semantic Web · Ontology · Digital Twin · Manufacturing Capacity

1 Motivation

The union of Joint European Submicron Silicon Initiative (JESSI) and the U.S. SEMATECH (derived from Semiconductor Manufacturing Technology) consortium in the early 90s of the last century is considered as the foundation of a strong collaboration within the semiconductor industry. The collaboration aims at identifying the main factors that influence development and manufacturing efficiency to improve the performance on a strategic level. Established for this purpose, the Measurement and Improvement of Manufacturing Capacity (MIMAC) pre-competitive project analyzed coefficients that led to drops in production capacity. The developed data sets, which also include a simulation model of an Infineon facility, are serving as reliable simulation reference models since then, yet the changing semiconductor environment necessitates adjustments of the model to a certain extent (Hassoun and Kalir 2017). Discrete-event and (multi) agent-based simulation approaches are leaving the tool and factory level towards the entire semiconductor supply chain, partially going beyond and covering the domains of value chains employing semiconductors. However, the core remains to cover the operations of manufacturing and development of lots of

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 24–32, 2020.
https://doi.org/10.1007/978-3-030-48602-0_3

semiconductor products (Fowler and Robinson 1995). With a multitude of processing steps and tool interrelations, the semiconductor manufacturing (Make) section of the supply chain is considered as highly complex. Hence, planning this supply chain is crucial for handling demand or supply volatility and managing global production flexibility.

Introducing Horizon 2020, the European Commission maintains the largest international collaborative research & innovation program. ECSEL, the Initiative on Electronic Components and Systems for European Leadership is the joint strategic approach reusing the JESSI experiences. Projects under the ECSEL umbrella – such as iDev40 and Productive4.0 – contribute to reaching the next level for semiconductor development, manufacturing and supply chain. Main reasons for the highly turbulent environment are shorter product lifecycles, high demand uncertainty and fluctuation, changing manufacturing processes as well as an increasing quantity of fabs involved in the globalized production of a single product (Chien et al. 2008). This includes an enhanced connection of the digital and physical world for smarter, automatable processes that are more flexible, scalable and versatile (Mönch et al. 2018). However, it is still unclear, to what extent the semiconductor industry will be influenced and which players are capable of driving the digital transformation.

Semantic Web technologies are a promising approach in order to ensure that the current B2B actors are staying in the driver seat in their field and are not taken over by B2C competition. Data is unified substantially in consistent ontologies by defining explicit semantics and can be enriched in a later phase. Furthermore, the machine understandable and interpretable structure improves collaboration within and between enterprises as well as on a workcenter layer (Baumgärtel et al. 2018). This paper describes a consistent digital twin approach that accompanies the semiconductor manufacturers with their supplier and customer tiers during the digital transformation by enhancing existing MIMAC standard data sets with Semantic Web technologies.

2 Background

The industrial importance and scientific attention of supply chain management concepts increased in the 1980s as strategic components were added, however, it initially emerged as logistics concept in the 1950s (Oliver and Webber 2012). As one of the most renowned approaches to manage supply chains, the Supply Chain Operations Reference Model (SCOR Model) contains recommendations for planning, controlling and monitoring (Bolstorff et al. 2007). The SCOR Model is endorsed by the Supply Chain Council (SCC) with its vast number of member companies and therefore serves as a harmonized industry standard across sectors. It furthermore allows efficient benchmarking procedures for supplier selection. The basic reference model includes the phases Plan, Source, Make, Deliver and Return, as depicted in Fig. 1. It is developed as a management tool, covering all business activities from the supplier's supplier to the customer's customer with the overarching goal to satisfy the customer's demand. The SCOR model is applied for describing, measuring and analyzing supply chain configurations. Furthermore, supply chain relevant sub processes are harmonized and become visible. Considering details within and between the sub phases, it is of major interest how to design and control the processes for material, information and value flow. The manufacturing (Make) phase is primarily processing material downstream and relevant data in both directions for controlling the production environment. Especially for the semiconductor industry and its various sensitive manufacturing systems, it includes a highly complex task that needs to be aligned effectively. This is achieved by the means of the Plan phase and incorporated processes. Within the Plan phase, major reasons for the high complexity of the semiconductor supply chain are aimed to be controlled, namely short lead times, frequent product innovations or volatile demand and supply behavior.

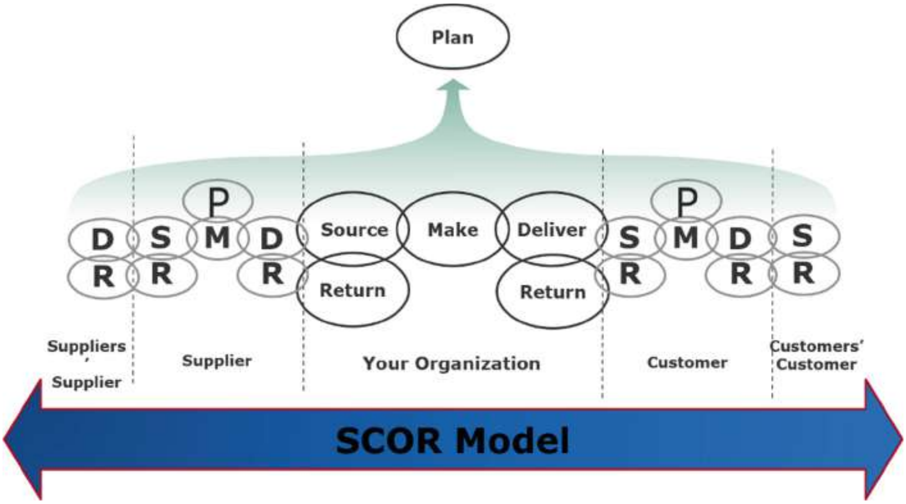

Fig. 1. The SCOR model. Cf. Bolstorff et al. (2007).

Addressing the needs of large and complex manufacturing environments, the joint JESSI/SEMATECH standard MIMAC incorporates the most relevant factors that influence fab capacity in a negative manner. It is based on surveys and interviews of member companies that reveal the importance of cycle-time and on-time delivery for semiconductor manufacturing. Furthermore, literature research and analysis of factory-level effects are conducted. Besides the identification of the most crucial factors, the results include a discrete-event simulation testbed for semiconductor manufacturing, capacity planning process reports and training tools for workshop participants. For the global experiment, after eliminating some variables during simulation, nine factors are consequently identified that have major effect on factory capacity. Namely, this is downtime, setup, yield, batching, operator availability, rework, operator cross-training, dispatch rules and hot lots. It incorporates capacity that is constrained by cycle time at different scenarios and calculated with different data sets. The complex and resource-expensive experiment provides lessons learned for the future in the form of characteristic curves. However, it is pointed out that each factory tends to behave differently due to individual factors that are not part of the experiment. Similar approaches in the area of semiconductor manufacturing scheduling improvement strategies mainly cover simulation- and agent-based control solutions (Mönch 2006), heuristic multi-criteria optimization approaches (Pfund et al. 2008), simulation models and policies (Singh and Mathirajan 2018) as well as rule-based optimization techniques that are based on the work from MIMAC (Mittler and Schoernig 2000).

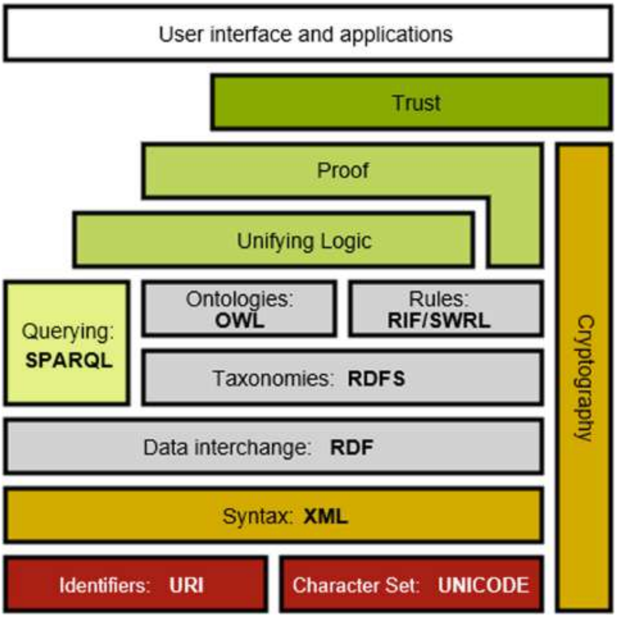

Fig. 2. The Semantic Web stack. Cf. Berners-Lee et al. (2001).

For overcoming the challenges that appear due to an increasing number of generated and processed data points, efficient data management strategies are necessary. Semantic Web Technologies are a promising approach that enable both data integration from heterogeneous sources and interpretability for machines. Semantic Web is an extension of the World Wide Web (WWW) framework that now connects data instead of hyperlinked documents. The building blocks of Semantic Web Technologies are depicted in Fig. 2, showing the Semantic Web Stack. Semantic Web applies the Resource Description Framework (RDF) to represent knowledge taxonomically, thus allowing to model resources with properties and linkages that identify them. Information is defined by the means of triples with each of the triples containing a subject, predicate and an object. The subject is the *thing* of interest, similar to the subject of a statement. The predicate, also referred to as property, links the subject with the object and hence defines their relationship. This relationship is either of type object property or data property. Here, an object property is applied to relate subjects to objects that are resources. In a similar way, a data property connects a subject with data-related information. A unique identifier (URI) is defined for each resource and property, referring to the address where the definition is stored. Ontologies summarize the knowledge about a certain domain and contain of classes – and presumably their subclasses –, individuals and a set of relations. Individuals, also known as instances, are the most individual objects that belong to a class. Resource Description Framework Schema (RDFS) extends RDF and offers a set of property and class vocabularies and hence facilitates the generation of hierarchies. Extending RDF and RDFS and expressing their constructs further, the Web Ontology Language (OWL) enables a variety of more descriptive properties, relationships and class descriptions. OWL is initially developed and now maintained by the World Wide Web Consortium (W3C). It serves as a common language for the Semantic Web community and additionally defines a broad set of standards. (Berners-Lee et al. 2001).

However, OWL adds even more value to the model, since reasoning may be applied. Reasoning is a powerful mechanism that automatically deduces implicit knowledge from the explicit knowledge graph. It is possible to review and trace each step of deduction, hence the knowledge gaining process is not hidden but rather clearly documented. For defining the area in which the inferences may or may not hold, constraints are necessary. Therefore, rule languages like the Semantic Web Rule Language (SWRL) or the Rule Interchange Format (RIF) are applied. The knowledge that is stored in an ontology is retrieved by queries, mainly based on the SPARQL (SPARQL Query Language for RDF). By setting filters, only relevant information will be passed to the user, maintaining a machine interpretable format. A general advantage of ontologies with regards to data management and collaboration of diverse teams is that information retrieval is supported for both humans and machines. This decreases the probability of misconceptions regarding the same bit of information and allows domain experts to readily share their knowledge in a standardized way. Furthermore, ontologies guarantee a certain level of quality in terms of consistency and approval. Automated queries moreover help to detect specific bits of information. By visualizing the graph structure in interactive interfaces, users can benefit from decision support and comprehensive search operations. (DuCharme 2013; Lacy 2006)

3 Approach Description

The semiconductor industry with its strong connection to digitalization approaches has the capacity to lead this disruptive change towards a data-driven smart development and manufacturing environment. In order to manage current issues such as described above, we propose the usage of Semantic Web technologies. This toolset enables the definition and maintenance of a controlled vocabulary of entities, including roles, processes and objects. Incremental improvements by simulation models may be reached by applying the powerful variety of Semantic Web technologies, especially incorporating a well-defined vocabulary of involved entities. Linked and openly available data sets are hence readable and interpretable by both machines and humans, which enables improved collaboration between computers and humans (Hitzler et al. 2009). The Digital Reference is a holistic Semantic Web based ontology that represents entities being relevant for semiconductor manufacturing and manufacturing employing semiconductors, including connections to related objects. Being developed as a digital twin, the Digital Reference accompanies the semiconductor manufacturers with their supplier and customer tiers during the digital transformation. As a digital representation of the physical entities in a semiconductor supply chain, the Digital Reference may support simulation models that are initially built on the MIMAC data set by adding more semantically unique content.

As a first step, the generic data model is depicted as a theoretical UML diagram and further clustered with regards to data content (Laipple et al. 2018). Moreover, solely entities that are related to supply chain simulation are chosen to be modeled as a database schema, including the MIMAC data. In the next step, Semantic Web technologies are applied. The fab data sets are transferred into a model based on the supply chain planning process for semiconductors and engineering integration. Facilitated by appropriate supply chain simulation models, a decision support framework for semiconductor value chains is developed. Similar to the initial MIMAC purpose, the identification of main independent variables is enabled. Increasing the granularity of the Digital Reference reveals four levels, as depicted in Fig. 3. The levels are related to simulation levels, with increasing magnification from top to bottom, the Digital Reference corresponding to the fourth level, accordingly. Each of the other layers contains smaller, partly problem- or domain-specific ontologies that serve as decision support knowledge bases exactly where needed.

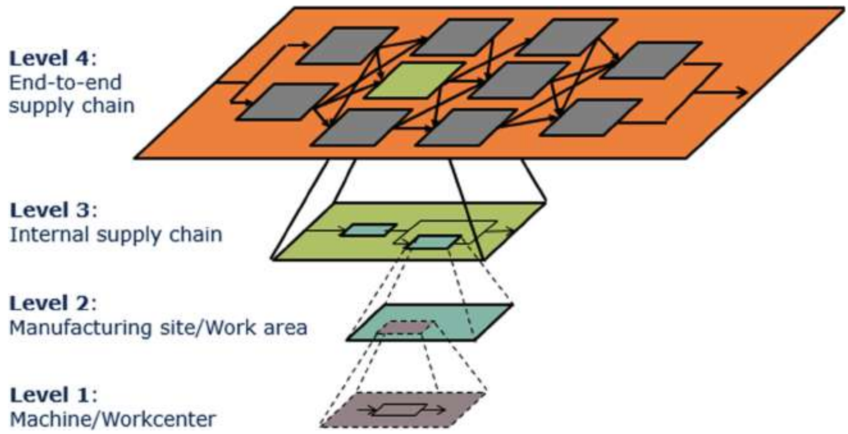

Fig. 3. The four levels of semiconductor operations. Own presentation.

The Digital Reference provides a fundamental understanding of the complex supply chain environment and facilitates inter-company collaboration. It incorporates a set of interrelated data points that is enhanced by reasoning and therefore provides insight into newly discovered implicit knowledge. Moreover, the Digital Reference emphasizes the connection between forecasting algorithms and the planning structure. For planning tasks in the supply chain domain with special focus on the manufacturing processes, overall communication between stakeholders is improved and a clear relation between assumptions in the planning system and the actual outcome may be provided. Depicted as a holistic and general platform, the alignment and interaction between algorithms and processed data is guaranteed.

4 Preliminary Results

By expanding the scope of the MIMAC data set and making it relevant in today's conditions, complexity is reduced and redundancy is eliminated by terms of the new model. Furthermore, the model follows a generic design and is hence flexible for adjustments and upcoming use cases. With the proposed model, high volume of data is being accessed and processed, facilitating knowledge extraction and information sharing. By splitting the holistic model up into smaller ontologies – each for a different scope, respectively – only relevant data is accessible for the respective user that can have several roles such as management, planning or operating. For instance, Fig. 4 shows a Level 3 ontology that represents an internal supply chain of a semiconductor manufacturer. Taking a closer look, one of its nodes is defining the actual demand including its properties due date, quantity and type for instance. Hence, the generic MIMAC model is enriched with detailed relationships and properties. By defining sub-ontologies while still maintaining a holistic top level ontology, the balance between a broad-scale extensibility and a very detailed specification is given. Since cooperation between enterprises within the supply network is increasing, both simulation and Semantic Web Technologies are considered crucial for enabling high volume data processing, real-time access to relevant knowledge and solution generation.

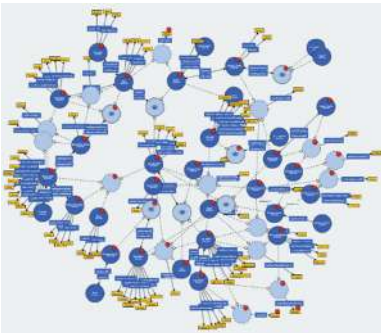

Fig. 4. Semiconductor operations third level ontology: internal supply chain. Own representation.

MIMAC data sets, utilized by agent-based or discrete-event simulation models still serve as a broadly adopted support solution in semiconductor manufacturing domains. Taking into account the digitalization efforts that are funded by large projects in the European Union, the next generation of semiconductor development, manufacturing, supply chain planning and deployment solutions is required to incorporate strategies to overcome Big Data hurdles. It is necessary to guarantee a seamless corporation, allow for automatic decision making processes and let machines and humans reliably interact in real-time. The Semantic Web based Digital Reference serves as a role model for tackling the digitalization challenges of the future. Extending the MIMAC standard in this direction, the semiconductor Plan and Make processes are supported, which hence can lead to a further overall fab capacity increase. This is equivalent to the initial goal of the joint JESSI/SEMATECH standard MIMAC, transferred to future demands.

5 Outlook

Despite the far-reaching advantages of this method, there are still open issues that may be addressed in future research. On the detailed levels of the model, the need for more data properties (such as key performance indicators for simulation) and greater availability of data for import arises. One option to overcome the latter issue is using qualified synthetic data. Yet, specific knowledge that is required for extracting it in favor of decision support needs to be implemented by the respective experts. At higher levels the need for linking projects or business partners arises. Furthermore, specific use cases are necessary to validate the approach and improve it further. Not every level is yet described in a holistic way, which leads to the aim of integrating a semiconductor planning and development ontology in the course of the iDev40 project. Moreover, a critical task is to ensure real-time accessibility of the high volume of data that is generated along the entire product lifecycle. Emerging technologies such as Semantic Web, Artificial Intelligence (AI) or Deep and Machine Learning (DL, ML) approaches will play a crucial role in manufacturing environments of the future.

References

Baumgärtel, H., Ehm, H., Laaouane, S., Gerhardt, J., Kasprzik, A. (eds.): Collaboration in supply chains for development of CPS enabled by semantic web technologies. In: 14th International Conference on Modeling and Analysis of Semiconductor Manufacturing (MASM) at Winter Simulation Conference 2018. Gothenburg, Denmark (2018)

Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. Sci. Am. Mag. **284**, 34–43 (2001)

Bolstorff, P.A., Rosenbaum, R.G., Poluha, R.G.: Spitzenleistungen im Supply Chain Management. Ein Praxishandbuch zur Optimierung mit SCOR; mit 33 Tabellen. Springer, Berlin (2007)

Chien, C.-F., Dauzere-Peres, S., Ehm, H., Fowler, J.W., Jiang, Z., Krishnaswamy, S., et al.: Modeling and analysis of semiconductor manufacturing in a shrinking world: challenges and successes. In: 2008 Winter Simulation Conference (WSC), Miami, FL, USA, 07–10 December 2008, pp. 2093–2099. IEEE (2008)

DuCharme, B.: Learning SPARQL: Querying and Updating with SPARQL 1.1, 2nd edn. O'Reilly, Cambridge (2013)

Fowler, J., Robinson, J.: Measurement and Improvement of Manufacturing Capacity (MIMAC). Final Report. In Technology Transfer # 95062861A-TR (1995)

Hassoun, M., Kalir, A.: Towards a new simulation testbed for semiconductor manufacturing. In: 2017 Winter Simulation Conference (WSC), Las Vegas, NV, 03–06 December 2017, pp. 3612–3623. IEEE (2017)

Hitzler, P., Kroetzsch, M., Rudolph, S.: Foundations of Semantic Web Technologies. Taylor & Francis Ltd. (Chapman & Hall/CRC Textbooks) (2009)

Lacy, L.W.: Interchanging Discrete Event Simulationprocess Interaction Models Using the Web Ontology Language - OWL. Dissertation. University of Central Florida. College of Engineering and Computer Science (2006)

Laipple, G., Dauzere-Peres, S., Ponsignon, T., Vialletelle, P.: Generic data model for semiconductor manufacturing supply chains. In: 2018 Winter Simulation Conference (WSC), Gothenburg, Sweden, 09–12 December 2018, pp. 3615–3626. IEEE (2018)

Mittler, M., Schoernig, A.K.: Comparison of dispatching rules for reducing the mean and variability of cycle times in semiconductor manufacturing. In: Inderfurth, K. (ed.) Operations Research Proceedings 1999, Selected Papers of the Symposium on Operations Research (SOR 1999), Magdeburg, 1–3 September 1999 pp. 479–485. Springer, Berlin (2000)

Mönch, L.: Agentenbasierte Produktionssteuerung komplexer Produktionssysteme. Deutscher Universitats-Verlag GWV Fachverlage GmbH, Wiesbaden (Wirtschaftsinformatik) (2006)

Mönch, L., Uzsoy, R., Fowler, J.: A survey of semiconductor supply chain models part I. Semiconductor supply chains, strategic network design, and supply chain simulation. Int. J. Prod. Res. **56**(13), 4524–4545 (2018)

Oliver, R.K., Webber, M.D.: Supply-chain management: logistics catches up with strategy. In: Klaus, P., Müller, S., (eds.) The Roots of Logistics, pp. 183–194. Springer, Heidelberg (2012)

Pfund, M.E., Balasubramanian, H., Fowler, J.W., Mason, S.J., Rose, O.: A multi-criteria approach for scheduling semiconductor wafer fabrication facilities. J. Sched. **11**(1), 29–47 (2008). https://doi.org/10.1007/s10951-007-0049-1

Singh, R., Mathirajan, M.: Experimental investigation for performance assessment of scheduling policies in semiconductor wafer fabrication—a simulation approach. Int. J. Adv. Manuf. Technol. **99**(5–8), 1503–1520 (2018). https://doi.org/10.1007/s00170-018-2414-y

Semiconductor Operations

Enhancing Prediction Quality of Fab Simulation by Advanced Cycle Time Modelling

Maximilian Dilefeld, Sebastian Rank(✉), and Thorsten Schmidt

Chair of Logistics Engineering, Technische Universität Dresden, 01069 Dresden, Germany
sebastian.rank@tu-dresden.de

Abstract. Simulation in the Semiconductor industry is an established method for planning greenfield projects and fab extensions as well as optimizing existing fabs and testing different scenarios for machine distribution with a constantly changing product mix. This paper discusses the importance of transportation and handling times within these simulations. Up to now, respective transportation and handling times are only considered marginally, oversimplified or are even neglected—mostly due to simulation run time performance. From experience this approach is known as error prone in respect to get valid simulation results. This paper discusses and presents strategies to increase the prediction quality of fab simulation models (e.g. for cycle time) through a more detailed modelling of transportation times. Data mining is considered as a possible approach for generating useful information with less effort then manual modeling.

Keywords: Material flow simulation · Fab simulation · Data mining · Knowledge discovery in databases · Cycle time · Transport system · AMHS

1 Motivation

Semiconductor fabrication plant (short: fab) simulation is an established approach for analyzing the capacity of an entire wafer fab and forecasting e.g. lot cycle time. In general simulation can be used for planning greenfield projects, existing fabs and fab extensions against the background of meeting new productivity requirements as well as testing different scenarios for machine distribution with a constantly changing product mix (Fowler et al. 2015; Rozinat et al. 2009).

This paper discusses the importance of transportation and handling times within fab simulation. Up to now, respective transportation and handling times are only considered marginally, oversimplified or are even neglected—mostly due to simulation run time performance. From experience this approach is known as error prone in respect to get valid simulation results. This paper discusses and presents strategies to increase the prediction quality of fab simulation models (e.g. for cycle time) through a more detailed modelling of transportation times.

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 35–42, 2020.
https://doi.org/10.1007/978-3-030-48602-0_4

2 State of the Art

Simulation is an established tool in semiconductor manufacturing with a wide variety of applications. In the semiconductor domain, usually two types of simulation are present: fab simulation and AMHS simulation (AMHS: automated material handling system). Fab simulation concerns about tools and their corresponding jobs which are to be processed on the tools in a specific order. AMHS simulation has the transportation and handling system in focus. So it realizes the transportation jobs between the tools.

Depending on the pursued goals the models representing the real-world systems are developed with specific aspects in mind (Fowler et al. 2015). Therefore, different models with a varying level of detail can exist to mimic the behavior of the whole system or to focus on specific parts of the system. Usually, these simulation models are created manually, which is time-consuming and prone to errors (Rozinat et al. 2009). Such models are based on human perception on reality rather than reality itself. The necessary meeting of assumptions and the selection of suitable key figures can be a challenging task.

In the semiconductor domain, usually discrete-event simulation models are applied. They are characterized by a sequence of events that mark changes of the examined system. The system is represented by a number of dynamic and static entities (resources). These resources typically have a limited capacity and the entities handled by these have to be ordered in accordance to specific strategies/rules. If a resource is occupied, entities join a queue to wait for processing. For realistic models the necessary times in which entities get transported from one resource to another are to be derived from the real-world transportation systems. These transportation times have to be considered within simulation studies (Fowler et al. 2015).

The automated material handling system (AMHS) for wafers gathered an essential role in semiconductor manufacturing as the increase in wafer size from 200 mm to 300 mm accompanied an increase of weight from 4 kg to 10 kg of the corresponding wafer carriers (Lin et al. 2003). The carriers used in AMHS are called front opening unified pods (FOUP) and usually have a capacity of 25 wafers. The most common system for transporting and handling FOUPs are overhead hoist transport (OHT)-systems, where fully automated vehicles travel along a unidirectional rail system built near the ceiling of the clean room/manufacturing plant. The different stations are accessed by lifting and lowering the FOUP. During this hoisting process the rail is blocked for other vehicles which can lead to delays and longer transport times. This is especially true when the OHT is working at high utilization.

In general, the AMHS should be designed to never become a limiting factor in semiconductor manufacturing (Jimenez et al. 2005). Therefore, transportation times should have a low impact on the lots' cycle times. In reality this cannot be achieved in every possible scenario due to the constantly changing product mixes (Arisha and Young 2005) and planning with the least amount of resources to meet the expected productivity (Jimenez et al. 2005). At high utilization or overload the impact of transportation and handling times rises significantly. This leads to the conclusion of giving transportation and handling times a significant consideration within the conduction of fab simulation. Surprisingly in traditional fab simulation the AMHS is not modelled at all, with less detail, or transport times are approximated by mathematical methods (Jimenez et al. 2005; Wang and Zhang 2016).

Combining or extending fab simulation models with a simulation model of the AMHS in detail would result in a sophisticated model with equipment in the form of physical models with a large scale, multiple machine types, and complex routing and dispatching strategies. These holistic models are challenging and time-consuming in terms of development, maintenance and application (Jimenez et al. 2008; Wang et al. 2018) and hence often limited by the capability of the used simulation tools. Therefore, very often individual models are used to represent the fab with its resources on the one hand and the AMHS on the other hand (Jimenez et al. 2008).

In this connection Jimenez et al. (2005) describe a common approach for reducing computation time of fab simulation models which tries to consider the AMHS: the application of from-to-matrices. These matrices hold the delivery times between source and destination locations. This method can produce accurate results for studies of fab capacities. The particular from-to-matrices can be obtained by detailed AMHS models that need to be created in advance and with a high expenditure of time. Other options with less effort, such as analytical AMHS models, lack of accuracy.

Considering runtime aspects, in a case study comparing a sophisticated holistic model (fab simulation with a fully integrated AMHS simulation model) with a simplified holistic model (fab simulation with mentioned from-to-matrices as representation for the AMHS) Jimenez et al. (2005) show a performance improvement by the factor 100 when using the mentioned from-to-matrices instead of a fully integrated model. The presented results are supposed to be of high accuracy but presuppose very precise transportation times held by the from-to-matrices as a basis for the simulation.

The time between the different process steps does not only consist of transportation and handling time, but also of the waiting times. Especially the latter varies greatly (Arisha and Young 2005) and depends on a wide range of factors and the dynamics of the system (Wang et al. 2018). These times have to be treated as randomly distributed values. Different statistical distributions can be used to represent these time variations. However, so far there are no universal solutions for the task.

3 Approaches for Enhancing the Accuracy of Transport Times Consideration in Fab Simulation

Even though there are efforts to provide "slim" AMHS models with comparatively high performance (e.g. Rank et al. 2016), holistic simulation models with a reasonable level of detail to represent the wafer fab accurately are often not able to be executed fast enough to allow efficient scenario testing (Jimenez et al. 2008). There are only a few attempts to improve prediction accuracy of fab simulation without complex AMHS modelling. Unfortunately, in general they do not perform very well in terms of accuracy and/or run time performance of the simulation model.

As mentioned, applying from-to-matrices can be a reasonable approach for modelling AMHS within fab simulation. In contrast to utilizing detailed AMHS simulation models, from-to-matrices can also be extracted from transport and waiting times based on historical datasets. When generalizing recorded data, the identification of relevant factors that influence the observed times is of great importance to avoid wrong interpretations and to be able to correctly estimate times for the fab simulation. Obviously, historical data is only available for existing systems. Therefore, this approach is not applicable in the planning phase of new plants.

Besides the common/popular way of combining fab simulation models with AMHS simulation models, Rozinat et al. (2009) propose an alternative approach for generating simulation models which might be useful in the semiconductor domain. In detail they apply process mining techniques to (semi-) automatically generate simulation models (see Fig. 1). The step of a first simulation model which can later be further evaluated and modified can be reached much quicker following this path. Besides the time savings the authors are arguing with an improved representation of reality because the development is based entirely on objective information. Additionally, the steps for generating the simulation model can easily be repeated when the observed system changes. It is pointed out that good knowledge about the examined process is of great importance for drawing the right conclusions. Another requirement is the existence of comprehensive recordings of historical data itself. The completeness and quality of this information have major influence on the resulting simulation model.

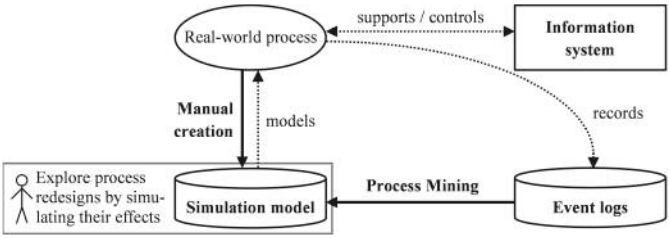

Fig. 1. Manually and (semi-) automatically created simulation models (based on Rozinat et al. 2009)

The term process mining refers to a number of different techniques to extract useful information from event logs (Rozinat et al. 2009). The basis for this are information collected by information systems from the real-world process. These systems usually coordinate the required process steps and record important events related to the performed activities. Process mining is used to discover and represent the causalities between the activities. Besides the performed activity itself event logs usually also include additional information like performance, equipment identifiers, time stamps and more. Often the activities are logged from different perspectives like scheduling, start and completion, resulting in multiple entries with different time stamps or one entry with multiple time stamps. The more information is contained in the data the more factors can be incorporated into the simulation.

Rozinat et al. (2009) point out that the usefulness of simulation models generated by this approach depends on the level of detail. A simple model can approximate the behavior of a system (e.g. routing or waiting time) based on probabilities. But it is not suitable for making predictions based on changes in the system. When starting from historical data usually a specific condition of the system is modeled. Only if different perspectives (e.g. control-flow, data, resources and time) are covered by the model, statements can be made how specific aspects influence the behavior of the whole system.

A similar approach within the scope of simulating transport systems is pursued by Wang and Zhang (2016). Big Data is deemed as a great opportunity to improve the prediction of relevant key figures in production systems. The necessity of a high level of automation and digitalization is met by the semiconductor industry. Within the approach itself, an artificial intelligence (AI) system treats the manufacturing system as a black box and makes predictions on key characteristics of the wafer lots. According to the authors, the applied neural networks can achieve higher prediction accuracy than conventional regression-based models. The considered characteristics used in AI and hybrid AI systems and their identification are presented as a major challenge. It is pointed out, that the quality of the input data of such systems has a great influence on the accuracy of the results. Collected raw data has to be pre-processed, especially when working with Big Data.

As already mentioned, the lot cycle time and hence the wafer fab's performance varies greatly depending on the manufacturing process itself as well as the transport between the different process steps. In this regard Wang et al. (2018) deal with the identification of candidate factors for cycle time forecasting based on historical datasets. The correlation of these factors on each other and on cycle time is analyzed with a regression-based model. From 774 candidates 108 factors are deemed correspondent to the cycle time for the presented system. The factor selection is stepwise from highest correlation to lowest. The factors are integrated successively into the forecasting model. The prediction error is calculated. Factors are incrementally added to the model until the subset of factors is sufficient. A relevant improvement in mean relative deviation is shown in dependence of the number of factors.

In both presented papers, the authors mentioned are using data driven concept to build complete, holistic models. Similarly, these approaches can be used to consider only a specific aspect of the system. In this case it might be possible to generate the mentioned from-to-matrices that afterwards can be implemented into established simulation methods which may already exist.

4 Extracting Transport Times from Historical Data

So far, applying from-to-matrices with corresponding transportation times is the choice for extending fab simulation models in order to get more realistic simulation results. This is mostly because none of the presented alternative approaches have been applied in the semiconductor domain and hence their capabilities are rather unclear. So, it will be investigated to what extent from-to-matrices can be extracted from historical data. In this regard there are multiple existing procedures which can be used for data analysis and exploration. The procedure applied in regard to this paper is based on the knowledge discovery in databases (KDD)-process introduced in 1996 by Fayyad et al. (for additional information see Ristoski and Paulheim 2016). For better understanding the KDD-process and hence the extraction of the from-to-matrices is described in the following. It starts with raw data, from which valuable knowledge and insights are gained following five steps: selection, preprocessing, transformation, Data Mining and evaluation and interpretation (Fig. 2).

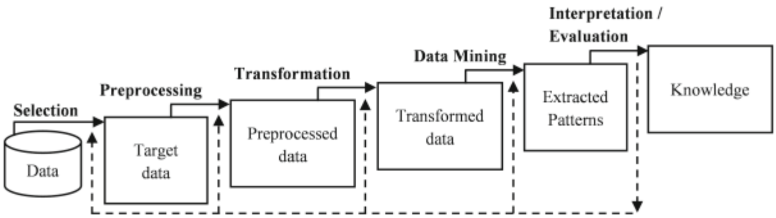

Fig. 2. KDD process (based on Ristoski and Paulheim 2016)

- Selection: In the first step an understanding of the application domain was developed, relevant prior knowledge was collected (which has been done in the preceding sections of this paper) and the goal of the end user was defined. In our case the goal was the generation of from-to-matrices containing the transport times faster than manually building an AMHS simulation model and with higher or at least similar accuracy. Based on this information the target data was chosen with reasonable sample size and relevant variables. Here the target data is a database containing transport logs from an existing AMHS. Knowledge about the different variables was collected in cooperation with experts on sight.
- Preprocessing: This step includes dealing with missing values, duplicates noise and error in the data. The quality of the dataset has to be analyzed. Incomplete or faulty entries have to be removed from the database or corrected if possible.
- Transformation: The data has to be converted into a form the later used algorithms can work on. The specific form dependents on the used programs/the used programming language. In our case the data was already tabular, which in most cases is an appropriate form. The raw data as well as the results later are stored in a MySQL database for easy access. Python with the "pandas-package" for data analyses is a common choice as a programming environment. The DataFrame-object from pandas is a convenient structure for working with tabular data.
- Data Mining: In this step different methods will be applied to the data to achieve the goals defined in the first step. The expectations and goals of the end user have to be considered in selecting appropriate methods and algorithms. One example is the level of detail. More simple models are less accurate but easier to interpret and therefore preferred in some circumstances.
- Evaluation and interpretation: In the last step the finding of the data analysis will be examined with respect to their validity. The usefulness of the gained knowledge will be evaluated. Visualization and presentation of the results is also a typical component of this step.

5 First Findings

An extract of log data from the AMHS was provided by Infineon Dresden. This data will be the basis for our data analysis on a larger scale. The structure of the provided data was analyzed and will further be discussed with experts from Infineon. The log files are available in tabular form with a total amount of 25 attributes. For the investigation particularly the different time stamps, source and destination location, the carrier-ID as well as the priority are of interest. Each row of the database represents a single transport operation.

One problem in respect to model from-to-matrices, which in a first analysis became apparent, is that many transport operations start or end at a storage unit. This phenomenon of low tool-to-tool ratios is well known in the semiconductor domain and has also been covered in a couple of other papers. Fischmann et al. (2008) assume a tool-to-tool-ratio as low as 20% to 40%. Jimenez et al. (2010) refer to a tool-to-tool-ratio between 25% and 47%. Heinrich et al. (2008) on the other hand assume a tool-to-tool ratio of 60%. The storage units have no identifiers in the database. Therefore, the location of the individual storage units is unknown. One possible solution to this issue is the combination of the transport operation tool-storage and storage-tool according to the pattern tool-storage-tool using the carrier-ID. This procedure will be discussed in cooperation with the experts from Infineon Dresden.

6 Results

The examined literature indicates that the quality of efficient prediction models is highly dependent on the determination of precise time tables for simulating the AMHS. It is shown that data driven concepts are likely to improve the prediction quality and are of interest of further research.

7 Discussion/Implications

A key finding of the research and the literature review is the identification of a lack of methods to accurately apply transportation and handling times within fab simulation. Even though transportation and handling times are essential to get high accuracy simulation results, several time and performance related challenges lead to the application of non-integrated models. So far, AMHS- and fab simulation models are treated as independent models. To overcome the issue, and as an alternative for holistic models, new approaches to apply transportation and handling times in fab simulation have to be found. Within the project iDev40, the ability of e.g. from-to-matrices and corresponding transportation and handling time distributions will be tested for application.

References

Arisha, A., Young, P.: Simulation in semiconductor manufacturing facilities. In: Fifth International Workshop on Advanced Manufacturing Technologies, vol. 9 (2005)

Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: From data mining to knowledge discovery in databases. AI Mag. **17**, 37–54 (1996)

Fischmann, C., et al.: Buffer management for automated material handling systems in semiconductor industries. In: Louca, L.S., Chrysanthou, Y., Oplatkova, Z., Al-Begain, K. (eds.) Proceedings of the 22nd Conference on Modelling and Simulation, ECMS 2008, pp. 423–427 (2008). http://www.scs-europe.net/dlib/2008/2008-0423.htm

Fowler, J.W., Mönch, L., Ponsignon, T.: Discrete-event simulation for semiconductor wafer fabrication facilities: a tutorial. Int. J. Ind. Eng. **22**, 5 (2015)

Heinrich, H., et al.: Pursuing the increase of factory automation in 200 mm frontend manufacturing to manage the changes imposed by the transition from high-volume low-mix to high-mix low-volume production. In: 2008 IEEE/SEMI Advanced Semiconductor Manufacturing Conference, pp. 148–155 (2008)

Jimenez, J.A., Mackulak, G., Fowler, J.W.: Efficient simulations for capacity analysis and automated material handling system design in semiconductor wafer fabs. In: Proceedings of the Winter Simulation Conference, Orlando, FL, USA, pp. 2157–2161. IEEE (2005)

Jimenez, J.A., Mackulak, G., Fowler, J.W.: Levels of capacity and material handling system modeling for factory integration decision making in semiconductor wafer fabs. IEEE Trans. Semicond. Manuf. **21**(4), 600–613 (2008)

Jimenez, J.A., et al.: AMHS factors enabling small wafer lot manufacturing in semiconductor wafer fabs. In: Proceedings of the Winter Simulation Conference, WSC 2010, Baltimore, Maryland, pp. 2575–2585 (2010). http://dl.acm.org/citation.cfm?id=2433508.2433826

Lin, J.T., Wang, F.-K., Wu, C.-K.: Simulation analysis of the connecting transport AMHS in a wafer fab. IEEE Trans. Semicond. Manuf. **16**(3), 555–564 (2003)

Rank, S., et al.: The correct level of model complexity in semiconductor fab simulation—lessons learned from practice. In: 2016 27th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), pp. 133–139 (2016)

Ristoski, P., Paulheim, H.: Semantic Web in data mining and knowledge discovery: a comprehensive survey. Web Semant. Sci. Serv. Agents World Wide Web **36**, 1–22 (2016)

Rozinat, A., et al.: Discovering simulation models. Inf. Syst. **34**(3), 305–327 (2009)

Wang, J., Zhang, J., Wang, X.: A data driven cycle time prediction with feature selection in a semiconductor wafer fabrication system. IEEE Trans. Semicond. Manuf. **31**(1), 173–182 (2018)

Wang, J., Zhang, J.: Big data analytics for forecasting cycle time in semiconductor wafer fabrication system. Int. J. Prod. Res. **54**(23), 7231–7244 (2016)

Visualization of Automated Material Handling System Components in Semiconductor Industry over the Lifecycle

Patrick Boden[1], Sebastian Rank[1(✉)], Thorsten Schmidt[1], and Martin Däumler[2]

[1] Chair of Material Handling and Logistics Engineering, TU Dresden, Dresden, Germany
{patrick.boden, sebastian.rank}@tu-dresden.de
[2] Fabmatics GmbH, Dresden, Germany

Abstract. Automated Material Handling Systems are highly complex which makes planning, monitoring and optimization a challenging task. Operators of an Automated Material Handling System (AMHS) have access to extensive data from the real system and additional data sources like material flow simulation studies. Until now, these data have mainly been evaluated statically in the form of predefined diagrams and reports. The aim of this article is to present a concept that allows interactive data visualization of semiconductor transport systems from planning up to their operational phase. Thereby the application of approaches from the field of Visual Analytics is of special interest. A high potential for a better understanding of the highly complex AMHS is seen here.

Keywords: AMHS · Lifecycle · Data visualization · Visual analytics

1 Motivation

Planning, monitoring and optimizing an Automated Material Handling System (AMHS) is highly challenging. Due to the diversity of the production equipment and their arrangement, efficient handling of logistical processes is of great importance. In order not to risk the production schedule high throughputs and short delivery times need to be achieved. To meet these requirements, AMHS have been established in 300 mm frontend factories. Ceiling-mounted Overhead Hoist Transport (OHT) systems are the most common systems in that area (see Hammel, Schmidt and Schöps 2012), but also other systems like Automated Guided Vehicles (AGV), Rail Guided Vehicles (RGV) and Conveyors are used (see Geng 2005).

Operators of AMHS have access to extensive data from the real system and additional data sources like material flow simulation studies (see Schmaler et al. 2017) or technical monitoring systems (see Siegel, Zhakov, Zhu and Schmidt 2018). Up to now, these data has mainly been evaluated statically in the form of predefined diagrams

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 43–53, 2020.
https://doi.org/10.1007/978-3-030-48602-0_5

and reports. They are often created by employees, such as IT experts or support staff, who are not deployed with the evaluation. This approach makes it difficult to react rapidly to current issues and delays necessary planning and optimization tasks, due to the delay of essential adjustments of reports.

The aim of this article is to present our concept "AMHSviz" which allows interactive data visualization of semiconductor transport systems from planning up to their operational phase. The concept should allow the use of data from typical AMHS sources and should be available for analysis by different stakeholders (e.g. operators, sustainers, shift leaders, engineers, and managers). Thereby the application of approaches from the field of visual analytics is of special interest.

2 Related Work

2.1 Data Visualization and Visual Analytics

As a result of progress in information technology operators of technical systems, such as transport systems, increasingly have access to large amounts of raw data for the analysis of system states. Data streams from different sources are often linked so that they can be described as demanding in the dimensions volume, variety and velocity and hence as Big Data (see Fasel and Meier 2016). In such an environment data visualization is a powerful but also challenging technique to explore data sets and communicate findings (see Keim, Qu and Ma 2013). A new trend towards visual analytics focuses mainly on the interaction- and algorithm-based exploration of large data sets. A high potential for a better understanding of the highly complex AMHS is seen here.

Data Visualization. Data visualization can be defined as "the representation and presentation of data to facilitate understanding" (see Kirk 2016). This definition already indicates the importance of data acquisition and preparation as preliminary stages of the final visualization. This aspect is reflected in the process model for the visualization of data from Fry (2004, see Table 1). In seven steps, the process of data visualization is described, whereby steps one to four are dedicated to the area of data processing and the remaining steps can be dedicated to the visualization itself. In the last step of the model the possibility of interaction is mentioned to give the user the possibility to control and explore the underlying data. In order to achieve visualization, it is necessary to go through the individual steps. In each step methods like from the field of statistics are selected manually. This makes classical data visualization a rather static procedure, which complicates the dynamic exploration of complex data structures.

Table 1. Process model for data visualization (based on Fry 2004)

Step	Subject matter	Related field
(1) acquire (2) parse	• The matter of obtaining the data • Providing some structure around what the data means and ordering it into categories	computer science
(3) filter (4) mine	• Removing all but the data of interest • The application of methods from statistics or data mining, as a way to discern patterns or place the data in mathematical context	mathematics, statistics and data mining
(5) represent (6) refine	• Determination of a simple representation, whether the data takes one of many shapes such as a bar graph, list, or tree • Improvements to the basic representation to make it clearer and more visually engaging	graphic design
(7) interact	• The addition of methods for manipulating the data or controlling what features are visible	information visualization and human computer interface

Visual Analytics. In contrast to conventional visualization approaches, visual analytics allows a more dynamic analysis of the data. Collaborative work is made possible and thus the cooperation between the stakeholders of an AMHS can be improved. Visual analytics refers to techniques that automatically process data and present it in the form of interactive graphics (see Thomas and Cook 2005). Automatic data processing and the application of knowledge discovery algorithms are the central points of this approach (see Keim et al. 2008). Combined by interaction mechanisms this creates the possibility to react flexibly to operational and strategic questions and to use human abilities such as expert knowledge or human perception for the analysis (see Fig. 1). An analysis based on the visual analytics process is characterized by a constant change between visual interaction and automatic data processing.

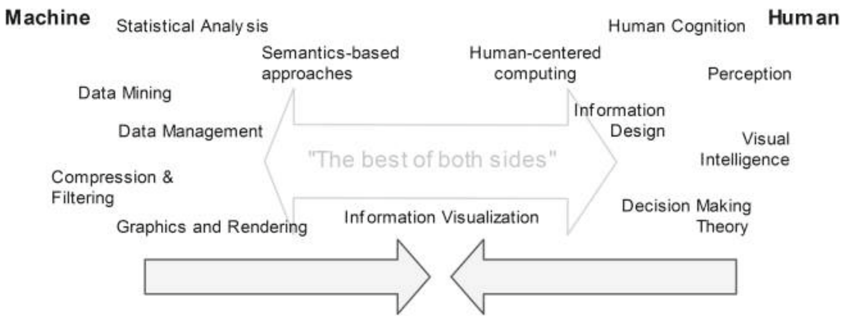

Fig. 1. Scientific disciplines integrated by visual analytics (Keim et al. 2008)

By using this approach in the context of AMHS planning and operating, the visual exploration of system specific data should allow a better understanding of the system and its dynamics. Since the field of visual analytics is still under development there is no commercial software for the analysis of an AMHS available yet. As a first step, towards a visual analytics application, a software tool for data visualization shall be created which allows the connection of AMHS specific data sources and the design of interactive visualizations. In general, the data is captured automatically and pre-processed using basic algorithms like mean value calculation. Our aim is to implement high sophistic algorithms from the fields of data mining and statistical analysis at a subsequent stage to allow a more diverse analysis.

2.2 AMHS Visualization

For the visualization of AMHS performance and corresponding key figures, experts from practice and scientific literature mainly discuss established diagram types such as bar charts, box plots and line diagrams based on statistical data processing. Typical examples can be found in the contributions of Sokhan-Saj, Gaxiola, Mackulak and Malmgren (1999), Potoradi, Boon, Mason, Fowler and Michele (2002) or Hammel, Schmidt and Schöps (2012). For the location dependent visualization of data some authors also use heat maps (see Kortus, Däumler and Schmidt 2018). In this connection, Erler, Lübke, Loose and Hammel (2018) introduced a so called "real time monitor".

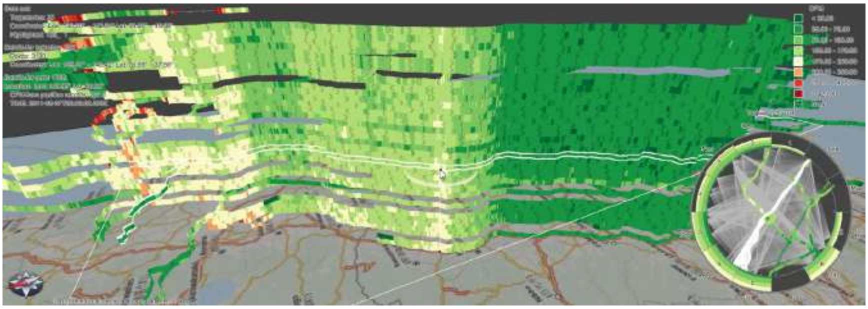

Fig. 2. Trajectory based visualization (Tominski, Schumann, Andrienko and Adrienko 2012)

The analysis of AMHS-specific literature reveals that novel, sophisticated visualization concepts such as the trajectory walls (see Fig. 2) or approaches from the field of visual analytics have not been established yet, even if necessary raw-data is available. In the semiconductor domain, data sources which provide information about transport orders, movement logs and error messages are the most important sources in order to judge about the condition of the AMHS. Furthermore the investigation of the AMHS vehicles, carriers and the remaining technical infrastructure are particularly objects of interest for a detailed analysis. Since AMHS typically have several hundred vehicles employed, the amount of data available is extensive.

As AMHS are very complex and come along with a long life cycle, different user groups (as already mentioned, e.g. operators, sustainers, shift leaders, engineers and managers) are involved in the planning and control process. Thus different user roles with different views and demands on corresponding AMHS-data emerge. In consequence different visualization solutions in accordance to the users' needs are required. In this context expert interviews have been conducted and requirements for visualization solutions were identified (see Rank, Boden and Schneider 2018). Some of the most important findings are the low hardware requirements. Furthermore, visualization solutions should also be device independent and responsive. Based on these findings, the following section presents and discusses possible software tools.

2.3 Visualization Software Tools

The increasing importance of the subject data visualization can be seen in scientific literature (see Tam and Song 2016) as well as in the increasing use of advanced software such as QlikView or Tableau. Both tools allow interactive visualization as well as exploration and are common in semiconductor industry. Since it is rather unclear, if the capabilities of the tools meet all the requirements of the mentioned user groups a systematic evaluation becomes necessary. This is particularly true because the variety of available software in the field of visualization is large and the tools have a different focus.

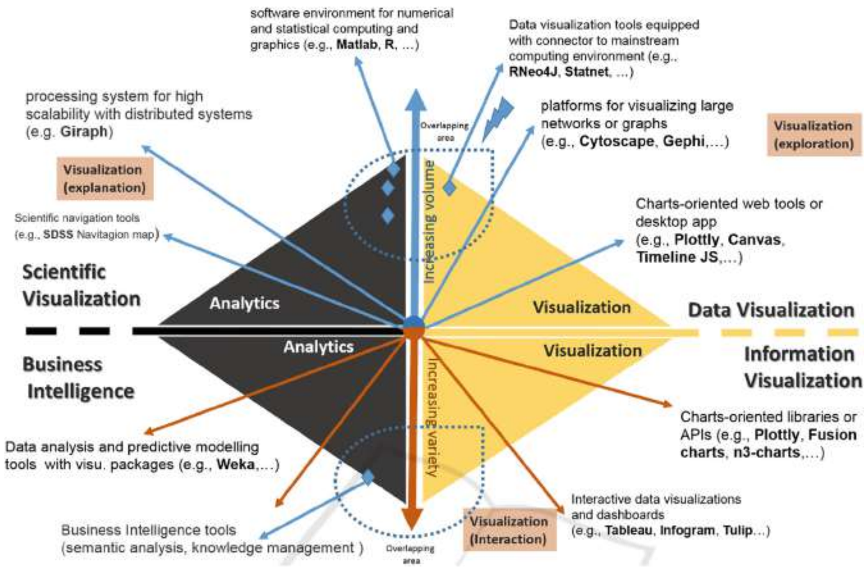

Fig. 3. Big Data visualization tools (Caldarola and Rinaldi 2017)

One possible option of such a taxonomy is illustrated by Caldarola and Rinaldi (2017, see Fig. 4). As a result, the software solutions investigated by the authors on the one hand differ in respect to their functional scope in the areas of analytics and visualization (x-axis). On the other hand, the software can be divided in accordance to the nature of the data that can be processed with the application. The applications are assigned on the basis of the data properties volume and variety (y-axis). Applications from the field of visual analytics can be assigned to the overlapping area along the variety axis. Due to this systematization, it is possible, to distinguish applications with extensive functions in the area of statistics (e.g. R) from applications with a clear focus in the design of interactive graphics (e.g. Tableau).

In their analysis, Caldarola and Rinaldi (2017) evaluated a variety of visualization solutions that can be used in the context of Big Data. Thereby, they not only focused on programs intended for end users, but also on software libraries that can be used by developers to create own illustrations and applications. The overview demonstrates a large variety of development tools and indicates that web-based applications are increasingly being developed. They can be used almost independently of an operating system. Typical visualization structures of web-based applications are charts, graphs or maps. An excerpt from the evaluated applications is shown in Table 2.

According to Fig. 3 and Table 2 most of the commercial visualizations tools (e.g. Tableau) are easy to use. They have a user-friendly user interface and they allow fast plot/graph generation and saving by simply clicking. On the other hand, they are restricted in terms of designing and defining new, highly sophisticated and self-designed plots. There are tools/libraries which focus on visualization and there are tools/libraries focusing on analytics. Apart from R, which has some major downsides in respect to plot interaction, none of the assessed tools show major strengths in both categories.

Table 2. Excerpt of the visualization tool survey (based on Caldarola and Rinaldi 2017)

Name	Usage	Software category	Visualization structure	Operating system	License
Tableau	Presentation	Desktop application	Chart, graph, map, etc.	Windows, OSX	Commercial
Infogram	Presentation	Desktop application	Chart, map, image and video	Windows, OSX	Commercial
Plotly	Presentation, developers	Web tool, JavaScript and Python library	Chart, plot, map	Web-based	Commercial, community
D3.js	Developers	JavaScript library	Chart, map	Web-based	Open source
Chart.js	Developers	JavaScript library	Chart	Web-based	Open source

Related to the results of Rank, Boden and Schneider (2018) none of the available tools meets the evaluated requirements. In order to avoid limitations of commercial software solutions (e.g. limited extensibility), the development of an own application called “AMHSviz” was initiated. It will be discussed in detail in the next section “Description of the Software Demonstrator”.

For the implementation of visualizations and interaction mechanisms the software library Plotly was chosen for AMHSviz. Plotly enables the integration of a wide variety of display formats and interaction possibilities. In order to guarantee the integration of highly specialized diagram types (such as trajectory walls) also other libraries can be integrated. The overall application is based on the frameworks Cordova and Electron. A detailed description of the application and its structure can be found in Däumler, Reith, Hochholzer and Schmidt (2018).

3 Description of the Software Demonstrator “AMHSviz”

A first result of the discussion in the sections before is the mentioned tool AMHSviz. It is an extension of a software tool based on the work of Däumler, Reith, Hochholzer and Schmidt (2018). They dealt with the analysis of simulation data of AMHS systems. In addition to the previous work, a possibility for the analysis of data from a real transport system was integrated. For this purpose, established AMHS data sources and possible interfaces were investigated. The data exchange between AMHSviz and the corresponding/dynamic databases can be done statically via database excerpts or log files, as well as continuously via the data exchange format JSON. The data is collected in a central data server and then processed by the visualization component into interactive graphics (see Fig. 5).

For data presentation, the AMHSviz allows the integration of various visualization types customized to the concrete application case. To meet the most important needs, the creation of reusable and stakeholder specific dashboards is possible (see Fig. 4) and first diagram types and interaction techniques were implemented. So far, graphic types such as radar plot, scatter plot or parallel coordinates can be used. The graphics allow different ways of interaction. For example, filters can be set or additional information can be displayed by hovering. As a result, the essential stakeholder requirements for a versatile and interactive visualization can be fulfilled.

The whole concept and a first demonstrator of AMHSviz were tested by applying AMHS-data. For this purpose, data from log files and the Murata AMHS Control System were automatically processed using an own implemented data wrapper. Investigating single vehicles or carrier categories (e.g. test or development) proved to be particularly informative and led, for example, to the identification of inefficient transport flows.

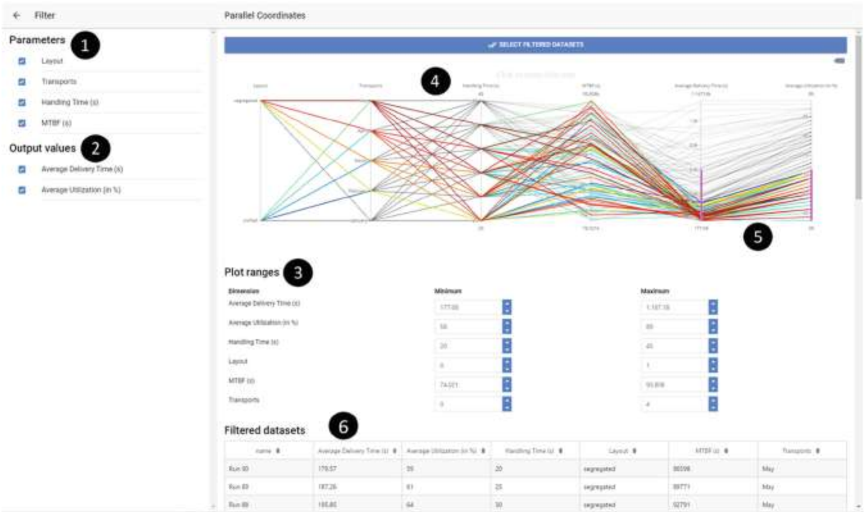

Fig. 4. AMHSviz—Screenshot of the user interface with (1) Selection of input parameters of interest (2) Selection of output values of interest (3) range selection for each axis (4) Parallel coordinates plot of the different data; color adjusted according to the order of the axes and the values on the axes (5) Manual interval selection [pinkish bars] of data according to values of the last two axes (6) List of selected runs (only a part is shown) (Däumler, Reith, Hochholzer and Schmidt 2018).

4 Discussion/Implications

This article describes the related work for the visualization of AMHS related data in the semiconductor environment. Typical AMHS data visualizations, possible software tools and relevant data sources were investigated. Dealing with the high diversity of stakeholders could be identified as an essential requirement for a visualization tool. Due to the high complexity of AMHS and the long life cycle, many people are engaged by the analysis of data. Through the application of concepts from the field of visual analytics, a promising way for the presentation of data and the interactive communication of results is seen. So far such a AMHS specific visualization tool does not exist.

A workflow/concept using an expanded software framework that allows evaluating extensive data collected from AMHS was developed—so called AMHSviz. Approaches from the field of visual analytics were integrated. AMHSviz provides a suitable possibility for different users ("roles") to react flexibly to operative and strategic questions over the life cycle of the system and to employ data from different sources.

Data that can be used for the analysis of transport systems have a strong time- and/or location-related character. One possibility for the further development of AMHSviz is the integration of data visualization approaches that enable a better analysis of such data. A comprehensive overview of visualization techniques especially for time-based data can be found at Aigner, Miksch, Schumann and Tominski (2011).

While comparatively simple algorithms for statistical analysis (e.g. calculation of mean values) have been implemented in the current demonstrator of AMHSviz, the integration of more sophisticated algorithms is planned for further development. In particular approaches from the field of data mining (such as cluster or sequence analysis) are to be examined for their suitability for the AMHS application case. In contrast to the algorithms implemented so far a more diverse analysis is aimed to better understand the connection between different system states. For example, the relationship between anomalies (such as traffic jams and technical disturbances) and the delivery time behavior could be investigated in more detail by customized data analysis and visualizations.

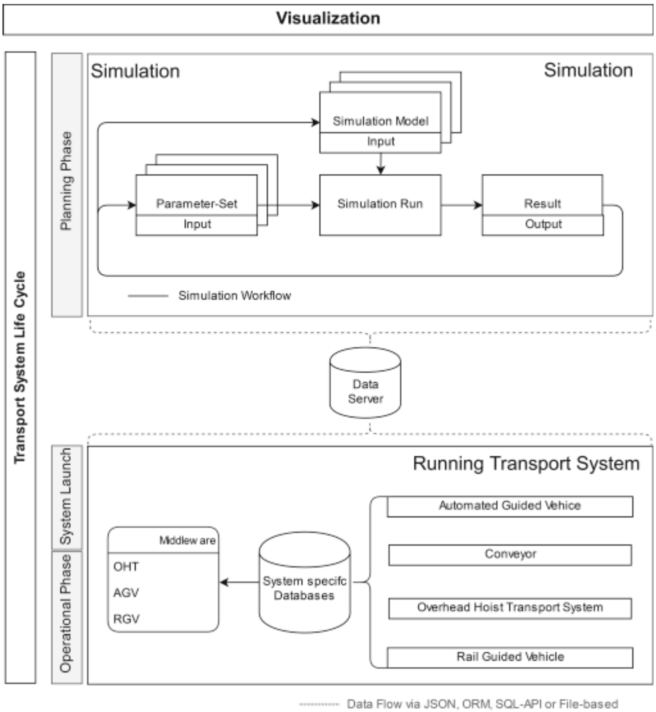

Fig. 5. Lifecycle visualization concept for AMHS (based on Däumler, Reith, Hochholzer and Schmidt 2018)

References

Aigner, W., Miksch, S., Schumann, H., Tominski, C.: Visualization of Time-Oriented Data. Springer, London (2011)

Caldarola, E., Rinaldi, A.: Big data visualization tools: a survey - the new paradigms, methodologies and tools for large data sets visualization. In: 6th International Conference on Data Science, Technology and Applications, pp. 296–305 (2017)
Däumler, M., Reith, K.-B., Hochholzer, S., Schmidt, T.: Approach on evaluating material handling simulation runs under consideration of different target groups. In: ASIM Symposium Simulationstechnik (2018)
Erler, M., Lübke, J., Loose, L., Hammel, C.: AMHS real-time monitor. In: 18th European Advanced Process Control and Manufacturing Conference (APC|M) (2018)
Fasel, D., Meier, A.: Big Data: Grundlagen, Systeme und Nutzungspotenziale. Springer, Wiesbaden (2016)
Fry, B.: Computational Information Design. Doctoral thesis. Massachusetts Institute of Technology, Cambridge (2004)
Geng, H.: Semiconductor Manufacturing Handbook. McGraw Hill Professional, New York (2005)
Hammel, C., Schmidt, T., Schöps, M.: Network optimization prior to dynamic simulation of AMHS. In: Proceedings of the 2012 Winter Simulation Conference, 1956–1966 (2012)
Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., et al.: Visual analytics: definition, process and challenges. In: Information Visualization - Human-Centered Issues and Perspectives, pp. 154–175 (2008)
Keim, D., Qu, H., Ma, K.-L.: Big-data visualization. IEEE Comput. Graph. Appl. **33**, 20–21 (2013)
Kirk, A.: Data Visualisation: A Handbook for Data Driven Design, 1 edn. SAGE Publications Ltd., London (2016)
Kortus, G., Däumler, M., Schmidt, T.: Applying the Discrete Network Design Problem (DNDP) for designing AMHS layouts in semiconductor fabs. In: 29th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), pp. 295–300 (2018)
Potoradi, J., Boon, O.S., Mason, S., Fowler, J., Michele, P.: Using simulation-based scheduling to maximize demand fulfillment in a semiconductor assembly facility. In: Proceedings of the 2002 Winter Simulation Conference, pp. 1857–1861 (2002)
Rank, S., Boden, P., Schneider, G.: Integrated Development 4.0: State of the art of visualization systems in semiconductor industry. ECSEL JU, Dresden (2018)
Schmaler, R., Hammel, C., Schmidt, T., Schoeps, M., Luebka, J., Hupfer, R.: Strategies to empower existing automated material handling systems to rising requirements. IEEE Trans. Semiconductor Manuf. **30**(4), 440–447 (2017)
Siegel, A., Zhakov, A., Zhu, H., Schmidt, T.: Trailer based rail monitoring in overhead hoist transport systems. In: 2018 International Symposium on Semiconductor Manufacturing (ISSM), pp. 1–4 (2018)
Sokhan-Saj, S., Gaxiola, G., Mackulak, G., Malmgren, F.: A comparison of the exponential and the hyperexponential distributions for modeling of move requests in a semiconductor fab. In: Proceedings of the 1999 Winter Simulation Conference, pp. 774–778 (1999)
Tam, N.T., Song, I.: Big data visualization: application in visualizing learning activities. In: Information Science and Applications (ICISA), pp. 399–408 (2016)
Thomas, J., Cook, K.: Illuminating the Path: A Research and Development Agenda for Visual Analytics. IEEE Press (2005)
Tominski, C., Schumann, H., Andrienko, G., Adrienko, N.: Stacking-based visualization of trajectory attribute data. IEEE Trans. Vis. Comput. Graph. **18**(12), 2565–2574 (2012)

A System Dynamics Approach for Modeling Return on Quality for ECS Industry

Bernhard Oberegger[1](✉), Andreas Felsberger[2](✉), and Gerald Reiner[1](✉)

[1] Vienna University of Economics and Business, Vienna, Austria
{bernhard.oberegger,gerald.reiner}@wu.ac.at
[2] University of Klagenfurt, Klagenfurt, Austria
andreas.felsberger@aau.at

Abstract. The Electronic Components and Systems industry (ECS) is characterized by long lead times and high market volatility. Besides fast technological development within this industry, cyclic market up- and downturns are influencing the semiconductor market. Therefore, adequate capacity and inventory management as well as continuous process improvements are important success factors for semiconductor companies to be competitive. In this study, the authors focus on a manufacturing excellence approach to increase front-end supply reliability and the availability of inventory within the customer order decoupling point. Here, development and manufacturing processes must be designed in a way that highest levels of product quality, flexibility, time and costs are reached. The purpose of this study is to explore the impact of return on quality in manufacturing systems. Therefore, multimethod simulation modelling including discrete-event and system dynamics simulation is applied.

Keywords: Multimethod simulation modelling · Return on quality · Semiconductor industry

1 Introduction

The Electronic Components and Systems industry (ECS) is characterized by large lead times and high market volatility. Besides fast technological development within this industry, cyclic market up- and downturns are influencing the semiconductor market (Mönch et al. 2011). Therefore, proper capacity and inventory management are key success factors for semiconductor companies to survive. The effects of rapidly advancing technology cause goods in inventory to depreciate very fast, which may lead to considerable losses in case of a steep market downturn. On the other hand, too little inventory of required products can cause a loss in terms of opportunity costs by not reaching customer satisfaction, service level and product availability if the customer demand cannot be satisfied due to the long term manufacturing lead times. In addition, the majority of fixed assets consists of very expensive front-end equipment. In this context this leads to high costs of wasted capacity (e.g., idling, bad product quality) in fab. Consequently, maximized front-end capacity utilization is an important constraint in semiconductor planning.

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 54–63, 2020.
https://doi.org/10.1007/978-3-030-48602-0_6

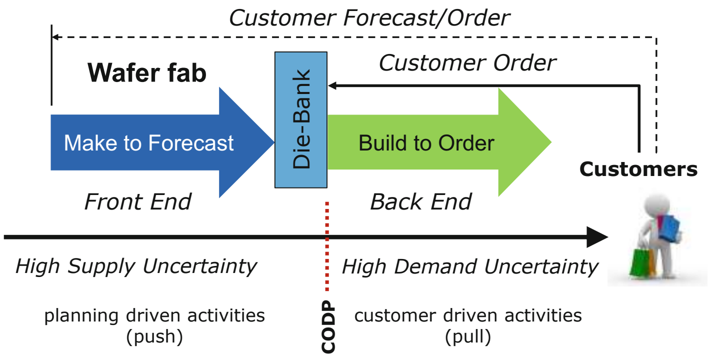

Fig. 1. Uncertainty in semiconductor manufacturing environment

Above, Fig. 1 shows that the internal semiconductor supply chain struggles with a high degree of supply and demand uncertainty. A key success factor of this highly competitive environment is process improvement. Within this work, the authors focus on a manufacturing excellence approach to increase front-end supply reliability and the availability of inventory within the customer order decoupling point (CODP). Manufacturing excellence is represented by superior key performance indicators, i.e. development and manufacturing processes must be designed in a way to reach an optimum between product quality, flexibility, time and costs (Yasin et al. 1999; Plunkett and Dale 1988).

To analyze the effects of process improvements we propose an advanced "return on quality" approach (Rust et al. 1995) considering,

- the accountability of quality efforts to reach a stable process (monitor, control, improve) and the related impact on product quality and process performance (lead time, service level, etc.),
- the complexity of the processes have the risk to spend too much on quality (over engineering) if the root cause of quality problems are not understood,
- quality investments are not equally valid and sometimes a repair process can be used to reach a Zero-Failure Culture (Schneiderman 1986), and
- non-quality costs, i.e., failure costs, productivity losses (material, energy, human resources) or customer dissatisfaction related costs can be high.

The need for scientific research in the field of an efficient and effective process improvement for the Electronic Components and Systems industry is motivated by the following research questions:

- *RQ1: Can production costs be reduced, when critical process deviation are identified at an earlier stage?*
- *RQ2: How does an early detection of defective products affect the lead-time, process quality and performance?*
- *RQ3: Understand the interaction between quality problems and lot sizing.*

2 Theoretical Foundations

For more realistic queueing models, exact analytical solutions become very difficult to achieve. Therefore, researchers usually use approximations to estimate cycle time. Accuracy of approximation models depends greatly on the actual distribution of the inter-arrival times and the service times. Accordingly, various approximations may exist for the same queueing model. For example, Buzacott and Shanthikumar (1993) listed three approximations for the G/G/1 queues. M/M/1 Models for example are not sufficient to accurately describe fab performance because of the oversimplified modelling assumptions. However, these models provide general insight on the relationship between WIP level, capacity and throughput, and provides intuitive directions on how to improve performance. Little's law (Little 1961) can describe a simple and stable process, where the inflow and outflow rates are identical. However, to analyse complex operation processes, require more than the long-run average values set by Little's law. For example, Wu (2005) applied Little's law and the G/G/1 queueing formula to estimate the variance of a simple factory with single toolsets. He found some basic properties that provide managerial insight on how variability affects the production performances. He also pointed out that the variability of bottleneck toolsets should be reduced to achieve fast cycle time. The following conditions apply for a stable process:

$$\rho = \frac{\lambda}{\mu} < 1, \tag{1}$$

where ρ is the utilization, λ is the unit arrival rate and μ is the service rate. The arrival and service rate is defined by

$$\lambda = \frac{1}{t_a}, \tag{2}$$

$$\mu = \frac{1}{t_e}, \tag{3}$$

where t_a is the inter-arrival time of units and t_e is the effective process time. The approximation for waiting time (CT_q) of a single process, is given by

$$CT_q = \frac{\rho}{(1-\rho)}\left(\frac{c_a^2 + c_e^2}{2}\right)t_e, \tag{4}$$

where c_a^2 is the squared-coefficient-of-variation (SCV) of the inter-arrival time (t_a), and c_e^2 is the SCV of the effective process time (t_e). The effective process time include the raw process time (t_s) and all time losses due to setup, breakdown, unavailability of operator, and any other source of variability (Wu 2005; Kingman 1961). The throughput rate of a single server is defined by

$$r = \frac{\rho}{t_e}, \tag{5}$$

and the cycle time is defined by

$$CT = CT_q + t_s. \tag{6}$$

The raw process time includes all the times a unit spends on a server, such as load/unload and processing (Fig. 2).

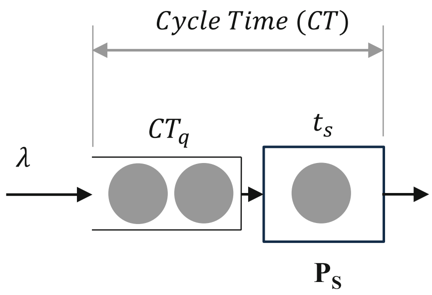

Fig. 2. Cycle time of a single server process

Analytical approaches based on queuing theory serve as a fundamental instrument to explore what-if questions. Queuing models make cycle time estimation based on the assumption of stochastic arrival and service processes. The arrivals (demand) of a unit to the workstation is a stochastic process where inter-arrival times are random variables following specific (unknown) distributions. Based on machine configuration and process requirements, processing times are also stochastic. Products that have already arrived but cannot be processed directly are placed in the waiting queue.

Most operational processes are characterized by dynamic operating conditions and complex interrelationships that cause changes in inputs, operations and outputs. Variability results in non-conformities that have a negative impact on an operations process. Product characteristics and product quality, as well as process attributes (e.g., process time, set-up time, process quality, equipment breakdowns and repairs) are subject to non-conformance (Klassen and Menor 2007). Thus, the reduction of variability improves operational performance, such as throughput, lead-time, customer service, quality, etc. (Hopp and Spearman 2001). Figure 3 shows examples of the effect of variability reduction on the cycle time and capacity utilization:

a) Lower variability leads to a reduced cycle time be unchanged capacity utilization.
b) Lower variability enables higher capacity utilization and therefore a higher throughput rate.

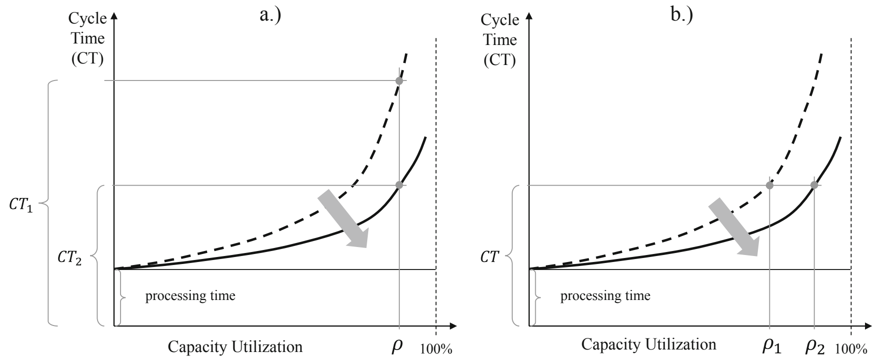

Fig. 3. Impact of process variability

3 Modelling Approach

The purpose of this article is to explore the meaning of return on quality in manufacturing systems. To accomplish this purpose a general model consisting of discrete-event simulation (DES) and system dynamics (SD) simulation models is described. In order to better understand the challenges and benefits of each simulation type the paper starts by reviewing the related work on the different simulation approaches. In general SD can be used to model problems at a strategic level, whereas DES is used at operational and tactical level (Tako and Robinson 2012). Different simulation approaches such as discrete event simulation (DES) and system Dynamics (SD) are excellent tool for visualizing, understanding and analysing the dynamics of manufacturing systems and thus supporting the decision-making process (Borshchev 2013). Both simulation approaches enhance the understanding of how systems behave over time (Sweetser 1999). However, contradictory statements are made regarding the level of understanding that users can gain from using these models. DES is more concerned with detailed complexity, while SD with dynamic complexity (Lane 2000).

These methods tend to be appropriate as long as changing circumstances and time have no significant impact on the managerial problem issue. However, the influence of certain elements on the overall problem definition forces scientists to take a dynamic system into consideration when applying solution methods. In practice, most problems derive from complex and dynamic systems which requires solution methods that are capable of handling these challenges. In the late 1950s MIT professor Jay W. Forrester developed a methodology for modelling and analysing the behaviour of complex social systems. The term “system dynamics” was introduced to the scientific world and attracted particular attention in recent years with the development of high-performance computer software (Rodrigues and Bowers 1996). The basic idea behind system dynamics is the identification and analysis of information feedback structures within a system and their representation with differential and algebraic equations in a simulation modelling environment (Homer and Hirsch 2006). This methodology has been applied to a number of problem domains, such as economic behaviour, biological and medical modelling, energy and the environment, theory development in the natural and social sciences, dynamic decision making, complex non-linear dynamics, software engineering and supply chain management (Angerhofer and Angelides 2000).

System dynamics models consist of feedback loops that represent information based interaction between elements of the simulated system. Feedback loops create a dynamic behaviour of a system, influence the values of stocks and flows and often lead to new unanticipated reactions of the system. They can either be positive (or self-reinforcing) or negative (or balancing). For example the inventory level in a production plant is increased by the flow of produced goods and decreased by the flow of shipped goods. The interaction between feedback loops, stocks and flows enables to replicate dynamic systems in a simulation environment (An and Jeng 2005).

Simulation modelling offers a great potential in modelling and analysing a systems behaviour. In our work, we use DES to simulate the improvement of early quality control. Furthermore, the impact on the holistic system shall be shown by improving the throughput time and the reduction of variability in process stability and quality. The approach presented in this paper not only enables accurate performance evaluation, but also suggests simple real time control rules under which the flexible manufacturing system has a regular and stable behavior when no major breakdown occurs, the production requirements are satisfied, and the bottleneck machines are fully utilized.

First results based on the theoretical background information led to the following selected expected impacts:

- Internal & external improvements based on overall framework.
- Reduction of variability influences the process stability, lead-time and gives potential improvements due to customer satisfaction.
- The company can react faster to uncertainty of order fulfilments and give adequate information to the customer, which results in better service performance and customer satisfaction.
- The company can reduce operating costs, production costs and enhances overall speed of production performance if the inspection and fault detection operations are improved (Felsberger et al. 2018).

Based on the above presented theoretical foundations we now illustrate the developed modelling framework. Our process improvement evaluation model is based on successively using a discrete event simulation model of the relevant processes that provides the basis for estimating the changes of input parameters for a system dynamics model (Fig. 4). The purpose of the discrete-event simulation model is to conduct experiments regarding these processes (e.g., Impact of early detection of systematic errors based on machine data). The DES model can be used to quantify relevant performance indicators (e.g., lead-time, equipment utilization, waiting times, WIP, throughput, service-level). In addition, our discrete-event simulation model analyzes risks such as front-end demand uncertainty as well as the impacts of these uncertainties. Furthermore, the methodology of system dynamics will be taken in consideration to present the dynamic changeovers of quality aspects by improving selected manufacturing processes (Reiner 2005). Existing system dynamics models already focus on the dynamics of quality improvement and quality costs in manufacturing companies (Mahanty et al. 2012; Kiani et al. 2009). Our model approach will be used to analyze the effects of process improvements (development and manufacturing) and analyze its dynamic behavior based on a Systems Dynamics approach in combination with process simulation for manufacturing processes for high-risk environments (Felsberger et al. 2018). This system dynamic model will have the ability to evaluate the impact of input parameters to production process performance as well as stability.

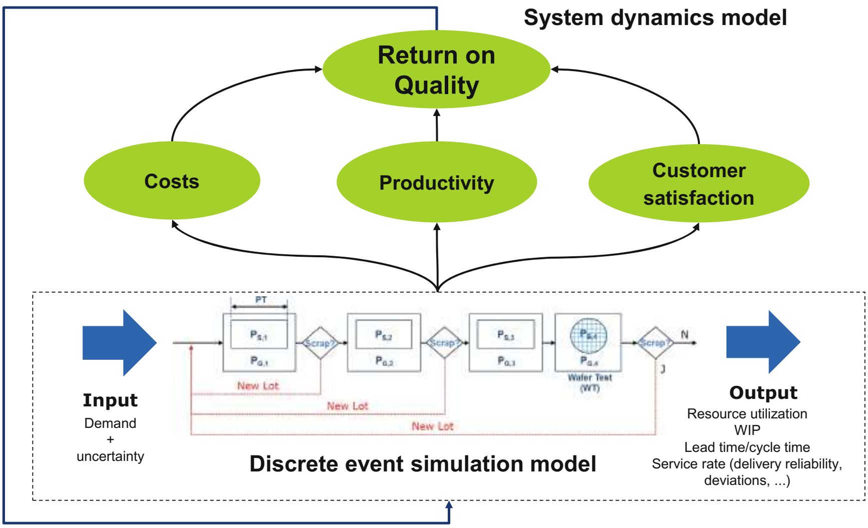

Fig. 4. System Dynamics approach in combination with process simulation

Inspired by a real front-end production process in the ECS industry we first start to develop a simplified model (Fig. 5) that highlights three so-called work centers or tool groups ($P_{G,1}$, $P_{G,2}$, $P_{G,3}$). Between these three selected work centers, additional activities such as transport activities and further processing steps take place. In our model, we consider these additional activities as a black box. The nature of such systems is characterized by several re-entrant flows at different stages in the manufacturing cycle. Single tools are the smallest unit of a work center (P_s). The machines within a work center provide similar processing capabilities. Therefore, the machines of a single work center can be often modeled as identical or unrelated parallel machines (Mönch et al. 2011). Contrariwise, jobs (Lots) cannot always be processed on all machines within a work center related to different process requirements. Finally, the last activity is the wafer test, were every circuit (die) is tested for its functionality. It is performed on the not yet divided wafer to detect faulty circuits. If the defective number of dies per wafer exceeds a certain limit, the entire wafer is discarded. If several wafers of a lot are faulty, the entire batch is discarded. Causes of yield reduction may consist of a common (random) or special (systematic) nature. Incorrect electrical parameters, faulty layers, poor alignment of the layers or poor matching of the chip design to the technology are examples of systematic errors (Hilsenbeck 2005).

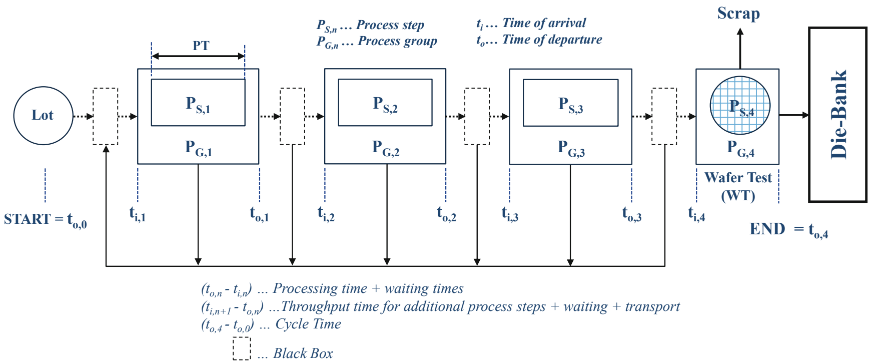

Fig. 5. Topology of the process model

4 Conclusion and Future Outlook

The model can be extended with several details showing effects of process improvement and variability reduction through an increased level of technology, i.e. the early detection of faulty items potentially leads to decrease of cycle times and therefore to an overall improved process performance. We expect that an investment in quality improvement, e.g. appraisal and prevention efforts leads to a positive impact on process quality (Zero Failure Culture) and performance (lead-time, service level). Furthermore, the benefits of investing in return on quality along with process improvement lead to significant impacts on overall manufacturing performance and therefore in profitability and customer satisfaction (Mahanty et al. 2012). Our process model therefore analyses these impacts on a strategical (SD) as well as on a tactical level (DES). Based on the DES model we intend to generate input data for the SD model, which will be subject of future work. In detail, this includes the development of the SD model and linking it to the DES model.

For example, the input parameters for productivity measures might be the potential increase of the overall improvement of speed lead-time. Finally, this improvement eventually ends up in a better customer satisfaction and retention. Costs for example are determined by earlier scrap detection and the reduction of expensive defect density test equipment. Considering Rust et al. (1995) we intend to measure:

- the impact on product quality and process performance (lead time, service level, etc.),
- the risk of over engineering if the root cause of quality problems are not understood,
- the justification of quality investments to reach a Zero-Failure Culture and
- the impact on non-quality costs, i.e., failure costs, productivity losses or customer dissatisfaction related costs.

References

An, L., Jeng, J.J.: On developing system dynamics model for business process simulation. In: Proceedings of the 37th Conference on Winter simulation, pp. 2068–2077 (2005)

Angerhofer, B.J., Angelides, M.C.: System dynamics modelling in supply chain management: research review. In: Proceedings of the 32nd Conference on Winter simulation, pp. 342–351. Society for Computer Simulation International, December 2000

Borshchev, A.: The Big Book of Simulation Modeling: Multimethod Modeling with AnyLogic 6, p. 614. AnyLogic North America, Chicago (2013)

Buzacott, J.A., Shanthikumar, J.G.: Stochastic Models of Manufacturing Systems, vol. 4. Prentice Hall, Englewood Cliffs (1993)

Felsberger, A., Rabta, B., Reiner, G.: Queuing network modelling for improved decision support in manufacturing systems. In: Proceedings of the 20th International Working Seminar on Production Economics, Innsbruck (2018)

Hilsenbeck, K.: Optimierungsmodelle in der Halbleiterproduktionstechnik. Doctoral dissertation, Technische Universität München (2005)

Homer, J.B., Hirsch, G.B.: System dynamics modeling for public health: background and opportunities. Am. J. Public Health **96**(3), 452–458 (2006)

Hopp, W.J., Spearman, M.L.: Factory Physics: Foundations of Manufacturing Management, 2nd edn. Irwin/McGraw-Hill, Boston (2001)

Kiani, B., Shirouyehzad, H., Khoshsaligheh Bafti, F., Fouladgar, H.: System dynamics approach to analysing the cost factors effects on cost of quality. Int. J. Qual. Reliab. Manag. **26**(7), 685–698 (2009)

Kingman, J.F.C.: The single server queue in heavy traffic. Proc. Cambridge Philos. Soc. **57**, 902–904 (1961)

Klassen, R.D., Menor, L.J.: The process management triangle: an empirical investigation of process trade-offs. J. Oper. Manag. **25**(5), 1015–1034 (2007)

Lane, D.C.: You just don't understand me: models of failure and success in the discourse between system dynamics and discrete-event simulation. Working Paper. London School of Economics and Political Sciences (2000)

Little, J.D.: A proof for the queuing formula: $L = \lambda W$. Oper. Res. **9**(3), 383–387 (1961)

Mahanty, B., Naikan, V.N.A., Nath, T.: System dynamics approach for modeling cost of quality. Int. J. Perform. Eng. **8**(6), 625–634 (2012)

Mönch, L., Fowler, J.W., Dauzere-Peres, S., Mason, S.J., Rose, O.: A survey of problems, solution techniques, and future challenges in scheduling semiconductor manufacturing operations. J. Sched. **14**(6), 583–599 (2011)

Plunkett, J.J., Dale, B.G.: Quality costs: a critique of some 'economic cost of quality' models. Int. J. Prod. Res. **26**(11), 1713–1726 (1988)

Reiner, G.: Customer-oriented improvement and evaluation of supply chain processes supported by simulation models. Int. J. Prod. Econ. **96**(3), 381–395 (2005)

Rodrigues, A., Bowers, J.: The role of system dynamics in project management. Int. J. Project Manag. **14**(4), 213–220 (1996)

Rust, R.T., Zahorik, A.J., Keiningham, T.L.: Return on quality (ROQ): making service quality financially accountable. J. Mark. **59**, 58–70 (1995)

Schneiderman, A.M.: Optimum quality costs and zero defects: are they contradictory concepts. Qual. Prog. **19**(11), 28–31 (1986)

Sweetser, A.: A comparison of system dynamics (SD) and discrete event simulation (DES). In: 17th International Conference of the System Dynamics Society, pp. 20–23, July 1999

Tako, A.A., Robinson, S.: The application of discrete event simulation and system dynamics in the logistics and supply chain context. Decis. Support Syst. **52**(4), 802–815 (2012)

Yasin, M.M., Czuchry, A.J., Dorsch, J.J., Small, M.: In search of an optimal cost of quality: an integrated framework of operational efficiency and strategic effectiveness. J. Eng. Tech. Manag. **16**(2), 171–189 (1999)

Wu, K.: An examination of variability and its basic properties for a factory. IEEE Trans. Semiconduct. Manufact. **18**, 214–221 (2005)

Automation of Cross-Factory Decision-Making Within Administrative Processes to Enhance Data Quality for Production

Jacob Lohmer[1(✉)], Christian Flechsig[1], Rainer Lasch[1], Germar Schneider[2], Dietrich Eberts[2], and Benjamin Zettler[2]

[1] Chair of Business Management, esp. Logistics, Technische Universität Dresden, Dresden, Germany
jacob.lohmer@tu-dresden.de
[2] Infineon Technologies Dresden GmbH & Co. KG, Dresden, Germany

Abstract. The efforts for digitalization and automation of manufacturing processes led to a neglected increase in manual administrative work tasks. Globalization, diversified supply chains, and the trend towards single-digit lot sizes add further complexity to administrative tasks and render a manual execution ineffective. Automation is a vital step for industries to stay competitive, but the path to it is not trivial. An in-depth analysis, optimization, and reengineering of old process flows is necessary to achieve the full optimization potential of automation activities. In this research, we indicate our approach that uses BPMN to standardize, analyze, and optimize processes from the Line Control department at the semiconductor fab of Infineon Technologies Dresden. Several improvements were proposed and are now at the implementation stage. However, some results are already visible, which we will discuss in the last section. The paper is concluded with limitations and an outlook into further research.

Keywords: Digitalization · Process automation · Administrative processes · BPMN · Lot-to-Order matching

1 Motivation

The efforts for digitalization in production environments have been considerably increased in the last years. Many manufacturing processes on the shop floor have posed a vital base for optimization and digitalization that enabled semiconductor manufacturers to produce complex products at low costs efficiently. These efforts have led to an increase in administrative back-office activities, which are often repetitive, time-consuming, and non-value adding (Strykowski and Wojciechowski 2012). Administrative processes have not been examined with the same amount of care and accuracy as production processes. This discrepancy is the result of a more sophisticated measuring process since the output of business processes is more complicated than manufacturing processes and cannot easily be measured by quality and quantity. Scientific methods like Lean Administration are utilized to automate business processes to reduce administrative efforts (Brenner 2018). However, the frequently used scientific methods

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 64–71, 2020.
https://doi.org/10.1007/978-3-030-48602-0_7

have not been customized to cope with the emerging trend of IoT and Big Data (Janiesch et al. 2017). In this context, an approach to identify and automate repetitive, administrative activities in order to shift the working efforts of the employees towards more innovative tasks would be expedient. Smart collaboration & knowledge management is a further area of interest.

The activities leading to this contribution are divided into two parts that are aligned and favor each other. On the one hand, we deal with the automation of administrative, manual, and repetitive activities, which we have identified and analyzed in the first place. On the other hand, we are developing a new scientific approach that enables the precise identification, analysis, and optimization of these activities, which can then be automated in a targeted manner. Globalization and diversified supply chains, as well as the trend towards single-digit lot sizes, have increased uncertainty in manufacturing environments and triggered volatile order quantities, especially for semiconductor fabs. Studies predict a further increase in order quantities with small lot sizes, which adds further complexity to the planning processes in semiconductor manufacturing (Mönch et al. 2017). As the usual administrative activities are still performed manually to a certain extent, a change to smaller but accumulating incoming orders is taking full effect at the administrative level. This observation was confirmed in our case study and added further impetus to the iDev40 project and our contribution. In this research, we indicate our approach to handling the increase in administrative activities that uses the Business Process Management and Notation (BPMN) to standardize, analyze and optimize processes from the Line Control department at Infineon Technologies Dresden. Robotic Process Automation (RPA) as a comparatively new method is also part of our approach as RPA can be utilized to automate simple, rule-based tasks without changing the underlying software architecture (Van der Aalst et al. 2018). Several improvements have been proposed and are now at the implementation stage. The results that are already visible will be presented and discussed in this contribution, which is concluded with limitations and an outlook into further research.

2 Description

Process management and optimization in business environments are mostly referred to as Business Process Management (BPM). BPM is a systematic approach to capture, design, execute, document, measure, monitor, and control non-automated as well as automated processes. The goal of this approach is always to sustainably achieve the goals that are aligned with the underlying enterprise strategy. BPM aims at identifying and eliminating non-value adding activities as well as the improvement of the process flow. This reduces costs and increases quality (Haddad et al. 2016; Houy et al. 2011). Different life cycle models of BPM have been developed in the literature; we focus on the one by Dumas et al. (2013) which is shown in Fig. 1.

At first, relevant processes are identified and precisely defined. The actual as-is process is then documented and modeled, taking into account the perception and expert knowledge of the concerned employees. The next step, process analysis, then identifies and quantifies issues associated with the as-is process to reveal optimization potential. The processes are redesigned into to-be models and subsequently implemented

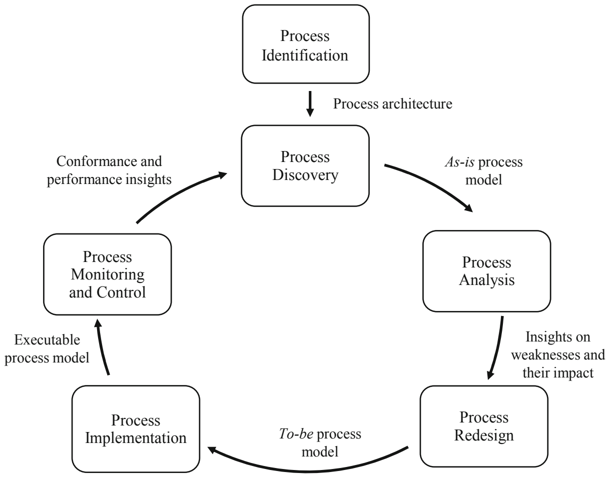

Fig. 1. Lifecycle model of BPM (based on Dumas et al. 2013)

(organizational change management and process automation). Continuous monitoring, analysis, and restarts, if appropriate, complete the BPM life cycle (De Weerdt et al. 2013; Dumas et al. 2013).

Several BPM standards have emerged, the interested reader is referred to Alotaibi and Liu (2017) and Ko et al. (2009) for detailed reviews. We used the renowned Business Process Model and Notation (BPMN), which offers a precise graphical presentation and also provides process execution methods. As one of the semiconductor industry partners in the iDev40 project, Infineon Technologies Dresden has been investigated for areas with potential for improvement regarding administrative, manual and repetitive processes. In particular, one of the central departments of production planning and control internally deals with the automation of existing administrative processes. This department was thoroughly examined, and many processes that could be automated were identified. The first process that is addressed is the allocation procedure of semi-finished products held in the Die Bank area of the semiconductor fab with specific and newly incoming customer orders, see Fig. 2 (based on Guhlich et al. 2015).

The primary step of wafer fabrication in the semiconductor production process is usually executed based on forecasts of customer orders rather than actual customer orders to be able to fulfill demands quickly. After the wafer test and sorting, the fabricated wafer lots are stored in the Die Bank inventory, which functions as the customer order decoupling point in most industry settings, this also applies to our case study. At this point, available lots are allocated to fixed customer orders for the upcoming weeks, a process called lot to order matching. The production mode therefore changes from make to stock setting to make to order setting after the lots are allocated to the customer orders. After lot to order matching, the lots are sent to the backend facility for assembly and test operations and subsequently stored in a finished

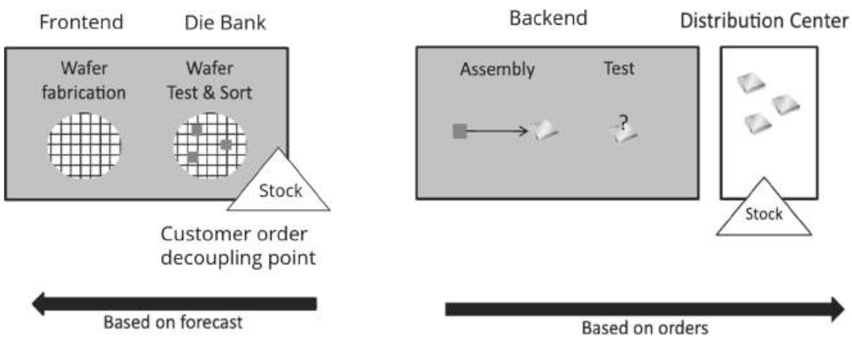

Fig. 2. Semiconductor production process (based on Guhlich et al. 2015)

goods inventory, usually at a distribution center. From the regional distribution centers, the products are distributed to the customer sites.

3 Results

The as-is process was analyzed, visualized, and transformed into a business process model using the Business Process Model and Notation (Camunda 2019; OMG 2013). In the as-is process in our case study, the internal process consists of the following steps: The customer demand for the upcoming week is aggregated by the respective product planners in weekly jour fixes and sent as internal orders in several worksheet formats to the colleagues from the line control department. The line control employees then have to apply data extraction and preparation techniques manually. There are several free text fields and different styles, depending on the responsible employee. The customer orders are imported into an MS Access pilot software tool, where the status of the orders is checked. Orders can be new, changed, unchanged, or canceled. The pilot software tool then returns the available lots for each customer order, which are manually allocated to matching customer orders based on a set of criteria that have shown good results in the past but are unable to cope with the upcoming complexity and uncertainty of a growing production site. The start scheduling of fixed customer orders is executed subsequently, where the overall capacity of the fab is estimated based on historical data from the last several weeks, and demand is divided by the number of days available. As the productive planners usually aggregate demands on Thursdays, the capacity of this day of the week is not fully utilized in most weeks. Therefore, this missing capacity is overcompensated on other weekdays (see Fig. 3).

Information was gathered in alignment meetings with several stakeholders regarding the as-is process, perceived shortcomings, and desirable capabilities of the system to be developed. We discovered that the process included many feedback loops and repetitive, unnecessary manual tasks. The BPMN as-is model (see Fig. 4 for an excerpt of the model) has been used in internal feedback rounds to optimize the process on hand and has been verified by internal operators. In the context of the digitalization of business processes, process analysis is indispensable for successful use, because it

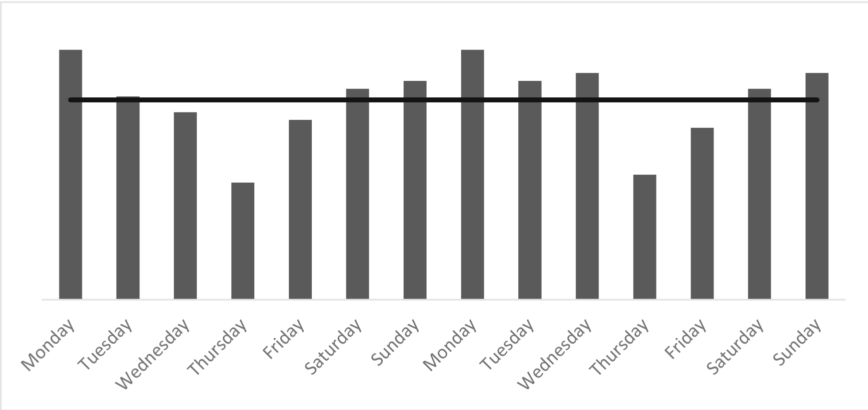

Fig. 3. Wafer starts from the Die bank sorted by weekdays (exemplary representation)

helps to uncover weaknesses and derives further improvement potential, thus paving the way for process automation. Additionally, process optimization is vital for any process that is targeted for automation, as a lack of optimization will complicate digitalization and automation. Optimized processes lead to higher profitability, which then could be further increased by digitalization.

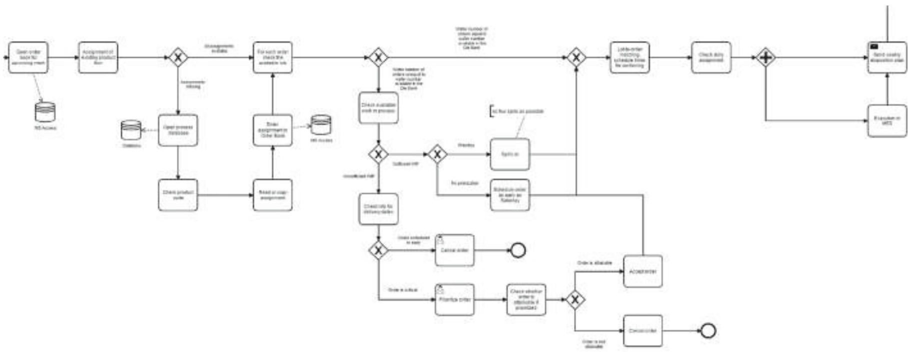

Fig. 4. Excerpt of the BPMN model of the as-is process using the Camunda Modeler (Camunda 2019)

Further interviews were arranged to consider the perspective of all process stakeholders. We have also defined appropriate objectives (e.g., capacity utilization, process efficiency) that can be verified before, during, and after automation. The target procedure for the whole Die Bank Sectioning process is currently in development and will take into account the needs of all process stakeholders and, perspectively, other partners in the supply chain. The roles and responsibilities of all partners within the core process have been defined and can be improved. In areas where full automation is not sufficient, we consider RPA as an alternative solution to perform activities ‘bottom-up’ using existing software tools. An approach to combine RPA and BPM would be helpful in this context, as our recent research has shown. This is also on our agenda for the coming months.

In order to optimize the allocation procedure, we are currently developing a rule-based scheduling system. The system works in real-time (utilizing the existing Real-Time Dispatcher repository in the fab) and assigns lots to customer orders based on a set of weighted criteria that is dynamically adjustable to the needs of the supply chain companies. After allocation, the lot starts are scheduled and equally distributed over the weekly time axis. In future research, we aim to include an operational capacity check in the optimization procedure that takes into account the actual uptimes and downtimes of resources.

To consider the scientific state-of-the-art in our system, a literature search was performed for existing literature concerning lot to order matching in semiconductor manufacturing, which returned a small set of articles (Boushell et al. 2008; Carlyle et al. 2001; Fowler et al. 2000; Kim et al. 2008; Knutson et al. 1999; Ng et al. 2010; Sun et al. 2011). The enforcement of a lot integrity policy seems to be the de-facto standard in industrial practice. Each production lot can only be allocated to a single customer order and lot splitting is not allowed (e.g., Ng et al. 2010). This differs from the industrial setting observed in our study, where the order of small wafer numbers is becoming increasingly popular. Therefore, we decided to incorporate the splitting procedure into the system and also to include the capacity constraints at the sorter resources in our considerations. The lot to order matching algorithm will ideally include a multi-objective optimization and will be designed to be flexible enough to respond adequately to various future requirements.

4 Discussion and Implications

The increasing complexity and the small order sizes in modern times have led to a shift of work from production to administration. Several tasks are still executed manually and are error-prone and suboptimal. Automation is essential for the industry to remain competitive, but the road to it is not always clear. In order to achieve the full optimization potential of automation activities, an in-depth analysis, optimization, and reengineering of process flows are essential. In this research, we used BPMN to analyze and optimize the lot to order matching process in the Line Control department of our case partner. Several improvements were proposed and are now at the implementation stage. Although the literature has presented a small set of contributions on the topic of lot to order matching in semiconductor companies, the study of a large semiconductor company has shown that the transfer between science and business is unsatisfying in this context. The process is executed by utilizing the employee's implicit knowledge rather than any sophisticated algorithm.

This study aims at enabling an efficient and time-saving process handling that is adjustable to the needs of the companies involved. Problems might arise as several internal and external partners need to agree on a specific and complicated process definition. The scope of the model might be too broad in some subareas leading to time-consuming feedback loops that could limit the achievable outcome. Furthermore, many partners need to be aligned to avoid local optimization and accomplish an overall optimized process.

In future research, we are dealing with the implementation of the developed procedure, test phases, and the productive roll-out. The focus will be on handling development lots to streamline this process and work towards a setup-optimized delivery procedure to optimize a larger share of the process. We also aim at developing a general approach to identify and systematically automate repetitive processes securely. The acquired knowledge from this process automation can be used for this scientific approach.

Acknowledgments. The project iDev40 has received funding from the ECSEL Joint Undertaking under grant agreement No 783163. The JU receives support from the European Union's Horizon 2020 research and innovation programme. It is co-funded by the consortium members, grants from Austria, Germany, Belgium, Italy, Spain and Romania. It is coordinated by Infineon Technologies Austria AG.

References

Alotaibi, Y., Liu, F.: Survey of business process management: challenges and solutions. Enterp. Inf. Syst. **11**, 1119–1153 (2017). https://doi.org/10.1080/17517575.2016.1161238

Boushell, T.G., Fowler, J.W., Keha, A.B., Knutson, K.R., Montgomery, D.C., Boushell, T.G., Fowler, J.W., Keha, A.B., Knutson, K.R.: Evaluation of heuristics for a class-constrained lot-to-order matching problem in semiconductor manufacturing. Int. J. Prod. Res. **46**, 3143–3166 (2008). https://doi.org/10.1080/00207540601001650

Brenner, J.: Lean Administration. Carl Hanser Verlag, München (2018)

Camunda: Camunda Modeler (2019). https://camunda.com/products/modeler/. Accessed 18 Mar 2019

Carlyle, M., Knutson, K., Fowler, J.: Bin covering algorithms in the second stage of the lot to order matching problem. J. Oper. Res. Soc. **52**, 1232–1243 (2001). https://doi.org/10.1057/palgrave.jors.2601222

De Weerdt, J., Schupp, A., Vanderloock, A., Baesens, B.: Process mining for the multi-faceted analysis of business processes - a case study in a financial services organization. Comput. Ind. **64**, 57–67 (2013). https://doi.org/10.1016/j.compind.2012.09.010

Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A.: Fundamentals of Business Process Management. Springer Verlag, Heidelberg (2013). https://doi.org/10.1007/978-3-642-33143-5

Fowler, J., Knutson, K., Carlyle, M.: Comparison and evaluation of lot-to-order matching policies for a semiconductor assembly and test facility. Int. J. Prod. Res. **38**, 1841–1853 (2000). https://doi.org/10.1080/002075400188627

Guhlich, H., Fleischmann, M., Stolletz, R.: Revenue management approach to due date quoting and scheduling in an assemble-to-order production system. OR Spectr. **37**, 951–982 (2015). https://doi.org/10.1007/s00291-015-0401-3

Haddad, C., Ayala, D., Uriona Maldonado, M., Forcellini, F.A., Lezana, Á.G.R.: Process improvement for professionalizing non-profit organizations: BPM approach. Bus. Process Manag. J. **22**, 634–658 (2016). https://doi.org/10.1108/BPMJ-08-2015-0114

Houy, C., Fettke, P., Loos, P., Van der Aalst, W.M.P., Krogstie, J.: Business process management in the Large. Bus. Inf. Syst. Eng. **3**, 385–388 (2011). https://doi.org/10.1007/s12599-011-0181-5

Janiesch, C., Koschmider, A., Mecella, M., Weber, B., Burattin, A., Di Ciccio, C., Gal, A., Kannengiesser, U., Mannhardt, F., Mendling, J., Oberweis, A., Reichert, M., Rinderle-Ma, S., Song, W., Su, J., Torres, V., Weidlich, M., Weske, M., Zhang, L.: The internet-of-things meets business process management: mutual benefits and challenges (2017). https://arxiv.org/abs/1709.03628

Kim, Y., Bang, J., An, K., Lim, S.: A due-date-based algorithm for lot-order assignment in a semiconductor wafer fabrication facility. IEEE Trans. Semicond. Manuf. **21**, 209–216 (2008)

Knutson, K., Kempf, K., Fowler, J., Carlyle, M.: Lot-to-order matching for a semiconductor assembly and test facility. IIE Trans. (Inst. Ind. Eng. **31**, 1103–1111 (1999). https://doi.org/10.1080/07408179908969911

Ko, R.K.L., Lee, S.S.G., Lee, E.W., Wah Lee, E.: Business process management (BPM) standards: a survey. Bus. Process Manag. J. (2009). https://doi.org/10.1108/14637150910987937

Mönch, L., Uzsoy, R., Fowler, J.W.: A survey of semiconductor supply chain models part I: semiconductor supply chains, strategic network design, and supply chain simulation. Int. J. Prod. Res. **7543**, 1–20 (2017). https://doi.org/10.1080/00207543.2017.1401234

Ng, T.S., Sun, Y., Fowler, J.: Semiconductor lot allocation using robust optimization. Eur. J. Oper. Res. **205**, 557–570 (2010). https://doi.org/10.1016/j.ejor.2010.01.021

OMG: Object Management Group. Business Process Model and Notation (BPMN), version 2.0.2 (2013). http://www.omg.org/spec/BPMN/

Strykowski, S., Wojciechowski, R.: Composable modeling and execution of administrative procedures. In: Kö, A., Leitner, C., Leitold, H., Prosser, A. (eds.) Advancing Democracy, Government and Governance - Lecture Notes in Computer Science, pp. 52–66. Springer Verlag, Berlin/Heidelberg (2012). https://doi.org/10.1007/978-3-642-32701-8

Sun, Y., Fowler, J.W., Shunk, D.L.: Policies for allocating product lots to customer orders in semiconductor manufacturing supply chains. Prod. Plan. Control **22**, 69–80 (2011). https://doi.org/10.1080/09537287.2010.490020

Van der Aalst, W.M.P., Bichler, M., Heinzl, A.: Robotic process automation. Bus. Inf. Syst. Eng. **60**, 269–272 (2018). https://doi.org/10.1007/s12599-018-0542-4

Human Factors in Semiconductor Manufacturing

Digitizing Human Work Places in Manufacturing Through Augmented and Mixed Reality

Michael Spitzer(✉), Manfred Rosenberger, Alexander Stocker, Inge Gsellmann, Matthias Hebenstreit, and Michael Schmeja

Virtual Vehicle Research Center, Inffeldgasse 21/A, 8010 Graz, Austria
michael.spitzer@v2c2.at

Abstract. It is no longer obvious that the share of knowledge work in production environments is increasing. However, this is not the only reason why more attention needs to be paid to the human factor in production. It is therefore increasingly important to provide workers in a production environment with the best possible support for their knowledge tasks by using modern information and communication technologies. In particular, the technical innovations in terms of augmented and mixed reality offer a great potential to be implemented in relevant applications in connection with the ongoing digitization or digital transformation of production. Against this background, this paper shows how current knowledge-based work processes in production environments can be best supported by augmented and mixed reality technologies. For this purpose, the paper reviews the state of the art of augmented and mixed reality technologies and then outlines in two concrete industrial use cases from the manufacturing domain how human work can be digitally augmented to facilitate knowledge-intensive production tasks.

Keywords: Augmented reality · Mixed reality · Knowledge-based work processes

1 Introduction and Motivation

Increasing demands for more innovative products and rising competition lead manufacturing companies to design more flexible and efficient production environments to sustain their competitiveness. As a result of a transformation process, the increasing automatization of production work has reduced the amount of manual work depending on the application domain, while the remaining manual work has become increasingly knowledge intensive (Campatelli et al. 2016). As knowledge has obviously been accepted as an important organizational asset, knowledge management systems– as a class of information systems – have since then greatly promoted the creation, transfer, and application of knowledge in organizational environments (Alavi and Leidner 2001). Though facilitating individual and organizational knowledge work by implementing technologies is per se not a new phenomenon (Stocker et al. 2012), recent developments of digital technologies including mobile technologies, big data analytics,

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 75–87, 2020.
https://doi.org/10.1007/978-3-030-48602-0_8

augmented, mixed and virtual reality seem to offer even more promising opportunities to facilitate knowledge-intensive tasks on the shop floor (Hannola et al. 2018).

Terms including 'Industry 4.0' or 'Digitalization' have become very popular in recent years and have become very successful in drawing the attention of senior decision makers. However, implementing related technologies to support knowledge work is not a management fashion, but can sustainably empower people and daily operations (Leyer et al. 2019). It is a grand challenge of any successful implementation project to improve current and future work practices of employees (Richter et al. 2018), which involves capturing and fully understanding the as-is situation, co-designing a to-be situation with the relevant stakeholders, and then kicking-off an iterative solution design supported by several digital prototypes with a rising degree of maturity. All this requires an integrated, interdisciplinary, participative, and agile approach, which allows identifying, analyzing, and supporting human work practices in a predominantly digital environment. It is hence crucial that digital work designers understand how and why things work before they can provide a digital solution to support work practices (Richter et al. 2018).

Augmented reality (AR) and mixed reality (MR) are two very promising digital technologies capable of facilitating workers in production environments and easing their work practices. Both technologies have a long history, since first concepts of heads-up, see-through, head-mounted displays in manufacturing reach back into the 1990s (Caudell and Mizell 1992). However, recent technological advancements of wearable technology devices including Microsoft HoloLens or Google Glass (Rauschnabel and Ro 2016) have again caught the attention of researchers to evaluate the adoption of AR and MR in industrial worker-centric use cases. These devices may not just offer increased usability, but also increased usefulness to be adopted in industrial use cases.

Within this paper the following research question will be answered: How can current work practices be transformed trough implementing AR and MR technologies finally enabling future digital workplaces? To answer this research question, the authors first performed a literature review on AR and MR in general, and on AR and MR in production environments in particular. Second, – drawing from own experiences – the authors describe use cases of AR/MR technologies in two different production-related scenarios, discuss their results and summarize their lessons learned.

Next the authors review the scientific state of the art on AR and MR in production environments.

2 A Brief Review of the Literature

2.1 Augmented, Mixed and Virtual Reality

Although Augmented, Mixed and Virtual Reality like to be thrown into the same pot, these are three completely different approaches. Thereby the differentiation between the real and the virtual environment as two poles lying at opposite ends of a Reality-Virtuality Continuum plays a major role for better understanding the different concepts, as outlined by Milgram and Colquoun (1999) and shown in the figure below.

According to Milgram and Kishino (1994), MR is a subset of (VR) related technologies that involve the merging of real and virtual worlds somewhere along the "virtuality continuum" which connects completely real environments to completely virtual ones (Fig. 1).

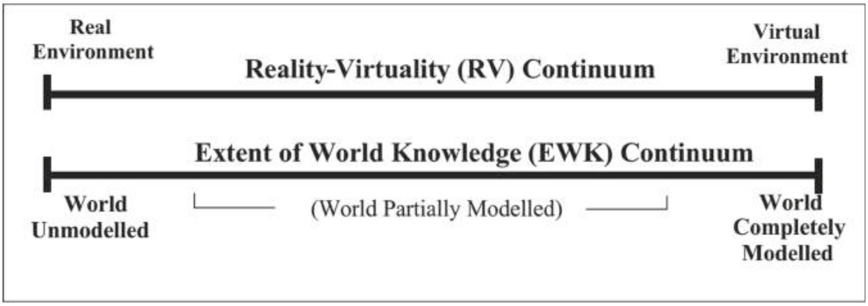

Fig. 1. Reality-Virtuality Continuum (Milgram and Colquoun 1999)

An early survey and one of the most cited papers on AR and its specific characteristics is provided by Azuma (1997) who defines AR as a variation of VR: While the user is fully immersed in VR and cannot see the real world around him, AR allows the user to see the real world with virtual objects superimposed upon or composited with the real world. Medical, manufacturing and repair, annotation and visualization, robot path planning, entertainment, and military air crafts are presented as applications in Azuma's paper. Azuma et al. (2001) published an updated survey paper considering the rapid technical advancements in the field of AR. In their review update they define an AR system as one that combines real and virtual objects in a real environment, runs interactively, and in real time; and registers (aligns) real and virtual objects with each other.

Another comprehensive survey of AR technologies and applications is provided by van Krevelen and Poelman (2010). According to them personal information systems, industrial and military applications, medical applications, entertainment, collaboration, and education and training are prominent applications domains for AR. A further survey on research and development in the field of AR is provided by Billinghurst et al. (2015)[1]. Education, architecture, and marketing are listed by them as prominent examples of typical modern-day applications. A more recent survey of mobile and wireless technologies for AR systems used for augmented ubiquitous computing is provided by Papagiannakis et al. (2008), covering application areas for mobile AR like virtual characters, cultural heritage, navigation and pathfinding, edutainment and games, collaborative assembly and construction, and maintenance and inspection.

VR, AR, and MR offer various potentials for innovative applications also besides digitizing manufacturing workplaces. However, manufacturing and related processes including e.g. machine maintenance or factory learning seem to be prominent industrial

[1] https://is.muni.cz/el/1433/podzim2015/PA198/um/59482554/A_Survey_of_Augmented_Reality.pdf.

applications areas for these technologies. Nevertheless, there is still a lot of misunderstanding of the differences between these concepts, which must be clarified. A widely used differentiation from a practitioner's viewpoint was provided by the software developer Julia Tokareva[2] and is shown in the figure below. According to her, VR immerses the user in a fully digital environment, while AR just overlays the real world with digital objects. MR additionally anchors digital objects to the real world (Fig. 2).

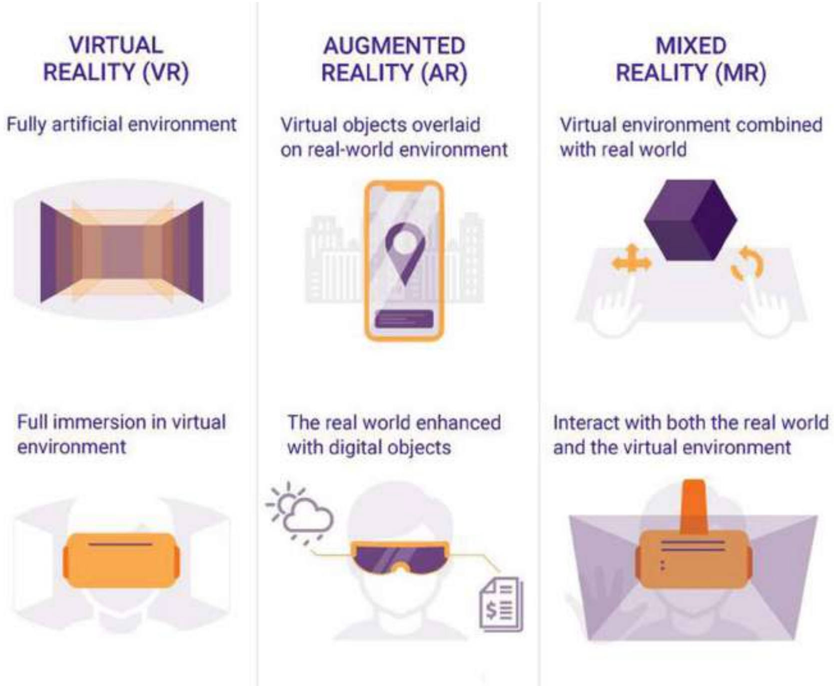

Fig. 2. VR, AR, and MR (Tokareva)

2.2 A Review of AR/MR in Industrial Use Cases

Following the scientific state-of-the-art on implementing augmented and mixed reality technologies in production(-related) environments, both hold a huge potential to support knowledge-based work. A series of researchers published scientific papers on how to apply AR/MR within industrial manufacturing and maintenance use cases.

Caudell and Mizell (1992) describe the design of a heads-up, see-through head-mounted display for human-involved aircraft manufacturing, as modern aircraft require

[2] https://www.forbes.com/sites/quora/2018/02/02/the-difference-between-virtual-reality-augmented-reality-and-mixed-reality/#423dea852d07.

a huge amount of manual efforts due to small lot sizes of parts and human skills required in many assembly tasks. AR can be for instance used to dynamically mark positions of drill holes inside an aircraft fuselage or to project graphical templates for location and orientation of composite cloth during the layout process. Neumann and Majoros (1998) describe an implementation for a maintenance scenario of a transport aircraft using AR technology, showing features that make AR attractive for manufacturing and maintenance. Friedrich (2002) outlines the potential for process and quality improvements for development scenarios (e.g. comparison of test and calculation of a crash rest, ergonomic layout design and flow visualization of pilot and passenger seats, and design and layout of cars), production/assembly scenarios (e.g. assembly of fresh-water system in aircraft, manual assembly in small batch production) and service scenarios (e.g. troubleshooting and service on production systems). Doil et al. (2003) describe how AR can be used to improve industrial planning processes, whereby an existing production environment can be augmented with virtual planning objects.

In their comprehensive paper Nee et al. (2012) review AR applications in design and manufacturing including AR collaborative design, robot path planning, plant layout, maintenance, CNC simulation, and assembly using AR tools and techniques. AR can provide manufacturing workers with hands-free access to context sensitive digital checklists to support assembly and quality control tasks in automotive manufacturing, thereby reducing process time and paper consumption (Stocker et al. 2017). Evans et al. (2017) present a prototype of a system for the Microsoft HoloLens to deliver spatially located AR assembly instructions.

Guhl et al. (2017) propose a concept for human-robot interaction using VR and AR on mobile devices including mobile phones and tablets as well as MR devices including the HoloLens, supporting human operators in interacting with and programming robots. Blaga and Levante (2018) present a developed scenario on human-robot collaborations using the HoloLens. Karlsson et al. (2017) introduce a decision support system using simulation and AR (HoloLens) to show a simulation model in 3D for an improved displaying of manufacturing information. Furthermore, AR and MR can support distant learning within a production environment, whenever a learner receives learning content as a video stream or as textual content from an instructor directly at the machine (Spitzer et al. 2018).

Following this review of the academic state-of-the-art, the authors provide two example cases, where augmented and mixed reality technologies will be used to support production-related processes.

3 An Implementation of Augmented and Mixed Reality in Two Production-Related Scenarios

After the literature review, the paper outlines two different scenarios of how digitally augmenting human work in production-related scenarios can improve human knowledge processes.

The first use case is in the additive manufacturing domain and will demonstrate how augmented and mixed reality can be used to support a worker in maintaining a

Fig. 3. Microsoft HoloLens

production machine (3D printer). Since a maintenance task often encloses a disassembly and assembly of a machine, this use-case partly overlaps with the second use-case. The second use case is in the automotive domain and will show, how to integrate mixed reality into a real-world construction and production workflow in the special machine design domain. As the second use case is still in its conceptualization phase, no demonstrator can be presented. The Microsoft HoloLens as shown in Fig. 3 is used for both use-cases as a representative of state-of-the-art Smart Glasses.

The Microsoft HoloLens is a fully untethered holographic eyewear. Virtual content is displayed on see-through holographic lenses directly in the field of view[3].

3.1 Maintaining a Production Machine with AR

Maintaining a production machine is a big challenge especially for inexperienced workers. Usually, there is a machine manual, instruction photos or videos available but it is still very difficult to gather all necessary information to be able to repair the machine. Photos or videos have the big disadvantage that the viewing angle is fixed by the creator of the material, sometimes it is necessary to see the maintenance instruction media from another viewing angle. Additionally, the creator of the instruction material often is a very experienced employee, therefore it is very difficult for him to put himself in the shoes of an inexperienced employee to provide suitable learning material. Written instructions occasionally cause ambiguity when shared across different countries, cultures and languages, especially when the maintenance process is very challenging, for example a maintenance process to clean the lens of an industrial laser cutting machine (Spitzer et al. 2017). Therefore, we implemented 3D maintenance animations augmented over the real machine. The advantage of this approach is that the viewing angle is not fixed and can be adjusted by walking around the machine while wearing the Microsoft HoloLens.

[3] https://docs.microsoft.com/en-us/windows/mixed-reality/hololens-hardware-details.

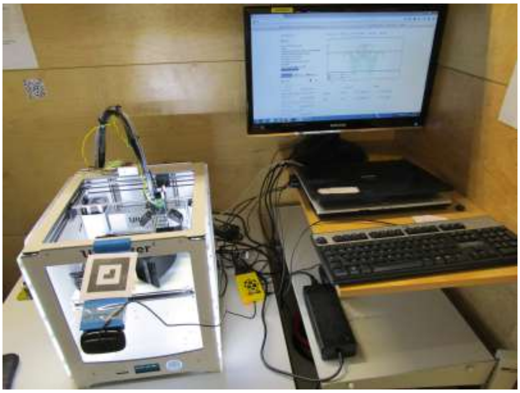

Fig. 4. 3D printer lab at Virtual Vehicle Research Center

We have a 3D printer lab in our research center which is used to print prototype car parts. Several departments of our research center are using the 3D printer. Figure 4 shows the Virtual Vehicle 3D printer lab.

The 3D printer is under heavy usage at our research center. Therefore, it is necessary to perform maintenance procedures. The glass plate has to be readjusted and cleaned, printer material has to be changed, the 3D printer nozzle has to be cleaned and many other maintenance tasks. We have some experts in the department, but when they are not available, inexperienced users have to perform the maintenance procedures. This situation is comparable to a real-world industry scenario. The expert can be ill, on vacation or on a business trip, in which case an inexperienced colleague must step in to perform the maintenance procedures. Figure 5 shows a capture of the field of view of a person who is wearing the HoloLens while performing the glass plate cleaning. The HoloLens UI can be placed in the room individually. The UI shows short text descriptions of the maintenance steps. The animation is augmented directly onto the 3D printer.

Fig. 5. 3D printer without HoloLens (left), 3D printer through HoloLens (right)

3.2 Automated Assembly Manual with AR

The second use-case is in the special engineering domain. The challenge in special engineering domain is that the production of such machines/products has very small lot sizes. Additionally, some machines only differ very slightly, hence, it is a challenge for the worker to assemble the right parts for a certain product. It is also a big challenge to communicate slight changes in the engineering CAD to the production worker. We will implement a software artifact to highlight such changes and to visualize the changes for the worker to minimize the error rate during assembly. Therefore, we develop an automated assembly manual with AR which automatically adapts to changes of parts and assembly groups.

There are a lot of people from engineering and production involved to create a new machine or product in general. The next section describes the different roles for this process. In some cases, one person could occupy multiple roles, depending on the company structure or on the complexity of the product. Because of non-disclosure of product development data (CADs, engineering documents) of our industry partners, we decided to abstract the use case to a Lego® Technic assembly which could be then mapped to the real industry use-case with adequate effort. This approach was already used in other projects (Spitzer and Ebner 2017).

Involved Roles in Special Engineering. This section describes the roles involved in technical aspects of product development. Other roles like business, management and logistics roles are not considered because we are focusing on optimizing the CAD/engineering to production workflow.

CAD Engineers. The CAD engineers construct the machine in CAD. They are aware of physical boundary conditions which must be considered. Additionally, they constantly do a reality check if the product can be assembled. They select appropriate parts, ideally standardized parts, and appropriate tools. The main goal is to reuse standardized parts and tools already available in enterprise resource planning (ERP) systems. Furthermore, they select and reserve space in the assembly for electronics. Also, technical documentation such as drawings, bill of material, calculations and reports are part of their work. The main challenges of CAD engineers are to use appropriate parts and design a product which can be assembled.

Assembly Manual Editors. The assembly manual editors translate the technical documentation of the CAD engineers to assembly instructions targeted at shop floor workers. They define the assembly sequence by mapping the engineering bill of materials (EBOM) to the manufacturing bill of materials (MBOM). Additionally, they select the appropriate manual type as for example printed CADs with some annotations, image or video-based instructions or 3D animations. The main challenges are to create an appropriate manual for a product which will be assembled for the first time. This usually is a very time-consuming process e.g. creation of 3D animated manuals. Very often, the used software does not fit for all kinds of manual types.

Assembly Planning Engineer. The assembly planning engineers ensure that the product can be assembled by the shop floor worker. They are responsible for error prevention, plausibility checks, collision detection and test. They validate the CAD engineers' tasks. Additionally, they estimate costs and time efforts of the assembly process. Also, the decision of whether the product will be assembled manually, partly automated or fully automated is on their head. They assemble the prototype first to identify assembly issues and trigger changes of the assembly manual if it is necessary. After assembly, they check if all flexible parts such as cables or tubes fit appropriately. The last step is then to identify the necessary skill level of the shop floor workers to be able to assemble the product. The main challenge of the assembly planning engineers is to identify all issues with the assembly manual and product assembly issues and give feedback to the assembly manual editors and CAD engineers.

Shop Floor Worker. The shop floor workers prepare the workplace, tools and parts. Their main task is to assemble the product. The main challenge is to identify revision changes and manual updates when the special engineering product is changed slightly.

As-Is Situation. The CAD engineer designs the product in CAD, the assembly manual editor creates the assembly manual using several tools as Microsoft Word, Visio, PowerPoint, several image editing applications and CAD tools. The assembly planning engineer validates the assembly manual and the engineering data and gives feedback to the CAD engineer and the assembly manual editor. If changes are necessary, the CAD and the assembly manual have to be adapted. The creation and adaption of the assembly manual is very time-consuming. The assembly planning engineer then communicates the work order to the shop floor workers. Figure 6 shows the process in detail.

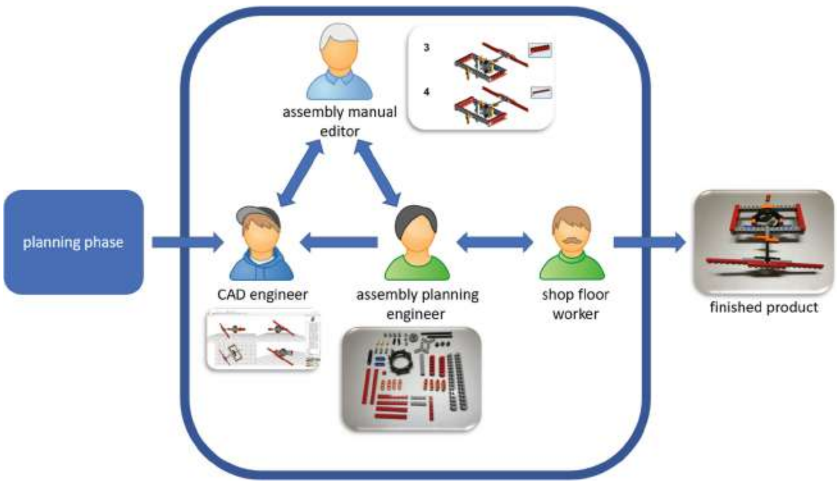

Fig. 6. As-is situation

To-Be Situation. The assembly manual should adapt to changes in the engineering data automatically. This will save a lot of time especially in the special engineering domain. The assembly planning engineer only has to communicate with the CAD

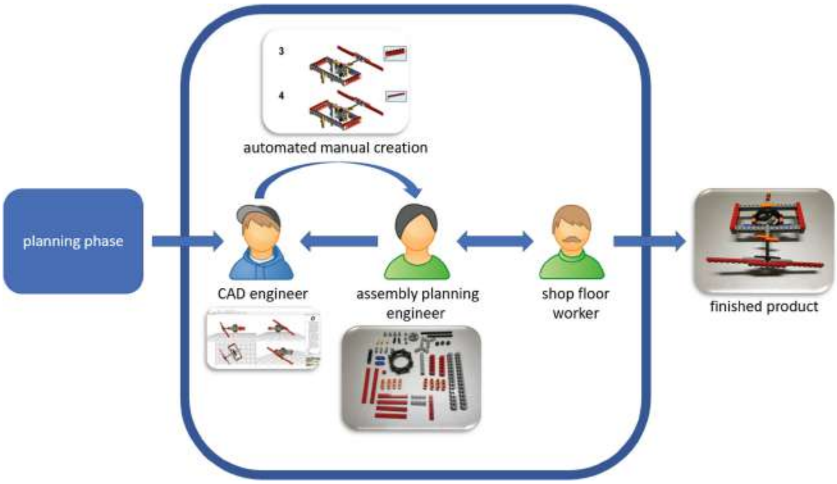

Fig. 7. To-be situation

engineers and the update of the assembly manual is triggered automatically when the CAD engineers change engineering data. Figure 7 shows the process in detail.

Main Challenges. The main challenge is to communicate small changes in engineering data to the shop floor worker. A lot of manual tasks must be performed to ensure data consistency. Such tasks should be automated as far as possible.

4 Conclusion, Discussion and Outlook

Following a review of the state of the art on AR in industrial use cases, the authors have provided insights into two different use cases, how human work can be digitally augmented to facilitate knowledge-intensive production tasks, maintaining a production machine with AR and automated assembly manual with AR. Thereby involved roles, the as-is situation (and challenge) as well as the envisaged to-be situation have been outlined, which is a useful practice when implementing smart factory technologies (Rosenberger and Stocker 2017).

Adopting AR/MR in production environments requires a change from the current as-is situation to the envisaged to be situation. In such a challenging change process, current work practices must be turned into digitally empowered work practices, while developed digital artefacts must support and transform current work places to enable digital work places in the best possible way (Richter et al. 2018). Such projects have a social perspective as information systems are treated as socio-technical systems and these initiatives must deliver (expected) benefits (Luna-Reyes et al. 2005). Users must adapt to new work settings and adopt new working practices. Active change management is a crucial factor for project success.

Implementing augmented and mixed reality in industrial use cases can have both theoretical and practical implications. For instance, digitally augmenting human work can enable context-aware access to process-relevant information and knowledge at the shop floor and allows cutting service and/or production times, while at the same time increasing product and process quality as empowered employees can make better-informed decisions. Workers making informed decisions may reduce their level of frustration, increase their job satisfaction (Schafler et al. 2018), and support retaining a productive flow of work. Context-relevant information displayed in the line of sight without media breaks, and seamless interaction across different IT tools becomes crucial for smooth operation and avoidance of cognitive overload. This will obviously generate not only a technical but also a social impact in factories.

Acknowledgments. This project has received funding from the Electronic Component Systems for European Leadership Joint Undertaking under grant agreement No 783163. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and Germany, Austria, France, Czech Republic, Netherlands, Belgium, Spain, Greece, Sweden, Italy, Ireland, Poland, Hungary, Portugal, Denmark, Finland, Luxembourg, Norway, Turkey.

This study has received funding from FACTS4WORKERS –Worker-centric Workplaces for Smart Factories-the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 636778.

In Austria the project was also funded by the program "IKT der Zukunft" and the Austrian Federal Ministry for Transport, Innovation and Technology (bmvit).

The publication was written at VIRTUAL VEHICLE Research Center in Graz and partially funded by the COMET K2 – Competence Centers for Excellent Technologies Programme of the Federal Ministry for Transport, Innovation and Technology (bmvit), the Federal Ministry for Digital, Business and Enterprise (bmdw), the Austrian Research Promotion Agency (FFG), the Province of Styria and the Styrian Business Promotion Agency (SFG).

References

Alavi, M., Leidner, D.: Knowledge management and knowledge management systems: conceptual foundations and research issues. MIS Q. **25**(1), 107–136 (2001)

Azuma, R.: A Survey of Augmented Reality, Presence: Teleoperators and Virtual Environments 6, 4, 355–385 (1997)

Azuma, R, Baillot, Y, Behringer, R, Feiner, S, Julier, S, MacIntyre B.: Recent advances in augmented reality. IEEE Comput. Graph. Appl. (2001). https://apps.dtic.mil/dtic/tr/fulltext/u2/a606245.pdf

Billinghurst, M., Clark, A., Lee, G.: A survey of augmented reality. Found. Trends® Hum.–Comput. Interact. **8**(2–3), 73–272 (2015). https://doi.org/10.1561/1100000049

Blaga, A., Levante, T.: Augmented reality for digital manufacturing. In: Proceedings of 26th Mediterranean Conference on Control and Automation (MED) (2018)

Campatelli, G., Richter, A., Stocker, A.: Participative knowledge management to empower manufacturing workers. Int. J. Knowl. Manag. **12**(4), 37–50 (2016)

Caudell, T.P., Mizell, D.W.: Augmented reality: an application of heads-up display technology to manual manufacturing processes. In: Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences, 7–10 Jan 1992

Doil, F., Schreiber, W., Alt, T., Patron, C.: Augmented reality for manufacturing planning. In: EGVE 2003 Proceedings of the workshop on Virtual environments (2003)

Evans, G, Miller, J., Pena, M.I., MacAllister, A., Winer, E.H.: Evaluating the Microsoft HoloLens through an augmented reality assembly application. Mechanical Engineering Conference Presentations, Papers, and Proceedings. 179 (2017)

Friedrich, W.: ARVIKA-augmented reality for development, production and service. In: Proceedings of the International Symposium on Mixed and Augmented Reality (2002). https://doi.org/10.1109/ismar.2002.1115059

Guhl, J., Tung, S., Kruger, J.: Concept and architecture for programming industrial robots using augmented reality with mobile devices like Microsoft HoloLens. In: Proceedings of 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA) (2017)

Hannola, L., Richter, A., Richter, S., Stocker, A.: Empowering production workers with digitally facilitated knowledge processes – a conceptual framework. Int. J. Prod. Res. **56**(14), 4729–4743 (2018)

Karlsson, I., Bernedixen, J., Ng, A.H.C., Pehrrsson L.: Combining augmented reality and simulation-based optimization for decision support in manufacturing. In: Proceedings of Winter Simulation Conference (WSC) (2017)

Leyer, M., Richter, A., Steinhüser, M.: "Power to the workers": empowering shop floor workers with worker-centric digital designs. Int. J. Oper. Prod. Manag **39**(1), 24–42 (2019). https://doi.org/10.1108/IJOPM-05-2017-0294

Milgram, P., Colquhoun, H.: A taxonomy of real and virtual world display integration. In: Ohta, Y., Tamura, H. (eds.) Mixed Reality, no. 418, pp. 5–30. Springer-Verlag Berlin, Heidelberg (1999). https://www.springer.com/de/book/9783642875144

Milgram, P., Kishino, F.: A Taxonomy of Mixed Reality Visual Displays. IEICE Trans. Inf. Syst. **77**(12), 1321–1329 (1994)

Luna-Reyes, L.F., Zhang, J., Gil-Garcia, R., Cresswell, M.: Information systems development as emergent socio-technical change: a practice approach. Eur. J. Inf. Syst. **14**, 93–105 (2005)

Nee, A.Y.C., Ong, S.K., Chryssolouris, G., Mourtzis, D.: Augmented reality applications in design and manufacturing. CIRP Ann. **61**(2), 657–679 (2012)

Neumann, U., Majoros, A.: Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance. In: Proceedings of the IEEE 1998 Virtual Reality Annual International Symposium (1998)

Papagiannakis, G., Singh, G., Thalmann, N.M.: A survey of mobile and wireless technologies for augmented reality systems. Comput. Animation Virual Worlds **19**, 3–22 (2008)

Rauschnabel, P.A., Ro, Y.K.: Augmented reality smart glasses: an investigation of technology acceptance drivers. Int. J. Technol. Mark. **11**(2), 123–148 (2016)

Richter, A., Heinrich, P., Stocker, A., Schwabe, G.: Digital work design - the interplay of human and computer in future work practices as an interdisciplinary (grand) challenge. Bus. Inf. Syst. Eng. **60**(3), 259–264 (2018)

Rosenberger, M., Stocker, A.: Eine vorgehensweise zur unterstützung der einführung von industrie-4.0-technologien. In: Burghardt, M., Wimmer, R., Wolff, C., Womser-Hacker, C. (Hrsg.). Mensch und Computer 2017 - Workshopband. Regensburg: Gesellschaft für Informatik e.V (2017)

Schafler, M., Lacueva-Perez, F.J., Hannola, L., Damalas, S., Nierhoff, J., Hermann, T.: Insights into the introduction of digital interventions at the shop floor. In: Proceedings of the 11th Pervasive Technologies Related to Assistive Environments Conference, Corfu, Greece, 26–29 June 2018

Spitzer, M., Ebner, M.: Project based learning: from the idea to a finished LEGO® technic artifact. assembled by using smart glasses. In: Proceedings of EdMedia: World Conference on Educational Media and Technology, pp. 196–209. Association for the Advancement of Computing in Education (AACE) (2017)

Spitzer, M., Nanić, I., Ebner, M.: Distance learning and assistance using smart glasses. Educ. Sci. **8**(1), 21 (2018)

Spitzer, M., Schafler, M., Milfelner, M.: Seamless learning in the production. In: Burghardt, M., Wimmer, R., Wolff, C., Womser-Hacker, C. (eds.) Mensch und Computer 2017 - Workshopband. Gesellschaft für Informatik e.V, Regensburg (2017)

Stocker, A., Richter, A., Hoefler, P., Tochtermann, K.: Exploring appropriation of enterprise wikis: a multiple-case study. Comput. Support. Coop. Work (CSCW) **21**(2–3), 317–356 (2012)

Stocker, A., Spitzer, M., Kaiser, C., Rosenberger, M., Fellmann, M.: Datenbrillengestützte checklisten in der fahrzeugmontage. Informatik-Spektrum **40**(3), 255–263 (2017)

Van Krevelen, D.W.F., Poelman, R.: A survey of augmented reality technologies, applications and limitations. Int. J. Virtual Real. **9**(2), 1–20 (2010)

Competency Requirements at Digitalized Workplaces in the Semiconductor Industry

Sophia Keil[1], Fabian Lindner[1](✉), Josef Moser[2], Rüdiger von der Weth[3], and Germar Schneider[4]

[1] Faculty of Business Administration and Engineering, Zittau/Goerlitz University of Applied Sciences, Zittau, Germany
{sophia.keil, fabian.lindner}@hszg.de
[2] Infineon Technologies Austria AG, Villach, Austria
[3] Faculty of Business Administration, Dresden University of Applied Sciences, Dresden, Germany
[4] Infineon Technologies Dresden GmbH & Co. KG, Dresden, Germany

Abstract. Automation and digitalization in manufacturing are used as means to speed up the production, reduce its costs, and improve the quality of the products. However, new systems and technologies that are applied to achieve these goals often require great adaptation efforts by the employees using them. This is especially the case in the semiconductor industry, where there is great competitive pressure, with the need for fast, flexible and innovative manufacturing processes. This paper examines at different workplaces at Infineon Austria and Dresden, how job characteristics and required competencies have changed so far due to digitalization, and how employees can be prepared to further changes in production. Therefore, a literature review, interviews, and observations are conducted and the results presented.

Keywords: Industry 4.0 · Competency requirements · Digitalized workplaces · Big data · Trainings · Education programs

1 Introduction

Semiconductor manufacturers face high competitive pressure regarding reduced production times, costs, and enhanced quality, as well as short innovation cycles, and the need for fast ramp-ups of new technologies (Kern 2016). The automation and digitalization of factories and manufacturing (e.g. by using handling robots), and business processes can offer opportunities to meet these challenges. Therefore, robots operating at the production line, and automated transport systems running through fabrication facilities are already widely used in the semiconductor sector. In this way, large amounts of digital information and data are generated during the whole manufacturing process, e.g. via sensors that are monitoring the machine processes and reporting on their status or specific parameters of the production to the responsible workers (Schöning and Dorchain 2014). Furthermore, information assistance systems such as head-mounted displays (HMD), that are or can be used in the production to aid the employees raise new questions about their use. Are these systems adding value to the

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 88–106, 2020.
https://doi.org/10.1007/978-3-030-48602-0_9

supply chain by increasing the efficiency and making the work tasks more comfortable for the workers – or do they result in more distraction and strain for the colleagues? Due to this ongoing automation and digitalization employees need to cope with knowledge and competency challenges at workplaces in Industry 4.0. This notion was introduced in 2011 by the German government's high-tech strategy and essentially stands for the intelligent linkage of machines and processes using ICT. It is also seen as the next evolutionary level of fabrication (Kagermann et al. 2013).

Within the European research project iDev40 (Integrated Development 4.0) one main focus is on the role of the human within complex and highly automated and digitalized working environments. Leading questions of this part of the project are about the design of human-centered future workplaces, the efficiency of human-machine interaction, knowledge sharing, smart collaboration, and training as well as educational programs to empower employees for Industry 4.0 scenarios.

In the following, Sect. 2 describes the research questions and the methodology used for answering them. Section 3 presents the results of the research and is divided into three subsections. The first one explains the characteristics of the three generic and production-related types of workplaces: operation, maintenance, and development workplace. The second subsection introduces relevant competencies in digitalized or Industry 4.0 working environments and compares them to the empirical findings at nowadays' workplaces in semiconductor fabrication facilities. Afterwards, measures for assessing and increasing the competency level of workers and engineers are briefly evaluated. In the end, the findings are discussed and an outlook on the topic is given.

2 Methodology

To contribute to the aforementioned aims of the respective iDev40 subproject, the following research questions (RQ) are proposed:

- RQ 1.1 – What characterizes automated and digitalized workplaces and work tasks in the semiconductor industry?
- RQ 1.2 – Which competency requirements must be met by employees in automated and digitalized working environments in the semiconductor industry?
- RQ 2 – How can employees in the semiconductor industry educationally and technically be empowered to meet these competency requirements?

I.e., the goal of this paper is to distinguish characteristic workplaces and to define the core competencies that workers in digitalized working environments in the semiconductor industry need. Additionally, an outlook will be given on possible (further) training opportunities. Besides studying the literature on competency requirements in Industry 4.0 and the high-tech sector in general, this paper focuses on empirical evidence. Therefore, semi-structured interviews were conducted with employees of different types of workplaces at Infineon Technologies in Villach, Austria (IFAT) on August 9th – 10th, 2018. The first part of the interview surveyed the workers' personal attitude towards and perspective on digitalization and automation at their workplace, and how it affects – or would affect – their daily working routine (Appendix A). In the second part, the employees were asked about their work with regard to the characteristic 5 V of Big Data commonly used in theory and practice: volume, variety, velocity, veracity, and value (Ilie-Zudor et al. 2015). As there is no common definition of when to speak of Big Data (Dautov and Distefano 2017), these five characteristics are mainly used to describe them. They represent challenges to the used technology, and result in “datasets that could not be perceived, acquired, managed, and processed by traditional IT and software/hardware tools within a tolerable time” (Chen et al. 2014). In this way, the interviews help to examine to which extent information and data play a role at each workplace.

In addition to the qualitative analysis of the interview transcriptions (Kuckartz 2016), the findings of several days of observation interviews (Kuhlmann 2009) of a maintenance workplace at the production site of Infineon Technologies in Dresden, Germany (IFD) between November 19th and 23rd, 2018 are considered. The application of these mixed methods approach towards answering the before proposed research questions is depicted in Fig. 1.

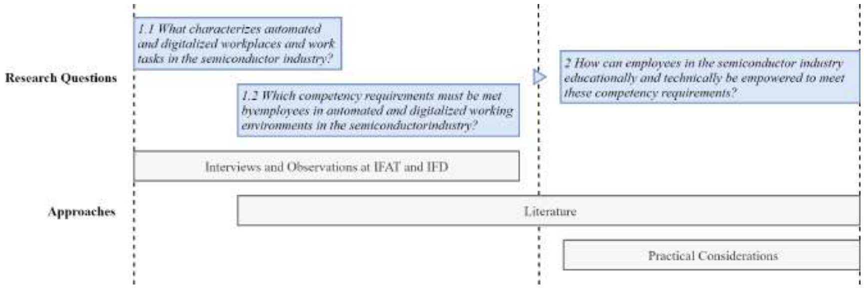

Fig. 1. Research design.

3 Results

3.1 Characteristics of Digitalized Workplaces in the Semiconductor Industry

In the beginning, three types of generic production-related workplaces at the semiconductor manufacturer Infineon Technologies are distinguished: production, maintenance, and development workplaces. The first one encompasses the workers on the shop-floor that are operating and monitoring several production machines (Fig. 2). Depending on the degree of automation on the shop floor, these employees are also more or less responsible for loading and unloading these machines, as well as fixing minor errors. Supervisors at production workplaces are furthermore accountable for procurement, prioritizing orders, organizing and leading the shift, as well as reporting to higher authorities.

Maintenance workplaces are highly characterized by monitoring, and supervision tasks, as well as repair activities – mechanically and electronically (Fig. 2). Due to the amount of IT and sensor systems that are controlling and operating the plants and the transportation systems, maintenance regarding software or sensor issues is increasing.

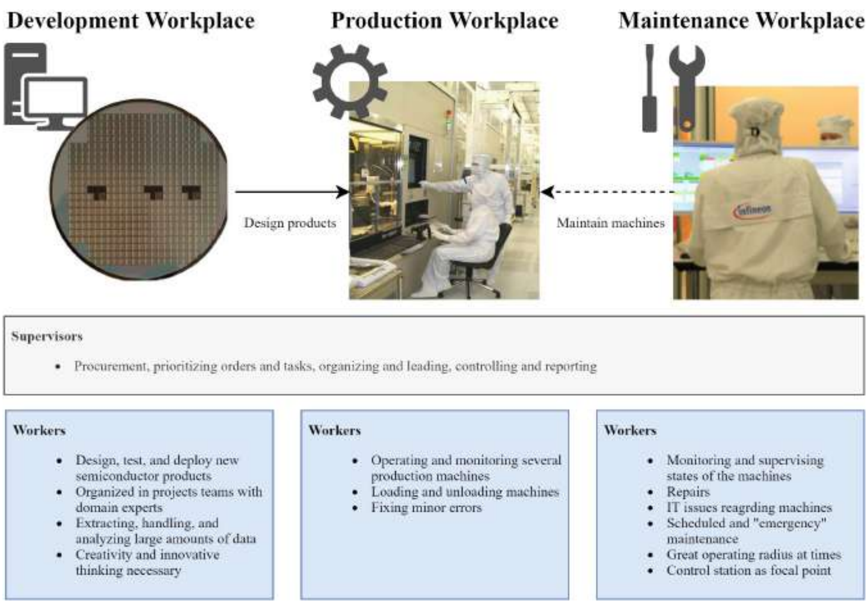

Fig. 2. Generic production and production-support workplaces at semiconductor manufacturers.

That's why part of the maintenance staff is specially trained for solving more complex IT issues as systems experts. Besides preventive maintenance activities on a regular basis before errors occur, the maintenance staff has to react quickly to unsuspected errors and fixing them to keep the production going. Different maintenance teams are responsible for different types of machines or systems. Due to this, the spatial area of action inside the fabrication facility differs as well. E.g., maintenance workers of the transportation systems are deployed all over the fabrication facility, whereas maintenance workers of a special type of machine might have a smaller operating radius – as generally similar types of machines are placed together in semiconductor fabrication facilities. When not on maintenance activities, the maintenance staff operates from within its own control room monitoring the respective systems. On supervision level, the work tasks and responsibilities are similar to those of the machine operators.

At development workplaces engineers design, test and deploy new semiconductor products in accordance with customer requests (Fig. 2). To do so, development teams are mainly organized along projects with experts for the respective tasks of the development process. At this workplace extracting, handling, and analyzing large amounts of data play an important role – electrical parameters and machine data can rise up to thousands of variables that can be examined theoretically to develop new product features or to evaluate them. Further characteristics of this workplace are the creativity and innovative thinking necessary to develop new products. However, routine tasks like data handling are still part of the job. Due to necessary expert knowledge and competencies for the development of new products these teams might be distributed over several production sites. Therefore, a successful collaboration cross-site necessitates certain communication and social skills of the employees. On the level of

	Development Status	Development Change	Production Status	Production Change	Maintenance Status	Maintenance Change
Physical Work	low	○	moderate	▽	moderate	▽
Concerns	moderate	△	moderate	△	low	○
Collaboration	high	△	low	▽	moderate	○
5Vs of Big Data						
Volume	high	△	low–moderate	○	moderate	△
Variety	high	△	moderate	○	high	△
Velocity	high	△	moderate	○	high	△
Veracity	moderate	△	moderate	○	moderate	△
Value	high	△	-	-	-	-

Change so far due to digitalization: ○ Stagnation △ Increase ▽ Decrease

Fig. 3. Status and change of several aspects due to digitalization at the respective types of workplaces in semiconductor industry based on own findings.

supervisors, the tasks and responsibilities again do not largely differ from those at the other workplaces – at least the amount of project management and monitoring activities is increased.

On the basis of the interviews and observations several aspects with regard to digitalization can be extracted that can be used for differentiation of these three types of workplaces (Fig. 3):

– The amount of physical work and activity during the job.
– The perception of digitalization of workplaces as a risk.

– The need for communication and collaboration.
– The 5Vs of Big Data:
  - The volume of information and data that play a role at the workplaces.
  - The variety of formats of information as well as the variety of sources of data.
  - The velocity of generating and processing information at the workplaces.
  - The veracity or quality of the information needed for the respective workflows.
  - The value generated through data.

Asking the employees about their opinion of digitalization affecting their workplaces already or in the future, one of the first aspects is the amount of physical work that has to be executed. At operating and maintenance workplaces physical activities still play a role. However, digitalization (and automation) affected the work of operators the most. At maintenance workplaces the amount of physical strain in general is still high due to mechanical reparations and the greater radius of action within the fabrication facility. Whereas digitalization did not affect the work of development engineers with respect to physical tasks – it still does not play a role. Machine operators seem to see the implementation of new technologies more skeptical as e.g. maintenance workers. I.e., they might be afraid of losing their jobs due to digitalization or to be moved to a new workplace that might also require some re-training. Furthermore, the colleagues fear overstrains due to too much digital noise like unnecessary information overflows. In this context, the age distribution of the workers matters as well – older employees seem to have greater concerns as younger ones that might already be very familiar with any kinds of digital technology as they grew up in a more or less digitalized world.

Further aspects of investigation are the 5Vs of big data (Ilie-Zudor et al. 2015) to compare the three types of workplaces by characteristics of the information flows needed for planning and decision making during work. With regard to the volume of information or data that must be processed at work, developers have to cope with the most – as described before – followed by the maintenance workplaces, where machine states from different systems come together at the control station or remotely. At both workplaces the amount of data increased and is expected to increase further due to digitalization as well. At operating workplaces volume of the data, like orders, is relatively low (to moderate). In any case, the supervisors of each workplace tend to have to manage more information as the workers, as they have additional organizing tasks regarding the shift respectively the team. Considering the variety of formats and sources of data, it is a similar picture as before. Both development and maintenance workplaces get their information from several different systems or sources and in very different formats from structured to unstructured. E.g., machine data like states and advanced process control data are structured, whereas phone calls for reporting an error to the maintenance staff are unstructured. Additionally, various IT systems might exist due to different manufacturers of the plants. So far, the ongoing automation and digitalization has led to increasing variety of sources of information due to new systems e.g., that were implemented.

However, the interviewees wish further uniform and structured information flows to ease their work. Operators as well face structured and unstructured information but in a smaller scale as the others. The velocity of information is especially high at development and maintenance workplaces, i.e. up to real-time due to advanced process control. However, this applies not to all information flows, so that it will probably increase further. In terms of veracity of the data, at all of the workplaces the information is available in more or less good quality that is sufficient for working with it. But unstructured information like phone calls in the maintenance control station are most of the time less accurate. Due to the increase of structured data used, the veracity of the information has also increased. Value creation through extracting, examining, analyzing and searching for useful patterns in raw data in the narrow sense of the fifth V of big data, is solely relevant at the development workplaces, and increasing with digitalization.

However, further similarities in all workplaces with regard to digitalization exist, when asked about the overall attitude towards digitalization at work: all of them see this development as a chance – despite some concerns mentioned before. Furthermore, all of the colleagues are aware of the constant change and the required flexibility to adapt to it – either as they already made experiences with it in the past since they are working in the semiconductor industry, or as they already grew up with ever faster changes in technology. For most of them there is also a great interest in technological changes in general, and they are open-minded towards testing new technologies. Nevertheless, there is also a consensus that tasks and activities do not necessarily become easier, more comfortable or less demanding due to digitalization. The employees expect them just to change and to develop further.

Additional interesting remarks were made on possible steps for implementing new technology. It was experienced as helpful if new technologies are implemented step by step, so that everyone can gradually adapt to the changes. Furthermore, the first pilot studies with a new technology should be conducted with persons that are very enthusiastic. In this way other employees can be influenced positively in their attitude. Still, coping with new technologies or new forms of communication probably rely on some degree on personal traits, as well. Besides all technological advancements, social contact and vis-à-vis communication seem still important for the well-being of the employees. If not possible otherwise, scheduled personal meetings should be installed therefore.

3.2 Required Competencies in Digitalized Working Environments

As the preceding chapter has shown, the types of workplaces in the semiconductor industry differ in several analogue and digital aspects. These specifics might require different competencies of the workers as well. In the following, four different approaches of defining digital competencies are therefore presented and compared (Fig. 4). These were selected due to varying methods used for finding and extracting digital or Industry 4.0-relevant competencies, as well as varying in the scope to which they are applied to, and in the competencies itself. In this way, different approaches and perspectives are considered. In addition, the results of these studies are juxtaposed with the findings and derivations of competencies out of the analysis of the interviews and observations at Infineon Technologies (IFX).

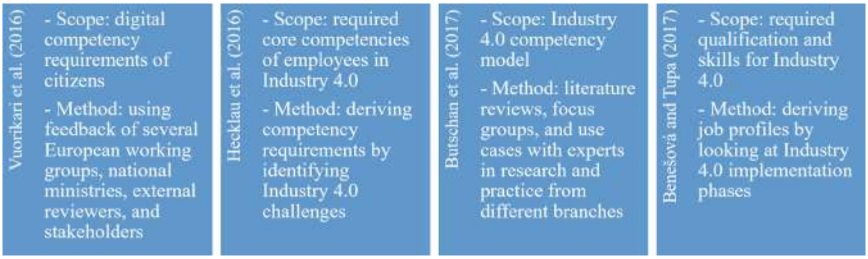

Fig. 4. Approaches of defining digital competencies.

The European Commission's in-house science service, the Joint Research Center (JRC), published in 2016 an update to the first in 2013 presented European Digital Competence Framework for Citizens (DigComp) (Vuorikari et al. 2016). Based on feedback of expert working groups from the European Commission, national ministries, external reviewers, and stakeholders, five areas of digital competencies were elaborated: information and data literacy (e.g. evaluating and managing data), communication and collaboration, digital content creation (among others: copyright and licenses, as well as programming), safety (e.g. protecting personal data and privacy or health and well-being), problem solving (creativity and identifying one's own digital competency gaps, among others).

Hecklau et al. (2016) derive four aggregated categories of required core competencies based on identified challenges arising with Industry 4.0 due to the literature: technical, methodological, social, and personal competencies. Technical competencies comprise such as state-of-the-art knowledge, media, or coding skills. Furthermore, creativity, problem, analytical, or research skills are necessary amongst other methodological competencies. The ability to transfer knowledge, leadership skills, flexibility, and the motivation to constantly learn are other social and personal competence requirements due to digitized and interconnected working environments.

Butschan et al. (2017) use a very thorough research design for developing their own Industry 4.0 competency model. They are applying mixed methods, starting with a literature review, and conducting a quantitative study as well as interviews with a focus group consisting of experts of different branches from research and practice. Finally, they are testing their findings in two use cases.

Benešová and Tupa (2017) use a different approach in presenting their findings. These authors describe necessary skills for relevant job profiles in Industry 4.0. Initially, a distinction is made between IT and production job profiles. Afterward, the job profiles of robot programmers, data analysts, or production technicians are further described in detail.

Several competencies required of workers for highly automated and digitalized working environments like Industry 4.0 are mentioned by several of these studies. Similarities and differences between these identified competencies can be seen in Table 1, where they are aggregated and subsumed under four categories for better visualization: technical, methodological, social and communication, and personal competencies. The differences between these studies are particularly noteworthy.

E.g., solely Butschan et al. (2017) explicitly mention eagerness to experiment, and openness for change as necessary competencies for production workers to adapt to changing processes. Especially interesting is the comparison of the derived competencies of these studies with the actual observations at IFX – differentiated by the three generic types of workplaces (Table 1). Such competencies like media skills or state-of-the-art knowledge apply to all of the workplaces, as well as the necessary openness for change to cope with the fast advances in digitalization.

Table 1. Comparison of competencies for digitalized working environments in literature and own findings – differentiated between development, production, and maintenance workplace.

Category	Competencies	References	Verified by interviews and observations		
			Dev.	Prod.	Main.
Technical competencies	State-of-the-art knowledge	Hecklau et al. (2016); Butschan et al. (2017)	x	x	x
	Technical skills	Hecklau et al. (2016); Benešová and Tupa (2017); Vuorikari et al. (2016)	x	x	x
	Process understanding	Hecklau et al. (2016); Benešová and Tupa (2017)	x	x	x
	Media skills	Hecklau et al. (2016); Benešová and Tupa (2017); Vuorikari et al. (2016)	x	x	x
	Coding skills	Hecklau et al. (2016); Benešová and Tupa (2017)	x		(x)
	Understanding IT security	Hecklau et al. (2016); Benešová and Tupa (2017); Vuorikari et al. (2016)	x		
	Multidisciplinary knowledge	Butschan et al. (2017)	x		x
Methodological competencies	Creativity and innovative spirit	Hecklau et al. (2016); Benešová and Tupa (2017); Butschan et al. (2017); Vuorikari et al. (2016)	x		x
	Entrepreneurial thinking	Hecklau et al. (2016)	x		
	Problem solving	Hecklau et al. (2016); Benešová and Tupa (2017); Butschan et al. (2017); Vuorikari et al. (2016)	x	x	x

(*continued*)

Table 1. (*continued*)

Category	Competencies	References	Verified by interviews and observations		
			Dev.	Prod.	Main.
	Conflict solving	Hecklau et al. (2016)	(S)		
	Decision making	Hecklau et al. (2016)	x	(S)	x
	Analytical skills	Hecklau et al. (2016); Benešová and Tupa (2017); Vuorikari et al. (2016)	x		x
	Statistical skills	Benešová and Tupa (2017)	x		
	Research skills	Hecklau et al. (2016)	x		x
	Efficiency orientation	Hecklau et al. (2016)	x	x	x
	Organizational skills	Benešová and Tupa (2017); Butschan et al. (2017)	x	(S)	x
	Conceptual strength	Butschan et al. (2017)	x		
Social and communication competencies	Intercultural skills	Hecklau et al. (2016)	x	(x)	(x)
	Language skills	Hecklau et al. (2016); Benešová and Tupa (2017)	x	(x)	(x)
	Communication skills and dialogue capability	Hecklau et al. (2016); Benešová and Tupa (2017); Butschan et al. (2017); Vuorikari et al. (2016)	x		x
	Networking skills and relationship management	Hecklau et al. (2016); Butschan et al. (2017)	x		x
	Ability to work in a team	Hecklau et al. (2016); Vuorikari et al. (2016)	x		x
	Ability to be compromising and cooperative	Hecklau et al. (2016); Benešová and Tupa (2017); Butschan et al. (2017)	x		
	Ability to transfer knowledge and to teach	Hecklau et al. (2016); Butschan et al. (2017)	x	x	x
	Leadership skills (staff development incl.)	Hecklau et al. (2016); Benešová and Tupa (2017); Butschan et al. (2017)	(S)	(S)	(S)

(*continued*)

Table 1. (*continued*)

Category	Competencies	References	Verified by interviews and observations		
			Dev.	Prod.	Main.
	Responsibility	Benešová and Tupa (2017)	x	x	x
	Reliability	Benešová and Tupa (2017)	x	x	x
Personal competencies	Eagerness to experiment	Butschan et al. (2017)	x		(x)
	Flexibility and mobility	Hecklau et al. (2016); Benešová and Tupa (2017); Butschan et al. (2017)	x		x
	Openness for change and adaptability	Butschan et al. (2017)	x	x	x
	Holistic thinking	Butschan et al. (2017)	x		
	Initiative	Butschan et al. (2017)			(x)
	Autonomy	Benešová and Tupa (2017)			
	Ambiguity tolerance	Hecklau et al. (2016)	x		
	Motivation to learn	Hecklau et al. (2016); Benešová and Tupa (2017); Butschan et al. (2017)	x	x	x
	Ability to work under pressure	Hecklau et al. (2016)		x	x
	Sustainable mindset	Hecklau et al. (2016); Vuorikari et al. (2016)	x	x	x
	Compliance	Hecklau et al. (2016)	x		

x = match, (x) = partially match, (S) = match on supervisor level

Whereas, conflict solving competencies with customers were not directly observed, but might be necessary foremost at development supervisor levels. Furthermore, coding skills were only relevant for developers for sure, and partially for maintenance workers that deal with IT issues. Also worth mentioning is taking the initiative as part of required personnel competencies (Benešová and Tupa 2017). This was also observed e.g. at IFD, but the initiative proposals of the workers for improvement were or could not always be acknowledged by the higher instances. A sustainable mindset – including considerations of health and environmental issues (Vuorikari et al. 2016) – were foremost observed at those workplaces, where the physical amount of work is higher, i.e. at production and maintenance workplaces (Fig. 3). The workers at these workplaces showed some concerns towards worsening their health by the application of digital technology.

Matching the experienced competency requirements at the respective workplaces with those mostly general formulated ones in literature, shows that not all of the types of workplaces necessitate all of the competencies – or at least in different gradations. All in all, the comparison of these workplaces show that the production workers seem to need fewer competencies as the maintenance staff and the developers, where almost all of the identified competencies were as well observed. However, these findings underlie the restriction that not having competencies observed or derived does not mean that they were not necessary. Furthermore, the extent to which certain competencies are required might depend on the level of automation of the fab and the used technologies as well.

For a more thorough examination of the required competencies, systematically assessing those using measurements is obligatory. Therefore, in the following section different means of assessing competencies are evaluated. In addition, the comparison of competencies in literature and as observed, point at the issue of clearly separating qualifications and competencies from one another. Qualifications foremost stand for the ability to act and react in routine work tasks. And competencies represent the ability to be empowered to confidently react to unknown issues as well, and to apply and transfer them in new situations in work (Sauter and Staudt 2016b). Furthermore, personality traits do not necessarily correspond to competence patterns. But in times of digitalization and ever faster innovation cycles of new technologies especially these competencies are of relevance to empower the employees for an effective adaption.

3.3 Measuring Digital Competency Gaps and Deriving Actions to Close Them

Methods for Competency Measurement. For thoroughly assessing competencies of the employees there exist several approaches in literature and practice, especially in human resource management, which are shortly presented in the following. The main goal of competency measures is to depict the current state of the competency of an employee, and to compare it with the desired one. This can happen for finding a suitable candidate for a job offering or for training employees for new tasks. For visualization of the results usually a radar chart is used. From a practitioners' point of view special interest lies on the practicability of the different methods in terms of necessary effort for applying them. Therefore, Table 2 presents different approaches ordered by the expected expenditure and complexity. More complex tools are assessment centers, whereas less complex ones are e.g. test procedures or self-assessments. However, the expected expenditure does not necessarily allow any conclusions to be drawn about the success of the methods in measuring competencies.

Table 2. Methods for competency measurement and relative complexity.

Method	Exemplary references	Complexity
Assessment Center (AC)	Sauter and Staudt (2016b); Eck et al. (2016)	
Biographical Methods - Certificates - Standardized Questionnaires - Standardized Interviews - Competency Biography *(Erpenbeck and Heyse)* - Competency Balance *(Psychological Models)*	Sauter and Staudt (2016b)	
Work Samples	Sauter and Staudt (2016b)	
Activity Analysis	Sauter and Staudt (2016b); Beuscher-Mackay et al. (2009)	
Interviews	Sauter and Staudt (2016b)	
Personality Models - E.g. "Big Five"	Sauter and Staudt (2016b)	
Computer Aided Simulations	Sauter and Staudt (2016b); Sauter and Staudt (Sauter et al. 2016a)	
External Assessment	Sauter and Staudt (2016b); Rietiker (2010)	
Self-Assessment	Sauter and Staudt (2016b); Hillebrand (2018)	
Test Procedures	Sauter and Staudt (2016b)	

After having identified potential competency gaps of people for certain Industry 4.0 or workplace requirements, the next question is how to raise these competencies effectively.

Training and Educational Approaches. Choosing the suitable educational and training methods for empowering and developing the competencies of the employees, different factors have to be considered. E.g., do the measures address employees or maybe trainees or future employees? Or how effective, complex, and time-consuming are they for the wished purposes? Besides classical pedagogical concepts for increasing competencies in digital transformation, training the workers by means of digital technologies becomes more important. Here, two suitable possibilities for this purpose are briefly mentioned: project-based learning (PBL) scenarios and learning factories.

PBL scenarios are characterized by five aspects (Krajcik and Blumenfeld 2006):

1. As starting point there is a practical problem to be solved,
2. students explore the problem on their own and learn important skills,
3. students and teachers work together to solve the problem,

4. students use (learning) technologies to give them knowledge beyond their current abilities, and
5. students develop products that solve the problem.

In accordance with this approach, Spitzer and Ebner (2017) for example applied CAD tools, 3D printing, and smart glasses to successfully teach students the usage of these tools and devices without prior knowledge. In addition, the students showed high motivation during the project.

For further hands-on education and training learning factories (Schallock et al. 2018) get more attention, as they allow the immediate application and visualization of digital technologies, as well as its use and possible issues. Abele et al. (2015) present a morphology of different types of learning factories that may vary in seven categories, such as operating model, purpose and targets, process, setting, product, didactics, and learning factory metrics. Concerning didactics, there are variabilities with respect to the learning targets (cognitive, affective, psycho-motorical), type of learning environment (greenfield, brownfield), role of trainer (instruction, demonstration, closed scenario, open scenario), evaluation, communication (onsite, remote), and others. This illustrates, that there are already plenty of possibilities that could be adapted to develop suitable environments for certain Industry 4.0 requirements.

4 Discussion and Implications of the Findings

The paper shows that the three generic types of workplace in semiconductor production – development, production, and maintenance workplace – differ in work- and information-related aspects like the necessary amount of collaboration, or the variety of sources of information needed for planning and decision making at work. Furthermore, digitalization so far has impacted the workplaces regarding these factors differently. The explorative interviews and observations allow as well to draw preliminary conclusions and make new hypotheses for further investigation and validation, e.g.:

– When are the workers used to new technologies (time of adaption), and when might new technologies be perceived as a burden?
– Does introducing new technologies and systems automatically mean extra effort by the workers that will only be reduced after they have adapted to it?
– How can information assistance systems help to enhance the human capabilities such as creativity at work and to eliminate frustrating factors?

A distinction seems to be possible between job specific activities that are contributing to the well-being and motivation of the workers, and those activities that are rather annoying and frustrating. Therefore, the focus of digitalization should be on processes and activities to eliminate those uncomfortable elements of the workflow, and not to extinct those that are perceived as motivating by the employees. Moreover, the personal traits of the employees seem to be an important factor for the adaption to new technologies.

As the employees are aware of the rapidly changing working environment due to digitalization, it is plausible that they are more open-minded towards ongoing training

– as they see the need for it. In designing appropriate training and education programs to enhance workers' digital competencies, a distinction should be made between people that are already working, and those that are still in vocational training or higher education. Hence, competencies must be made measurable to identify competency gaps and the need for action for individual employees, firstly. Therefore, several methods of assessing competencies have been presented by differentiating them in terms of effort in practice. The comparison of existing competency models with the observations at the workplaces in the semiconductor industry so far have shown that there are similarities, but differences as well. And again, not all of the literature-based competencies for Industry 4.0 are relevant to all of the workplaces, or at least in different forms.

Afterward, further education programs in-house and university courses aligned with these changing job profiles due to digitalization can be developed to close the identified competency gaps. Future academic teaching might also make use of digital technologies and integrate them into the curriculum. E.g., some new course formats that combine digital and analog teaching and learning aspects proved to lead to better results of the students (Handke 2018). Especially the results regarding 5V of Big Data can be used to conceptualize information assistance systems to support the workers in their activities, and to evaluate where such systems would make sense – and where they would not. Example cases of testing innovative assistance systems are the use of Google Glasses or gesture recognition in the quality assurance of BMW (QZ-onlinde.de 2014; QZ-online.de 2015). Nonetheless, a critical view of such digital assistance systems is mandatory to avoid digitization for its own sake. They should enhance workers' skills at reasonable costs for the company without overstraining its employees.

For further validation of the results of this paper, further observations and interviews at the aforementioned three types of workplaces will be conducted at semiconductor manufacturers at different maturity levels of digitalization. In addition, a digitalization lab at the Zittau/Goerlitz University of Applied Sciences will be installed to test implementing Industry 4.0 technologies in hands-on higher education and training programs.

Acknowledgement. The project iDev40 has received funding from the ECSEL Joint Undertaking under grant agreement No. 783163. The JU receives support from the European Union's Horizon 2020 research and innovation programme. It is co-funded by the consortium members, grants from Austria, Germany, Belgium, Italy, Spain and Romania. It is coordinated by Infineon Technologies Austria AG.

Appendix A

iDev40 – Interview

August 9, 2018, Villach

iDev40

Preamble:

Within our work package 4 of the project "iDev40" we are concerned, among other things, with finding design solutions to develop the workplace of the future. To do this, it is important that we take personal assessments of the employees, to get to know how to proceed in a sensible way. For this purpose, we would like to know your opinion about the impact of digitisation on your work. At this basis, we can design digitisation solutions in such way that they are useful for your work.

Procedure and questions:

Part 1

1. Thank you for your time. First, a few general questions:
2. What is your role in the company?
3. Can you briefly describe your typical tasks on a working day?
4. Does digitisation already play a role at your workplace?
5. What do you generally think of the term digitisation in the working environment?
6. In what way could digitisation be useful for your work? Can you imagine, that at certain points this means extra work for you? Do you see the topic as an opportunity or risk to your occupation?
7. How do you think digitalisation will affect your personal work in the future?
   a. Are your tasks becoming simpler or more complex? (Why?)
   b. Are your tasks becoming more or less interesting? (Why?)
   c. Do your tasks become more comfortable or less comfortable to perform? (Why?)
   d. Is more or less cooperation with colleagues necessary? (Why?)

Part 2 - Big Data

When answering the following questions, think of the most important (or most common) task you have.

Volume

8. Do you have a large amount of data and information to process?

Variety = variety of data (sources, structures)

9. Are the types of data you deal with very different?
10. How many different sources does the data come from?
11. Is the data available in raw form or does it need further processing?

Velocity = speed of change in data provision and data use

12. How often do you need to retrieve this data?

Veracity = are data available in the right quality, in the right place, at the right time

13. Do you have all the information you need for your work (If “no”: Which ones are missing?)
14. Are the data you use for your work available in the required quality?

(Replacement question: Is there generally information that you have to search for a long time, or that is not available so that you can deal with it?)

15. Is the data available at your workplace exactly when you need it?
16. Do you often have to involve third parties in retrieving your data?
17. Is there information that you wish you had available because then your work would be easier?

Value = creating value from data for companies

18. With the things we were just talking about, can you imagine changes through digitalisation?

References

Abele, E., Metternich, J., Tisch, M., Chryssolouris, G., Sihn, W., ElMaraghy, H., et al.: Learning factories for research, education, and training. Procedia CIRP **32**, 1–6 (2015). https://doi.org/10.1016/j.procir.2015.02.187

Benešová, A., Tupa, J.: Requirements for education and qualification of people in industry 4.0. Procedia Manufact. **11**, 2195–2202 (2017). https://doi.org/10.1016/j.promfg.2017.07.366

Beuscher-Mackay, S., Jeske, T., Steiger, P., Hinrichsen, S., Schlick, C.M.: Tätigkeitsanalyse und Personalentwicklungskonzepte. In: Michael Schenk, C.M.S. (ed.) Industrielle Dienstleistungen und Internationalisierung: One-Stop Services als erfolgreiches Konzept, pp. 219–265, Gabler, Wiesbaden (2009). https://doi.org/10.1007/978-3-8349-8794-5_12

Butschan, J., Nestle, V., Munck, J.C., Gleich, R.: Kompetenzaufbau zur Umsetzung von Industrie 4.0 in der Produktion. In: Mischa, S., Lars, G., Sebastian, B. (eds.) Betriebswirtschaftliche Aspekte von Industrie 4.0, pp. 75–110. Springer Fachmedien Wiesbaden, Wiesbaden (2017). https://doi.org/10.1007/978-3-658-18488-9_4

Chen, M., Mao, S., Liu, Y.: Big data: a survey. Mob. Netw. Appl. **19**(2), 171–209 (2014). https://doi.org/10.1007/s11036-013-0489-0

Dautov, R., Distefano, S.: Quantifying volume, velocity, and variety to support (big) data-intensive application development. In: Jian-Yun, N., Zoran, O., Toyotaro, S., Rumi, G., Raghunath, N., Chonggang, W. (eds.) 2017 Proceedings of IEEE International Conference on Big Data, Boston, Piscataway, 11–14 December 2017, pp. 2843–2852. IEEE (2017)

Eck, C.D., Jöri, H., Vogt, M.: Das Assessment-Center-Verfahren und seine Durchführung. In: Claus, D.E., Hans, J., Marlène, V. (eds.) Assessment-Center: Entwicklung und Anwendung – mit 57 AC-Aufgaben und Checklisten, pp. 41–112. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-47742-7_2

Handke, J.: Digitale Hochschullehre – Vom einfachen Integrationsmodell zur Künstlichen Intelligenz. In: Dittler, U., Kreidl, C. (eds.) Hochschule der Zukunft, pp. 249–263. Springer Fachmedien Wiesbaden, Wiesbaden (2018)

Hecklau, F., Galeitzke, M., Flachs, S., Kohl, H.: Holistic approach for human resource management in industry 4.0. Procedia CIRP **54**, 1–6 (2016). https://doi.org/10.1016/j.procir.2016.05.102

Hillebrand, A.: Studie 2 – Fragebogenstudie. In: Ariane, H. (ed.) Welche Kompetenzen zeichnen einen Experten aus? Entwicklung eines expertenspezifischen Kompetenzmodells in einem produzierenden Unternehmen, pp. 109–223. Springer Fachmedien Wiesbaden, Wiesbaden (2018). https://doi.org/10.1007/978-3-658-22546-9_4

iDev40: Project Work. Work Plan (2019). http://www.idev40.eu/project-work/work-plan. Accessed 1 July 2019

Ilie-Zudor, E., Ekárt, A., Kemeny, Z., Buckingham, C., Welch, P., Monostori, L.: Advanced predictive-analysis-based decision support for collaborative logistics networks. Supp. Chain Manag. **20**(4), 369–388 (2015). https://doi.org/10.1108/SCM-10-2014-0323

Kagermann, H., Wahlster, W., Helbig, J.: Recommendations for implementing the strategic initiative INDUSTRIE 4.0. Final Report of the Industrie 4.0 Working Group. Securing the future of German manufacturing industry (2013)

Kern, E.-M.: Verteilte Produktentwicklung. In: Lindemann, U. (ed.) Handbuch Produktentwicklung, pp. 455–481. Hanser, München (2016)

Krajcik, J.S., Blumenfeld, P.C.: Project-based learning. In: Keith Sawyer, R., (ed.) The Cambridge Handbook of the Learning Sciences, pp. 317–333. Cambridge University Press, Cambridge, New York (2006)

Kuckartz, U.: Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung. 3., überarbeitete Auflage. Weinheim, Basel: Beltz Juventa (Grundlagentexte Methoden) (2016). http://www.beltz.de/de/nc/verlagsgruppe-beltz/gesamtprogramm.html?isbn=978-3-7799-3344-1

Martin, K.: Beobachtungsinterview. In: Kühl, S. (ed.) Handbuch Methoden der Organisationsforschung. Quantitative und qualitative Methoden. 1. Aufl. Wiesbaden: Verl. für Sozialwiss./GWV Fachverl, pp. 78–99 (2009)

QZ-onlinde.de (ed.): Qualitätskontrolle per Geste (2014). https://www.qz-online.de/news/uebersicht/artikel/industrie-4-0-qualitaetskontrolle-per-geste-873180.html. 21 Sept 2018

QZ-online.de (ed.): BMW. Qualitätssicherung mit Google-Brille (2015). https://www.qz-online.de/news/uebersicht/nachrichten/bmw-qualitaetssicherung-mit-google-brille-976281.html. 21 Sept 2018

Rietiker, J.: Auswahl von Personal. In: Werkmann-Karcher, B., Rietiker, J. (eds.) Angewandte Psychologie für das Human Resource Management: Konzepte und Instrumente für ein wirkungsvolles Personalmanagement, pp. 215–237. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-12481-5_11

Sauter, W., Staudt, A.-K.: Kompetenzmessung in der Praxis. In: Sauter, W., Staudt, A.-K. (eds.) Kompetenzmessung in der Praxis. Mitarbeiterpotenziale erfassen und analysieren, pp. 33–44. Springer Fachmedien Wiesbaden, Wiesbaden (2016a). https://doi.org/10.1007/978-3-658-11904-1_4

Sauter, W., Staudt, A.-K.: Methoden der Kompetenzmessung. In: Sauter, W., Staudt, A.-K. (eds.) Kompetenzmessung in der Praxis. Mitarbeiterpotenziale erfassen und analysieren. 1. Aufl. 2016, pp. 7–21. Springer Gabler (essentials), Wiesbaden (2016b)

Schallock, B., Rybski, C., Jochem, R., Kohl, H.: Learning factory for industry 4.0 to provide future skills beyond technical training. Procedia Manuf. **23**, 27–32 (2018). https://doi.org/10.1016/j.promfg.2018.03.156

Schöning, H., Dorchain, M.: Data mining und analyse. In: Thomas, B., ten Hompel, M., Vogel-Heuser, B. (eds.) Industrie 4.0 in Produktion, Automatisierung und Logistik. Anwendung, Technologien, Migration, pp. 543–554. Springer Fachmedien Wiesbaden, Wiesbaden (2014)

Spitzer, M., Ebner, M.: Project based learning. From the idea to a finished LEGO® technic artifact, assembled by using smart glasses. In: Proceedings of EdMedia, World Conference on Educational Media and Technology 2017, Association for the Advancement of Computing in Education (AACE), pp. 196–209 (2017)

Vuorikari, R., Punie, Y., Carretero, S., van den Brande, L.: DigComp 2.0. The digital competence framework for citizens. update phase 1: the conceptual reference model. EUR 27948 EN. Publication Office of the European Union, Luxembourg (2016)

Training the Human-in-the-Loop in Industrial Cyber Ranges

Stela Kucek and Maria Leitner(✉)

Center for Digital Safety & Security, AIT Austrian Institute of Technology, Vienna, Austria
{stela.kucek.fl,maria.leitner}@ait.ac.at

Abstract. With the trend of automation in manufacturing and the advancements of technologies, knowledge, skills and abilities of the workforce should develop accordingly. Current training technologies often do not provide hands-on training and exercises. Hence, training methods and technologies need to adapt to support the new requirements and progress. Cyber ranges are virtual environments that mimic realistic networks and systems and can be used for e.g., training, exercises or research. While current state of the art focuses mostly on technical designs and developments, this paper focuses on assessing and integrating the human-in-the-loop in industrial cyber ranges (i.e. cyber ranges with specific scenarios that can be found in Industry 4.0, manufacturing or related topics). We describe the human-in-the-loop in cyber ranges and outline an example application scenario. Furthermore, we discuss challenges in relation to the implementation in cyber ranges. For future work, we will utilize this design and development scheme for further advancements of industrial cyber ranges and its components.

1 Introduction

With digitization in industry and the shift towards smart manufacturing, industrial control systems (ICS) and their interconnectedness has increased the likelihood of potential attacks. In recent years, attacks such as Stuxnet or Crashoverride have demonstrated that security is a key factor in ICS. Waslo et al. (2017) state that enhancing digital capabilities in the production and supply chain processes can lead to new cyber risks. For example, exploiting SCADA (Supervisory control and data acquisition) system vulnerabilities may be a potential threat (Ralston et al. 2007).

However, organizations in many domains such as manufacturing, energy, or logistics are facing the challenges with the increased speed of evolving threats and vulnerabilities as well as preventing, detecting and mitigating (cyber) security incidents. The aim of many organizations is to increase their resilience by increasing their preparedness and response times of (cyber) incidents. When an unexpected or unwanted event (i.e. incident) occurs in a highly automated environment, the responsible human operator should be able to react in a fast and efficient manner. For this fast and efficient response, new training methods need to be assessed to enable dynamic, real-time training and exercises. In fact, Weyer et al. (2015) state that new teaching and training platforms have to be developed in order to train and support new qualifications

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 107–118, 2020.
https://doi.org/10.1007/978-3-030-48602-0_10

in cyber-physical systems (CPS) (often also referred to as Industry 4.0 – the fourth industrial revolution). These new platforms would support the adequate training and preparation of employees (from administrative staff, engineers, computer security incident response (CSIRT) teams to the management board) to increase security awareness and skills as well as providing novel approaches to capability management in organizations. This would support organizations, as they often provide inhouse training to increase the skills and abilities of the workforce as it is in general challenging to find highly-skilled workers (see (Neuman 2009)). Current training methods support only rarely realistic simulations of the business operations, technology and their impact (e.g., e-learning).

In this paper, we introduce cyber ranges as a training platform to the manufacturing domain in order to support the aforementioned new qualifications and challenges. The NIST (National Institute of Standards and Technology, U.S. Department of Commerce 2018) specifies cyber ranges as interactive, simulated representations of an organization's network, system tools, and applications that are connected to a simulated Internet level environment. Hence, cyber ranges can simulate information technology (IT) and operational technology (OT) infrastructures, benign and malicious users as well as (cyber) incidents (Frank et al. 2017). On top of these virtual environments, training scenarios are practiced (Davis and Margath 2013; Pham et al. 2016). The advantage is that this hands-on training is conducted in a safe environment and not the production systems. In the context of cyber-physical systems in manufacturing, cyber ranges can be utilized to train a target audience e.g., to operate an infrastructure, to practice relevant procedures, as well as to handle security-related incidents. In this paper, we aim to describe how cyber ranges can be adapted for training and exercises in manufacturing incorporating the human-in-the-loop.

Therefore, we review current trends and aspirations in technical training and exercises in manufacturing and introduce cyber ranges as a technology platform for training and exercises. With this, organizations may enable hands-on training for employees (e.g., engineers, incident response teams, managers, etc.). This training enables more practice than usual training situations (e.g., lecturers or e-learning) and supports the training of technical (e.g., knowhow and practice of networks and systems) and organizational topics (e.g., contingency processes or hierarchical structures). The advantages of using virtual environments are not only the flexibility and reusability of scenarios but also that training can be conducted anywhere (using an Internet-enabled environment). Hence, training can be conducted on premises but also with virtual and distributed teams. Based on current state of the art, we describe cyber ranges and potential application areas and introduce the human-in-the-loop concept. Furthermore, we outline an application scenario that showcases the concept of the human-in-the-loop in cyber ranges. We discuss current challenges for the implementation and evaluation in cyber ranges.

The rest of this paper is structured as follows: Sect. 2 outlines related work and background on cyber ranges, training technologies in manufacturing and the human-in-the-loop. Section 3 introduces cyber ranges and their applications. In Sect. 4, the human-in-the-loop is described within industrial cyber ranges. Section 5 presents an example scenario that demonstrates the concept. Section 6 discusses the content and implications of the approach and concludes the paper.

2 Related Work

Cyber security management has become very important in ICS. For example, Knowles et al. (2015) provide a review of cyber security management in ICS. Fabro et al. (2016) suggest that one of the five key countermeasures of ICS security is to *manage the human*. This includes providing ICS security training for all operators and administrators. As this paper centers on training technologies, we briefly review technical hands-on training methodologies and technologies in the context of CPS and cyber ranges and do not focus on e-learning, augmented or virtual reality approaches. Weyer et al. (2015) state that there is a need for new teaching and training platforms due to the need of new qualifications in Industry 4.0. This can be seen in the recent developments.

2.1 Physical Environments: Learning Factories

Schallock et al. (2018) describe a design of a learning factory that focuses not only on technical skills, but also trains decision making, group work and performance monitoring skills of the production staff. They present a conceptual design that encompasses theoretical, as well as practical aspects applied in a German learning factory. Simons et al. (2017) propose a similar concept, namely a holistic, fully automated Industry 4.0 learning factory. These works provide a physical environment (or single physical devices such as PLCs and robots) to demonstrate and train production processes and not a virtual environment such as provided with cyber ranges or other virtual testbeds.

2.2 Virtual Environments: ICS Testbeds and Cyber Ranges

Research and industry has developed virtual environments for training and experimentation in ICS. Holm et al. (2015) provide a systematic review of 30 ICS testbeds. Common methods for implementing such environments are hardware, simulation, emulation and virtualization (Holm et al. 2015). Typical testbed components are the control center, communication architecture, field devices, and the physical process. The authors suggest using taxonomies to make ICS testbeds more tangible and comparable. Common objectives are to enable vulnerability analysis, education and tests of defense mechanisms (compare (Siddiqi et al. 2018)). While the work of (Holm et al. 2015) focuses on systemizing and categorizing ICS testbeds, in this paper, we formulate and integrate a human-in-the-loop concept for industrial cyber ranges.

Cyber ranges are virtual environments that simulate or mimic IT and OT systems and networks. Many cyber ranges are used in the public domain for research, training, exercises and a review of cyber ranges can be found in (Davis and Magrath 2013). Cyber ranges do not automatically include ICS, however, many of the cyber ranges started to integrate components or subsystems that represent OT. Please refer to Sect. 3 for more information on cyber ranges.

Note that in practice other approaches to training for ICS exist, but either use different technologies (physical devices, virtual reality) or have a different goal (education, technical testing of the system, evaluating low-level processes).

2.3 The Human-in-the-Loop Concept

The concept of a "*human-in-the-loop*" (HitL) has been a well-known principle, representing a model where the interaction of a human actor is required (Karwowski 2006). Its application can be found in well-known examples such as flight or marine simulators. Recently, its application can be found in many other domains such as aviation automation (Bilimoria et al. 2018), system design and modeling (Smith et al. 2018), virtual reality (Sherman and Craig 2018) and machine learning (Warrier and Devasia 2017). The HitL has been also introduced to cyber-physical systems in (Schirner et al. 2013) where body or brain sensors interact with an embedded system and a physical environment. Stouffer et al. (2014) suggest that the HitL concept can be used as a form of control and supervision of the system rather than only automated supervision in ICS. This includes also manual modes of ICS where humans completely control the systems (opposite to open-loop or closed-loop control systems). As a typical ICS contains numerous control loops, human intervention and supervision may be required and a specific skillset for the interaction with these systems. Hence, this mix of control systems might bring a new perspective to training incorporating the semi-automated control systems that require human supervision and interaction. This would align with Weyer et al. (2015) and their proposed need for new qualifications in smart manufacturing. In addition, with the HitL concept, we aim to address the significance of a human operator in ICSs, as well as their competences and skills required in this context.

Hence, this brief review of related work displays that there is a need for dynamic and hands-on training in cyber security in ICS and a vivid research on environments that support various purposes such as education, training or research. While many of these papers focus on the design and development of infrastructures, to the best of our knowledge, none of the papers integrate the human-in-the-loop in these systems. Surprisingly, we also found little or no literature on the concept of hands-on technical exercises in cyber ranges in manufacturing. However, dynamic exercises will become more essential, particularly for coping with daily or unexpected situations (e.g., cyber incidents).

3 Cyber Ranges for Smart Manufacturing

In general, cyber ranges are virtual environments that simulate systems and critical (cyber) security incidents on these systems. Cyber ranges can simulate various technologies such as IT and OT infrastructures. They have been widely used in cyber security training and education, cyber security competitions and challenges, national cyber security exercises and more (Davis and Magrath 2013). In addition to existing ICS testbeds (see Sect. 2), many cyber ranges integrate ICS components as part of a larger and complex system and to mimic various attack vectors. Hence, ICS testbeds may overlap with or be part of cyber ranges. Furthermore, realistic scenarios are simulated on top of these environments. This can include multiple aspects such as benign and malicious users, network traffic, external stakeholders or sites, security incidents and more. Figure 1 displays a general scheme of components of cyber ranges.

However, each individual cyber range may have a similar or different configuration. It can be seen from the figure that there is a simulated infrastructure of networks and systems, supported processes or contingency plans, simulated stakeholders, simulated external sites and participants (blue and red teams). The blue teams can be, for example, the trainees of a training and the red team can be also trainees who aim to infiltrate the system (or could be replaced with a program).

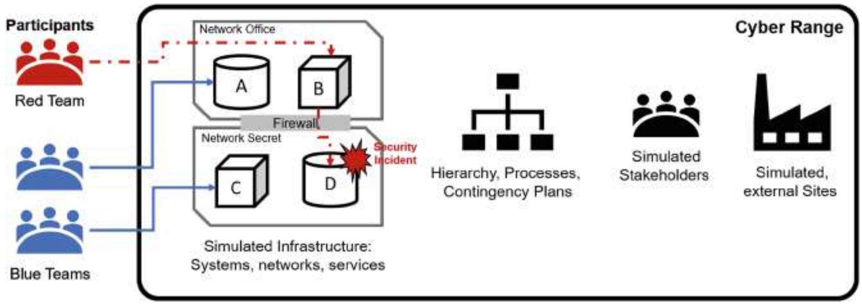

Fig. 1. Cyber range schema

This setup allows for manifold activities to be conducted on, the two most common activities are currently:

- Training and exercises: Cyber ranges can provide a multidimensional way of training cyber security or information security staff and general staff. In particular, this hands-on training is heavily requested by organizations (Frank et al. 2017) for cyber security staff. In addition, training and exercises on cyber ranges allow for dynamic situations, like the real world. Training supports security awareness trainings but also technical trainings and exercises. In manufacturing this could be e.g., training daily or rare situations of business operations.
- Research and experimentation: Cyber ranges support experimentation and testing of software, hardware, as well as human reactions and therefore can be used in research. Many cyber ranges have been developed that have this background (see e.g., (Davis and Magrath 2013)). For example, the virtualization of certain automation processing components (and their impact on production in case of incidents) or the handling of stressful situations could be reviewed in cyber ranges.

The advantage of using a cyber range is that situations and strategies can be tested within a safe and isolated virtual environment and the outcome does not affect production. Also, the environment allows a lot of practice allowing theory to be manifested with practical and hands-on work. The cyber range approach can be applied and adapted to any industrial domain e.g., from water supply, logistics, energy to manufacturing. Hence, in cyber-physical systems, cyber ranges can be utilized to e.g., exercise contingency plans in case of security incidents, operating infrastructures, testing relevant procedures and others. They may be further called “industrial cyber ranges” as they might have specifics and scenarios focusing only on smart manufacturing and related challenges.

4 The Human-in-the-Loop in Cyber Ranges

In this section, we aim to describe a new concept for training and building a capable and highly skilled workforce in Industry 4.0. The core of this idea is to simulate real situations in a safe environment with the use of user- and skill-specific scenarios in the context of modern industry. We have introduced cyber ranges and their application in Sect. 3. An industrial cyber range can be composed of virtual components that represent IT and OT infrastructures. Apart from the system itself, there is a human actor (further called participant) interacting with the system, in the industrial context often referred to as an operator (or engineer), and a participant can be a single person or a team. They can be part of various target groups: from manufacturing engineers, control engineers to the shift managers and managers on duty. Applying the HitL principle to industrial cyber ranges, the focus shifts to the operators or engineers and their interaction with the system. In fact, in the training within an industrial cyber range, each trainee or participant is a human-in-the-loop.

The interaction is typically enabled within a scenario. For example, the participant can be confronted with situations that reflect daily business routines but also rare situations (e.g., cyber incident response, cyber attacks, production shutdowns, etc.). Within these scenarios, operators are performing their routine tasks or processes such as interactions with ICS components and machines, communication with other (distributed) teams, review of news and monitoring of controls. The human-in-the-loop with all these interactions is depicted in Fig. 2.

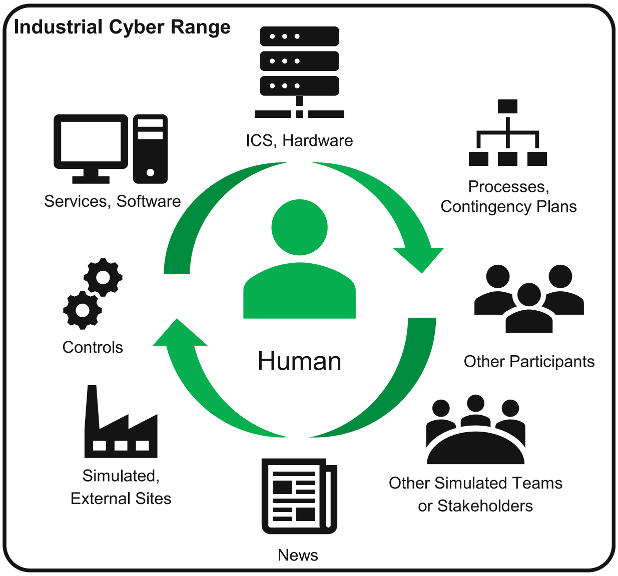

Fig. 2. Human-in-the-loop in industrial cyber ranges

It can be seen from Fig. 2 that the participant is the central point of all actions. All interactions within the cyber range such as with other participants, other simulated actors, controls, ICS and others revolve around the participant.

In a cyber security training or exercise, the participant interacts with the cyber range to manage certain situations. A typical training could e.g., focus on improving reaction times, assessing the quality of decision making and dealing with unwanted or unexpected situations. We outline an example of a scenario for the case of manufacturing in the following section.

5 Application Scenario

This section sketches a (fictive) example scenario that demonstrates how training on the cyber range can be designed. A scenario is a situation where participants need to handle business operations. The participants use the systems and interfaces provided in the cyber range to resolve the situation following regular business standards and procedures. The actions performed by the participant are to be observed and later evaluated. Example scenarios are the failure of one or more critical system components, denial of service, social engineering, and malicious code exploitation (Falco et al. 2002; Tuptuk and Hailes 2018). The following fictive example is based within the manufacturing domain but can also occur in other domains where the same technology is utilized.

5.1 Background

In this scenario, a control system of the robot arm is manipulated from a source that has access to the control device of the robot or the monitoring device communicating with the control system. This scenario stems from examples of potential vulnerabilities of automation systems such as outlined in (ICS-CERT 2015, 2018). In this scenario, the relevant parameters passed to the control device of the robot are incorrect, i.e. the diameter of the coating does not match the size of the semiconductor material currently processed by the robot. As a result, the semiconductor robot misplaces the coating on the semiconductor material which may have a high impact on the production. Potential consequences of this incident are e.g., faulty semiconductor material, overflow of the coating material or damaged surfaces.

5.2 Setup

The goal of this scenario is for the trainee to rehearse an incident handling protocol in case of identified manipulations of an automation component, i.e. robot arm handling semiconductor material. The trainee of our scenario is a control engineer (ENG) who is supervising the automation process via software on their workstation in a fictive organization. The ENGs main task is to manage the incident scenario (described above). Apart from the ENG, there are several roles that act as supporting roles to the ENG. One of them is a local operator team (LOT), i.e. engineers working nearby with physical access to the robot. Next, there is an incident management and response (INC) team that the ENG can contact for technical support and incidents. Lastly, the manager on duty (MOD) is a shift manager.

5.3 Script

The scenario is planned by using a script that consists of the timeline and the associated steps to be performed by the participants (see Table 1). It can be seen from the table that the ENG must perform many steps throughout the scenario to assess the situation, investigate the incident and resolve the incident. INC, MOD and LOT can be simulated stakeholders or active participants in this scenario.

Table 1. Scenario script

Roles	Steps
ENG	1. Identifies unexpected, abnormal behavior of the ICS component on the ENG workstation
ENG	2. Stops or pauses the ICS component 2.1. The controls (i.e., start, stop, pause) and parameters for the arm movement are provided on the ENG workstation interface. ENG uses the provided controls; presses the STOP button to prevent further damage 2.2. The system does not respond to the termination attempts from the interface. ENG informs a MOD. ENG requests a LOT that is close to the location of the incident for assistance 2.3. No LOTs are available, the ENG terminates the arm physically (e.g., by switching off the hardware or unplugging)
ENG	3. Reports the incident to INC and MOD. Documents the circumstances and consequences of incident, measures taken, and status of situation. Sends information to INC
INC	4. Analyzes incident and responds with feedback. There is a defined incident response protocol for the incident. INC sends the recovery steps to the ENG: (1) Assuming the robot arm is off, first identify misconfigured parameters and correct them (2) Check the default parameters for correctness (3) Perform a malware scan of the control-monitor software to detect a malicious program if present (4) If malware was found, respond in the same thread of the initial report and the issue will be forwarded or solved by the support team (5) If no malware was found, start the robot arm again, but monitor continuously to determine if the issue persists (6) If the issue persists, respond in the same thread of the initial report and the issue will be forwarded to the support team
ENG	5. Performs steps provided in the INC feedback and responds with required details that were missing in the initial report

5.4 Implementation in the Cyber Range

The implementation in the cyber range is challenging as many ICS technologies have not been virtualized and therefore, many testbeds develop simulations or emulations to simulate the industrial process or the field devices that are related to it (see e.g., (Holm et al. 2015)). Therefore, we plan to simulate the workstations for ENG using a control monitor interface of the handling robot and for INC and MOD as shown in Fig. 3, which displays the simplified architectural scheme of the scenario. More components such as field devices, control stations, etc. could be integrated in order to make the scenario more realistic. It can be seen from the figure that the ENG workstation is connected to the control center that transmits the ongoing status to the ENG workstation. Furthermore, the control center is connected to the ICS component 1 (e.g., a programmable logic controller, PLC) that mimics the behavior of such an automation processing system (compare smart factory[1] visualizations for example). Additional network communications or connections, even though not outlined in the scheme, are imaginable and would contribute to the realism of the scenario. All communication between participants is conducted by email and telephone.

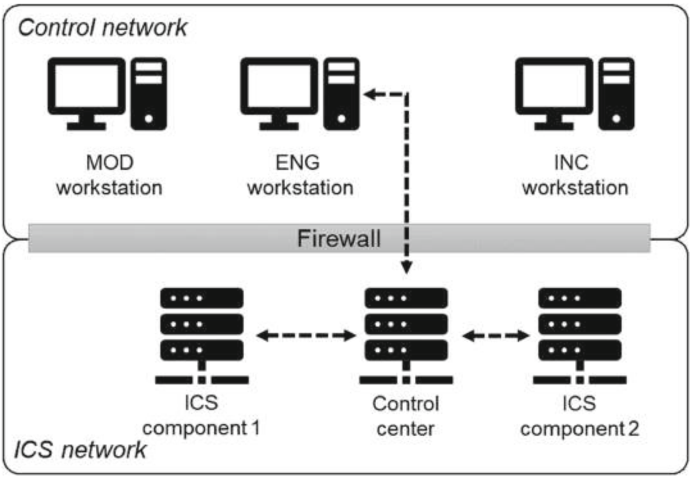

Fig. 3. Simplified scheme of example environment

6 Discussion and Conclusions

Training on a cyber range enables the participants to gain capabilities, to practice new, established or changed routines, as well as to test and evaluate processes and procedures in a safe and isolated environment. Even if the participant does not need to apply

[1] http://www.starteasy.io/ (last access on January 30, 2019).

the gained skills in daily business, the hands-on training will raise awareness and increase the preparedness of participants for these situations. Hence, when these situations occur, participants may react more goal-oriented and efficiently and therefore unwanted situations may be quicker resolved.

One big challenge for cyber ranges and ICS testbeds is the virtualization and simulation of the manufacturing environment. As stated in (Holm et al. 2015), only some technologies have been virtualized and only a few are gradually virtualized to testbeds. Hence, it would be helpful if vendors and others consider providing simulation models or virtualized technologies for research purposes. Another challenge is the simulation of the field devices and the physical process itself (Holm et al. 2015). The utilization of certain devices and process depends heavily on the application scenario (and the defined purpose and goals). In each scenario, designers would have to evaluate if the simulation of field devices or the physical process is really required to support the goals of the training.

Another challenge is the measurement and evaluation of (1) the success of training and exercises, (2) the team performance, (3) the individual participant's performance and (4) the scenario and technology utilized in the cyber range. This is particularly interesting when using different systems and networks in different scenarios. First examples for the evaluation of cyber exercises exist (Schepens et al. 2002). More research needs to adequately assess evaluation measures and procedures for technical training.

In conclusion, the utilization of cyber ranges as training environments to prepare employees for regular and unexpected challenges of business operations is a well-established platform in the cyber security domain. In this paper, we introduced industrial cyber ranges and the integration of the human-in-the-loop. With this, we emphasize on the individual human operator in (manufacturing) systems and the importance of human-machine interactions. We described an application scenario in the semiconductor industry to highlight how the human is the central point of communication and how many activities are conducted in such a hands-on training. With this dynamic and real-time training, we aim to achieve highly-skilled teams that can manage various situations, handle relevant systems, and cope with potential failures. This contributes to the resilience of organizations. For future work, we will use this concept as a basis for designing and implementing a manufacturing use case in the cyber range. After implementation, we plan to evaluate the use case with practical hands-on sessions.

References

Bilimoria, K.D., Hayashi, M., Sheth, K.: Human-in-the-loop evaluation of dynamic multi-flight common route advisories. In: 2018 Aviation Technology, Integration, and Operations Conference, AIAA AVIATION Forum. American Institute of Aeronautics and Astronautics (2018)

Davis, J., Margath, S.: A survey of cyber ranges and testbeds (No. DSTO-GD-0771). Cyber Electronic Warfare Division, DSTO Defence Science and Technology Organisation, Edinburgh, South Australia 5111, Australia (2013)

Fabro, M., Gorski, E., Spiers, N., Diedrich, J., Kuipers, D.: Recommended practice: improving industrial control system cybersecurity with defense-in-depth strategies. Department of Homeland Security's NCCIS and ICS-CERT (2016)
Falco, J., Wavering, A., Proctor, F.: IT security for industrial control systems (No. NIST IR 6859). National Institute of Standards and Technology, Gaithersburg, MD (2002)
Frank, M., Leitner, M., Pahi, T.: Design considerations for cyber security testbeds: a case study on a cyber security testbed for education. In: IEEE 3rd International Conference on Cyber Science and Technology Congress (CyberSciTech). Presented at the CyberSciTech, pp. 38–46. IEEE, Orlando (2017)
Holm, H., Karresand, M., Vidström, A., Westring, E.: A survey of industrial control system testbeds. In: Buchegger, S., Dam, M. (eds.) Secure IT Systems. Lecture Notes in Computer Science, pp. 11–26. Springer, Heidelberg (2015)
ICS-CERT: Rockwell Automation 1766-L32 Series Vulnerability [WWW Document] (2015). https://ics-cert.us-cert.gov/alerts/ICS-ALERT-15-225-02A. Accessed 29 Jan 2019
ICS-CERT: Omron CX-One [WWW Document] (2018). https://ics-cert.us-cert.gov/advisories/ICSA-18-338-01. Accessed 29 Jan 2019
Karwowski, W.: International Encyclopedia of Ergonomics and Human Factors, 2nd edn., 3 Volume Set. CRC Press, Boco Raton (2006)
Knowles, W., Prince, D., Hutchison, D., Disso, J.F.P., Jones, K.: A survey of cyber security management in industrial control systems. Int. J. Crit. Infrastruct. Prot. **9**, 52–80 (2015)
National Institute of Standards and Technology, U.S. Department of Commerce: Cyber Ranges. National Institute of Standards and Technology, US (2018)
Neuman, C.: Challenges in security for cyber-physical systems. In: DHS Workshop on Future Directions in Cyber-Physical Systems Security, pp. 22–24. Citeseer (2009)
Pham, C., Tang, D., Chinen, K., Beuran, R.: CyRIS: a cyber range instantiation system for facilitating security training. In: Proceedings of the Seventh Symposium on Information and Communication Technology, SoICT 2016, pp. 251–258. ACM, New York (2016)
Ralston, P.A.S., Graham, J.H., Hieb, J.L.: Cyber security risk assessment for SCADA and DCS networks. ISA Trans. **46**, 583–594 (2007)
Schallock, B., Rybski, C., Jochem, R., Kohl, H.: Learning factory for Industry 4.0 to provide future skills beyond technical training. In: Procedia Manufacturing, "Advanced Engineering Education & Training for Manufacturing Innovation" 8th CIRP Sponsored Conference on Learning Factories (CLF 2018), vol. 23, pp. 27–32 (2018)
Schepens, W.J., Ragsdale, D.J., Surdu, J.R., Schafer, J., New Port, R.: The cyber defense exercise: an evaluation of the effectiveness of information assurance education. J. Inf. Secur. **1**, 1–14 (2002)
Schirner, G., Erdogmus, D., Chowdhury, K., Padir, T.: The future of human-in-the-loop cyber-physical systems. Computer **46**, 36–45 (2013)
Sherman, W.R., Craig, A.B.: Chapter 3 - the human in the loop. In: Sherman, W.R., Craig, A.B. (eds.) Understanding Virtual Reality. The Morgan Kaufmann Series in Computer Graphics, 2nd edn, pp. 108–188. Morgan Kaufmann, Boston (2018)
Siddiqi, A., Tippenhauer, N.O., Mashima, D., Chen, B.: On practical threat scenario testing in an electric power ICS testbed. In: Proceedings of the 4th ACM Workshop on Cyber-Physical System Security, CPSS 2018, pp. 15–21. ACM, New York (2018)
Simons, S., Abé, P., Neser, S.: Learning in the AutFab – the fully automated Industrie 4.0 learning factory of the university of applied sciences Darmstadt. In: Procedia Manufacturing, 7th Conference on Learning Factories, CLF 2017, vol. 9, pp. 81–88 (2017)
Smith, A., Kumar, V., Boyd-Graber, J., Seppi, K., Findlater, L.: Closing the loop: user-centered design and evaluation of a human-in-the-loop topic modeling system. In: 23rd International Conference on Intelligent User Interfaces, IUI 2018, pp. 293–304. ACM, New York (2018)

Stouffer, K., Lightman, S., Pillitteri, V., Abrams, M., Hahn, A.: Guide to Industrial Control Systems (ICS) Security, p. 255. NIST Spec. Publ. (2014)

Tuptuk, N., Hailes, S.: Security of smart manufacturing systems. J. Manuf. Syst. **47**, 93–106 (2018)

Warrier, R.B., Devasia, S.: Data-based iterative human-in-the-loop robot-learning for output tracking. In: IFAC-Paper, 20th IFAC World Congress, vol. 50, pp. 12113–12118 (2017)

Waslo, R., Lewis, T., Hajj, R., Carton, R.: Industry 4.0 and Cybersecurity. Deloitte Insights (2017)

Weyer, S., Schmitt, M., Ohmer, M., Gorecky, D.: Towards Industry 4.0 - standardization as the crucial challenge for highly modular, multi-vendor production systems. In: IFAC-Paper, 15th IFAC Symposium on Information Control Problems in Manufacturing, vol. 48, pp. 579–584 (2015)

Smart Software and Hardware Solutions for Semiconductor Manufacturing

A Novel Software Architecture for Mixed Criticality Systems

Ralf Ramsauer[1], Jan Kiszka[2], and Wolfgang Mauerer[1,2(✉)]

[1] Technical University of Applied Sciences Regensburg, Regensburg, Germany
ralf.ramsauer@oth-regensburg.de,
wolfgang.mauerer@othr.de
[2] Siemens AG, Corporate Research, Munich, Germany

Abstract. The advent of multi-core CPUs in nearly all embedded markets has prompted an architectural trend towards combining safety critical and uncritical software on single hardware units. We present a novel architecture for mixed criticality systems based on Linux that allows us to consolidate critical and uncritical parts onto a single hardware unit. CPU virtualisation extensions enable strict and static partitioning of hardware by direct assignment of resources, which allows us to boot additional operating systems or bare metal applications running aside Linux. The hypervisor *Jailhouse* is at the core of the architecture and ensures that the resulting domains may serve workloads of different criticality and can not interfere in an unintended way. This retains Linux's feature-richness in uncritical parts, while frugal safety and real-time critical applications execute in isolated domains. Architectural simplicity is a central aspect of our approach and a precondition for reliable implementability and successful certification. While standard virtualisation extensions provided by current hardware seem to suffice for a straight forward implementation of our approach, there are a number of further limitations that need to be worked around. This paper discusses the arising issues, and evaluates the suitability of our approach for real-world safety and real-time critical scenarios.

Keywords: Mixed criticality · Raltime · Virtualisation · Hypervisor · Linux

1 Introduction

Software for safety-critical systems requires strict certification, and uncritical parts must not interfere with critical ones. Reliability of the software is crucial while the amount of software, measured in Lines of Code (LoC) is a limiting factor for certification processes.

Obtaining a functional safety certification for a kernel like Linux that contains millions of lines of code is obviously a challenging enterprise OSADL Project: SIL2LinuxMP (OSADL 2014), yet product vendors do not want to miss the capabilities of Linux in mixed-criticality systems. We present a novel architectural approach

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 121–128, 2020.
https://doi.org/10.1007/978-3-030-48602-0_11

that satisfies both goals, safety for critical parts and feature-richness for uncritical parts: Jailhouse[1], a Linux-based partitioning hypervisor.

Jailhouse transforms symmetric multiprocessing (SMP) systems to asymmetric multiprocessing (AMP) systems by inserting virtual barriers to the system and I/O bus. From a hardware point of view, the system bus is still shared, while software is allowed to only access resources within its scope.

Jailhouse is enabled from a standard Linux running on bare-metal hardware (cf. Fig. 1). It takes control over all hardware resources described in a system configuration file, reassigns them back to Linux and lifts Linux in the state of a virtual machine (VM). The hypervisor core of Jailhouse acts as Virtual Machine Monitor (VMM). Jailhouse does not fit into the usual classification of hypervisors Formal requirements for virtualizable third generation architectures (Goldberg 1973), it can be seen as a mixture of Type-1 and Type-2 hypervisors: It is a bare-metal hypervisor that runs on raw hardware without an underlying system level, but requires Linux to initialise hardware before it takes global control over the whole system.

Unlike other real-time partitioning approaches like XtratuM Partitioned Embedded Architecture based on Hypervisor: The XtratuM approach (Crespo et al. 2010) or PikeOS Evolution of the PikeOS microkernel (Kaiser and Wagner 2007) that aim to manage hardware resources and hence forbid direct access, Jailhouse only supports direct hardware access. Instead of using complex and time-consuming (para-)virtualisation Xen and the Art of Virtualization schemes (Braham et al. 2003) for emulation of device drivers, Jailhouse uses virtualisation extensions only for isolation purposes and does neither provide a scheduler nor virtual CPUs. It is a signalbox for direct routing of hardware devices to isolated domains, called »cells«. Only resources that are essential for a hardware platform and that cannot be partitioned in hardware are virtualised.

For creating new isolated domains, Jailhouse removes hardware resources[2] from Linux (also called the root cell) and reassigns them to isolated domains, called non-root cells. Virtualisation extensions ARM Architecture Reference ManualSecure Virtual Machine Architecture Reference ManualIntel virtualization technology (ARM 2013; Uhlig et al. 2005; AMD 2005) guarantee strict isolation: any access violation, for instance prohibited access to certain memory areas, wake up (trap Formal requirements for virtualizable third generation architectures) the hypervisor (Popek and Goldberg 1974), which eventually stops execution. Certain instruction executed by the guest cause traps and must be handled by the hypervisor.

Since Jailhouse remaps and reassigns resources, the hypervisor will not get active after setting up and starting all cells under ideal conditions. The following circumstances require hypervisor intervention:

- Cell management (e.g., create, start, stop or destroy cells)
- Access violations (memory, I/O ports)
- Interception of non hardware virtualisable resources (e.g., parts of the ARM Generic Interrupt Controller)
- Trapping on certain CPU instructions (e.g., x86 *cpuid*)

[1] Available at https://github.com/siemens/jailhouseunderGPLv2.

[2] E.g. CPU(s), memory, (PCI) devices, …

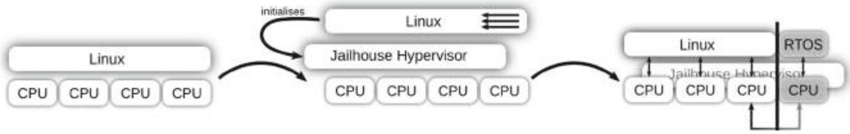

Fig. 1. Activation sequence of the Jailhouse hypervisor. After Linux has placed and started the hypervisor, an additional real-time operating system is started in an isolated critical domain.

On common bare-metal hypervisors, interrupts are dispatched by the hypervisor and reinjected into the guest. On Intel x86, we make use of Interrupt Remapping support and directly map hardware interrupts to cells without trapping the hypervisor: interrupts arrive directly in the assigned cell. This results in lower interrupt arrival times and interrupt latencies, which is beneficial for appliances with hard real-time requirements.

In this way, a safety-certified (minimalist) operating system or bare-metal application can run on a single multi-core system in parallel to Linux. The minimalist approach of Jailhouse results in only a few thousands lines of code for the core parts, which simplifies any certification process.

The rest of this paper is structured as follows: First, we present the hardware partitioning techniques of Jailhouse. We implement a multicopter demonstration platform and run critical parts software (the flight stack) in a jailhouse cell. We give architectural overview of our platform and brief introduction to our hardware setup. Afterwards, present obstacles that appeared during the implementation on real hardware, and present possible solutions.

2 Architecture

To activate the hypervisor (cf. Fig. 2), Linux must be booted with a predefined amount of reserved memory for the hypervisor and for additional non-root cells. After loading the hypervisor binary to its destination inside this memory area, the hypervisor startup code is entered by each CPU and the VMM is initialised.

After the hypervisor is initialised, non-root cells can be created. A non-root cell consists at least of one CPU and a certain amount of memory that can be preloaded by the root cell with a secondary operating system. Linux shuts down selected CPUs and calls the Hypervisor to create a new cell by providing a cell configuration. The VMM creates this new isolated domain by removing resources from the root cell and reassigning them to the newly created domain. Other resources like PCI devices, memory-mapped devices or I/O ports, can be exclusively assigned to a cell. After the cell has been started, it can reject any further tries on modifying its state. This ensures inadvertent modifications of critical domains.

Generally, Jailhouse allows guests to share physical pages with the root cell. Besides enabling inter cell communication, the mechanism also allows for sharing memory-mapped I/O pages, which, if desired, allows us to access hardware resources

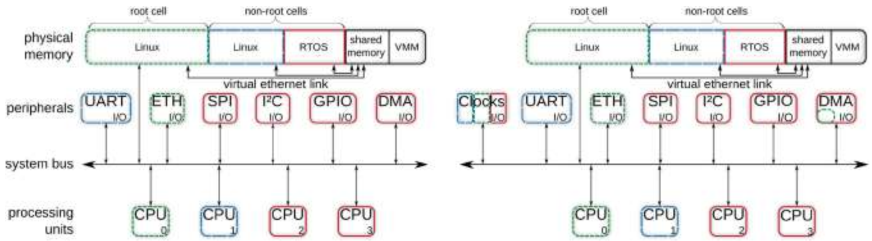

Fig. 2. Ideal vs. real hardware partitioning: Under ideal conditions (left), devices can exclusively be mapped to a cell. In reality (right), functionalities of some peripheral devices may be required in multiple domains or overlap. While the system bus is still shared, the Jailhouse hypervisor takes care that cells will only access resources within their scope. Safe communication between critical and uncritical domains is enabled by shared memory.

from within multiple domains. Such concurrent access is, however, not arbitrated by Jailhouse and needs to be addressed appropriately by the guests[3].

Figure 2 shows a possible partitioned system layout for three cells: the Linux root-cell, an additional Linux non-root cell and a bare-metal real-time operating system. As mentioned before, communication between cells is realised by memory regions that are shared between two cells, together with a signaling interface. This ensures a minimal code footprint. Jailhouse does not emulate any driver functionality, but device drivers may, for instance, use these means to establish a virtual high-performance ethernet connection between two cells. Depending on hardware support, signaling is implemented based on a virtual PCI device through Message-Signaled Interrupts (MSI-X) or legacy interrupts. On systems without PCI support, Jailhouse emulates a generic and simple PCI host controller.

Jailhouse currently supports 64-bit x86 (Intel and AMD), ARMv7, and ARMv8 architectures. Several operating systems were already successfully ported to run as Jailhouse guests. Linux can act as a Jailhouse guest on all supported architectures, the root file system is provided in memory as initial ramdisk. Let us remark that we successfully ported the RTEMS real-time operating system for the ARM architecture with limited efforts. Additionally, a port of FreeRTOS already exists[4].

3 Jailhouse Multicopter Platform

To prove the suitability of Jailhouse for industrial use cases, we implemented a fully functioning multicopter platform for demonstration purposes. We chose this platform as its requirements are similar to industrial appliances as they arise, for instance, in semiconductor manufacturing or when collaborative tasks between machines and humans need to be performed: The flight stack, a highly reliable and safety-critical part of the system, is responsible for balancing and navigating the aircraft. Sensor values

[3] This technique is mainly used for debugging purposes.

[4] https://github.com/siemens/freertos-cell.

must be sampled at high data rates, processed, and eventually be used to control rotors. The control loop is governed by different flight modes, such as a manual mode, stabilised mode or automatic modes like position hold. For a safe and reliable mission, the control loop must respond deterministically. System crashes may result in real crashes with severe consequences.

This obviously requires a real-time capable operating system. We ported the whole critical flight stack to a Jailhouse cell, while uncritical tasks still benefit from the Linux ecosystem and will not interfere with the flight stack in an unacceptable way. Remaining cells can serve any uncritical payload, such as communication with the ground station or camera tracking.

In the critical domain, a second tailored and minimalist Linux operating system with the PREEMPT_RT Internals of the RT Patch real-time kernel extension is executed. As flight stack, we chose the Ardupilot project. No modifications (besides board support and missing hardware drivers) are required. This underlines that existing applications can be deployed in a Jailhouse setup with little effort.

For controlling a multicopter platform, several sensors and actuators are connected to different inter-board buses and peripherals: gyroscopes, compasses, GPS, RC-control receiver and motor control form the controlling circuit. This requires access to SPI, I^2C, UART and GPIO hardware devices from the critical cell. A simplified architectural overview of the partitioned system is shown in Fig. 2.

As hardware platform, we chose an NVIDIA Jetson TK1[5] with a quad-core Cortex-A15 ARMv7 CPU with virtualisation extensions. The TK1 is connected to an Emlid Navio2[6] sensor shield. The system is divided into two parts: two cores are assigned to the uncritical part, the other two to the critical one.

We remove resources that are required for controlling the platform from the root cell and reassign them to the critical domain. The flight stack always controls the machine, even if uncritical cells misbehave. A crash in an uncritical cell does not cause a crash of the critical appliance. The functioning of this architecture is a solid testament to the suitability of Jailhouse for implementing real-time safety-critical systems that are based to a large extent on existing components.

4 Requirements on Partitioning Hardware

Despite the real-world practicability of our approach, we discovered limitations that are caused by hardware design. While every of these limitations can be worked around in software, the issues should be addressed by hardware manufacturers in future to provide optimal base components for mixed-criticality systems. Every workaround results in extra functionality in the hypervisor code, which contravenes the original goal of a most reduced minimal footprint, and also leads to slower response times. Such interception are, of course, contrary to the envisioned partitioning concept.

[5] http://elinux.org/Jetson_TK1.

[6] https://docs.emlid.com/navio2/.

4.1 Memory-Mapped I/O

Peripheral devices are usually accessed by reading from or writing to dedicated physical memory addresses. Those addresses are backed by the registers of the particular device. The typical page size of almost all modern architectures is 4 KiB or more, and represents the finest granularity of memory that can be assigned to a cell without the need for trapping and dispatching access.

While 32 or more bits for physical addresses provide enough space to place different devices on different pages, hardware manufacturers often place multiple devices on one single page, even different types of devices.

This is problematic for hardware partitioning, since only entire memory pages can be assigned to a cell without the need for trapping and dispatching on memory access. Jailhouse implements subpaging, a technique where the hypervisor allows for mapping memory areas to guests that are smaller than the page size. When subpaging is enabled for a certain memory area, Jailhouse will trap on any access to that page and either permit access or crash the cell because of access violations. This leads to noticeable and undesired slow-downs.

4.2 Indivisible Hardware Resources

Placing different devices on different physical memory pages is not always sufficient for hardware partitioning: functionality of a single device might be needed in two cells. Typical devices that are required in multiple cells are DMA controllers, system clock and reset controllers, or GPIO devices. Jailhouse supports sharing of physical memory pages, but it does neither moderate access nor understand the underlying hardware access protocol: Jailhouse will not ensure that parameters are not overwritten by other cells.

Most devices provide full functionality without DMA transfers. In real-time contexts, where I/O response time matters more than I/O throughput, DMA controllers should either be exclusively assigned to a single cell or should not be used if possible. Shared DMA access from different cells requires partitionable DMA controllers.

GPIO devices should exclusively be assigned to a single cell as well. As long as they are not partitionable, accesses have to be dispatched by the hypervisor.

Clock and Reset controllers allow for gating and ungating of device clocks, to select a particular clock source, and to select a prescaler for the clock. They also allow for setting and clearing reset lines of devices. Such clock and reset controllers are usually organised as a single hardware device that controls all available devices of a system. An uncritical cell that has access to the clock and reset controller can therefore deactivate or reset resources that are assigned to a critical cell, and influence the behaviour of the whole system.

One software based solution is to gate and initialise all devices, and then prohibit any further access to the clock and reset controller. While this solution would actually be straight forward, many existing drivers make the assumption that a clock and reset controller is always present and (de-)assert resets during runtime. Other device drivers, like SPI, UART or I^2C driver need to change their speed or baud rate during runtime, which requires them to access the clock and reset controller as well.

As long as clock and reset registers of all devices are bound to a single common clock and reset controller device, it is not possible to partition them without paravirtualisation or dispatching in the hypervisor. This solution is efficient for practical purposes, since clock and reset controllers are usually accessed very rarely compared to regular accesses of a device. The disadvantage is the variety of clock and reset controllers and their different protocols.

Even if shared access to clock and reset controllers is admissible, existing clock driver code is usually not prepared to run on partitioned hardware: available resources are often hard encoded in driver code, and clock drivers often reset or disable all existing system clocks during startup.

4.3 Erroneous Hardware Behaviour

Hardware misbehaves. During the implementation of our demonstration platform, we observed that accessing registers of devices with ungated clocks causes an immediate freeze of the whole system. This misbehaviour occurs in all Tegra-based platforms up to tegra186[7] and is caused by flaws in the hardware design. This problem can be fixed in software by trapping on affected memory areas when their clock gets ungated to guarantee the stability of the rest of the system.

5 Conclusion

Partitioning hypervisor techniques are promising and can be used in mixed-criticality scenarios. By using standard operating systems, we minimised the effort that is required for porting existing legacy payload applications. A minimalist hypervisor core simplifies certification efforts.

We successfully demonstrated the usability of hardware partitioning. However, hardware manufacturers need to change design aspects with respect to the demand that the hardware can be partitioned. Any software-based workarounds lead to more preventable hypervisor code and more hypervisor logic.

This demand requires software engineers and hardware manufacturers to strengthen their focus on Hardware-Software Co-design, in particular when it comes to building mixed-criticality systems that will gain increasing importance in many manufacturing domains.

References

OSADL: Open Source Automation Development Lab, OSADL Project: SIL2LinuxMP (2014)

Popek, G.J., Goldberg, R.P.: Formal requirements for virtualizable third generation architectures. Harvard University, Cambridge (1974)

Crespo, A., Ripoll, I., Masmano, M.: Partitioned embedded architecture based on hypervisor: the XtratuM approach (2010)

[7] http://www.mail-archive.com/jailhouse-dev@googlegroups.com/msg01522.html.

Kaiser, R., Wagner, S.: Evolution of the PikeOS microkernel (2007)
Barham, P., Dragovic, B., Fraser, K. et al.: Xen and the art of virtualization (2003)
ARM: ARM Architecture Reference Manual (2013)
AMD: Secure virtual machine architecture reference manual (2005)
Uhlig, R., Neiger, G., Rodgers, D., et al.: Intel virtualization technology (2005)
Rostedt, S., Hart, D.V.: Internals of the RT patch (2007)

Smart Platform for Rapid Prototyping: A First Solution Approach to Improve Time-to-Market and Process Control in Low-Volume Device Fabrication

Martin Schellenberger(✉), Sabrina Anger, Markus Pfeffer, Volker Häublein, Georg Roeder, and Anton Bauer

Fraunhofer Institute of Integrated Systems and Device Technology IISB, Erlangen, Germany
martin.schellenberger@iisb.fraunhofer.de

Abstract. Fraunhofer IISB offers prototyping services for electron devices. Such prototyping builds on proven processes and designs, amending them with customer-specific specifications and additions. The challenge is to efficiently combine existing expertise and proven process/device modules with "on the fly" R&D results and novel findings. In this paper, we present a first solution approach for a smart platform for rapid prototyping in order to improve time-to-market in low-volume device fabrication. This solution covers research of a digital twin for optimized management of wafer containers, development of intelligent algorithms for process and equipment control as well as a novel approach for "smart experiments" to achieve an accelerated transfer of process development to ISO-certified production.

Keywords: Digital twin · Data analytics · Virtual factory · Smart experiments · Real Time Control and Planning 4.0

1 Motivation

The Fraunhofer IISB runs the π-Fab (see Fig. 1), which comprises a continuous silicon CMOS process line in an industry-compatible environment, completed with specific equipment for processing of silicon carbide. Here, prototyping services for electron devices (i.e., power devices, CMOS devices, passives, sensors and MEMS) and processes are offered and performed.

The main challenge for such a prototyping line is to address a multitude of customer requirements in a flexible manner, e.g., with regard to material, device layout, or functionality. While most custom designs build and rely on a proven and optimized set of process steps, it is the mere nature of "prototyping" to replace or augment existing process steps with novel or modified ones, which are tailored to the respective customer's needs. This requires the smart combination of a certified library of proven process or device modules with results from "on the fly" research and development. As any low-volume production site with high product diversity, the π-Fab faces the challenge of an ever-changing product-specific knowledge base. Maintaining such a

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 129–135, 2020.
https://doi.org/10.1007/978-3-030-48602-0_12

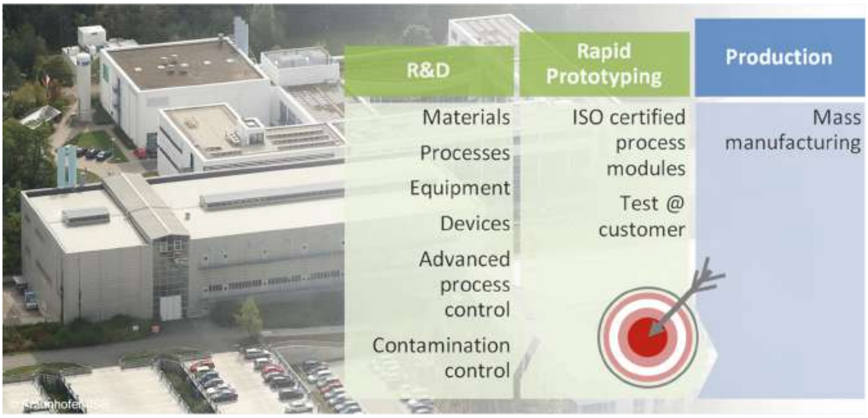

Fig. 1. Fraunhofer IISB's π-Fab covers R&D and ISO-certified rapid prototyping of semiconductor devices. Customers may transfer results from rapid prototyping to mass manufacturing at other sites.

flexible environment requires effective learning and implementation strategies in order to achieve highest quality and repeatability for the customers. Such a smart, effective and flexible platform for prototype device fabrication has to consider knowledge management, smart experiments (Davenport 2009), digital twins, smart logistics (Sturm 2006) and smart organization.

2 New Concepts for Manufacturing Control and Fast Process Transfer

To achieve substantial progress in these areas, Fraunhofer IISB carries out focused research and development within the European project iDev4.0 ("integrated Development 4.0") to evolve its π-Fab towards a smart platform for rapid prototyping. The research approach is to elaborate, and in the long run to implement, a digital duplication of the π-Fab (see Fig. 2).

The concept of using such a digital copy of the physical system to perform real-time optimization is often referred to as a "digital twin" (Söderberg 2017): The digital twin can be used to collect and store data and information from the real world, learn from it and turn the results into knowledge and actions for the real world again.

The first step towards the "π-Fab digital twin" will focus on smart experiments for accelerated process transfer and flexible concepts for manufacturing control, tailored to a rapid prototyping environment. Respective results will be used to augment existing π-Fab elements in a smart manner to enable new perspectives for the prototype fabrication of small lot-sizes with high product diversity on different substrate materials.

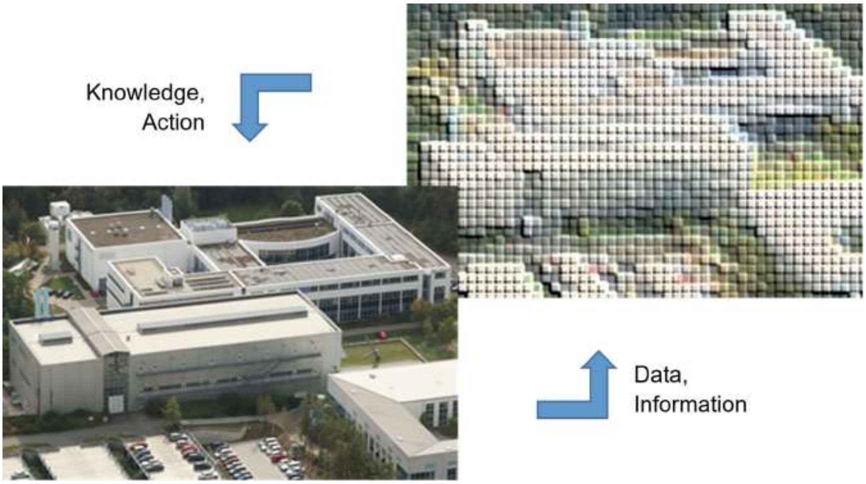

Fig. 2. The digital twin of Fraunhofer IISB's π-Fab collects data, extracts information from it and turns the obtained information into knowledge and actions.

Thus, R&D effort of Fraunhofer IISB and collaboration with project partners in iDev4.0 focuses on three areas:

1. Virtual Factory Cluster:
   - Research of a local digital twin for an optimized management of wafer containers in a combined development and manufacturing environment. This digital twin will later become a part of the overarching "π-Fab digital twin".
   - Development of intelligent algorithms for process and equipment control in a rapid prototyping environment.
2. Smart Experiments: Research of a novel approach for "Smart Experiments" to achieve an accelerated transfer from process development to ISO-certified production.
3. Real Time Control and Planning 4.0: Elaboration of flexible concepts for manufacturing control, tailored to a rapid prototyping environment.

Respective results will be used to augment existing π-Fab elements in a smart manner to enable new perspectives for the prototype fabrication of small lot-sizes with high product diversity on different substrate materials (see Fig. 3).

WORK PLAN	1. The Virtual Factory Cluster	2. Smart Experiments	3. Real Time Control and Planning 4.0
	Data-driven strategies for optimized wafer container management **APC-strategies per process in rapid prototyping**	**Smart experiments in the context of rapid device prototyping**	**Smart logistics and flexible manufacturing control for rapid prototyping**
ACHIEVEMENTS	• Data and corresponding data-bases for statistical analysis of wafer contamination identified • First investigations started based on data sets of three different data bases provided by Infineon Dresden	• Most promising application identified: Etch equipment • First data sets available • First concept for "smart experiments" developed	• Functionalities and require-ments for smart logistics and flexible manufacturing control concepts (e. g. adaptive process flow planning/ execution based on real-time experiment results) defined
CURRENT ACTIVITIES & NEXT STEPS	• Modeling of wafer container contamination behavior depending on lot process flow and transportation routes • Starting actual APC-strategy development	• Data analysis • Refinement of "smart experiments" concept • Approach alignment with project partners	• MES specification and evaluation considering specified functionalities and requirements

Fig. 3. First solution approach to evolve the π-Fab towards a smart platform for rapid prototyping

3 Results

Within the first period of the iDev4.0 project, several steps have been taken to realize the above described solution approach. Specific focus was on smart logistics and contamination control (area 1 and 3) and smart experiments (area 2).

3.1 Smart Experiments

In the flexible prototyping environment of the π-Fab with a wide mix of products, low product volumes as well as both R&D and production evolving in parallel in the same manufacturing line, an efficient use of any available bit of information and data is essential to quickly transfer latest R&D results into certified prototyping. To achieve this objective, we started research on the novel concept of "smart experiments". So far, two important steps were taken:

1. A group of experts from the equipment, process and data analytics domains identified in a joint effort the most challenging processes that require accelerated process transfer. As a result, the focus will be on an etch cluster tool, comprising three chambers dedicated to specific etching steps.
2. In parallel, a concept for smart experiments was developed. Figure 4 summarizes the initial concept, while the alignment of the approach with the other project partners is ongoing.

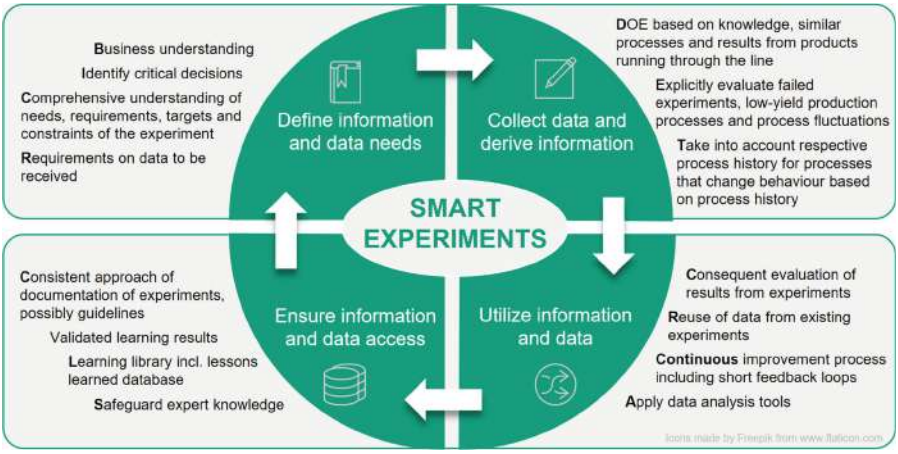

Fig. 4. The concept of "smart experiments".

By applying this approach, a quick identification of a high quality and comprehensive set of experiments is aimed for, resulting in a decrease of design and learning cycle time. The first concept evaluation and validation approach considering dry etch processes will show, whether the aspired 20% reduction of integration-time for newly developed processes becomes possible.

3.2 Smart Logistics and Flexible Manufacturing Control for Rapid Prototyping

For high quality and reliable products, it is common in industrial high volume production to use only precisely defined processes and to monitor the processes closely by means of various software systems, such as a manufacturing execution system (MES). This is a distinct difference to Fraunhofer IISB's prototype fabrication of electron devices, where a multitude of customer requirements with regard to material, device layout, or functionality is being addressed in the process pool. This flexibility is a matter of principle and leads to continuous variations of process steps and process flows in the prototype fabrication. In order to optimize the entire value chain, from order to customer and back again for process/device improvements, flexible manufacturing control techniques have to be developed and applied.

Fraunhofer IISB is working within iDev4.0 on smart logistics and flexible manufacturing control concepts for rapid prototyping. Functionalities and requirements for adaptive process flow planning, process execution and process parameter adjustment based on results of real-time and smart experiments are currently being defined. The developed concepts will be later implemented in the new manufacturing execution system, which will be rolled out as part of the "Research Fab Microelectronics Germany/Forschungsfabrik Mikroelektronik Deutschland (FMD)".

Another important challenge in manufacturing and prototyping, respectively, is the implementation of smart logistics for wafer carriers. Based on the π-Fab's flexibility, various wafer sizes and types are being handled in different contamination protocol zones and the customer may determine the points of entry and exit in the process line. To overcome this challenge, innovative carrier monitoring strategies are currently developed by Fraunhofer IISB utilizing the concept of digital twins. During the project, a prototype for the data-driven contamination monitoring of FOUPs in the production line at Infineon will be developed. All the techniques learned reaching from the analysis equipment to the different contamination sources in a power semiconductor fab to data management of carrier data will be taken into account when implementing an optimized wafer container monitoring in the Pi-Fab. This could be for example the introduction of carrier changes or adding time coupling after certain steps.

Figure 5 shows the principle of the intended concept for a digital twin of a wafer container. The tracking information of the wafer and the corresponding wafer container will be combined with the analysis of data from in-line and off-line contamination control procedures to improve carrier logistics with regard to contamination. First investigations have been started with data sets provided by an industrial project partner.

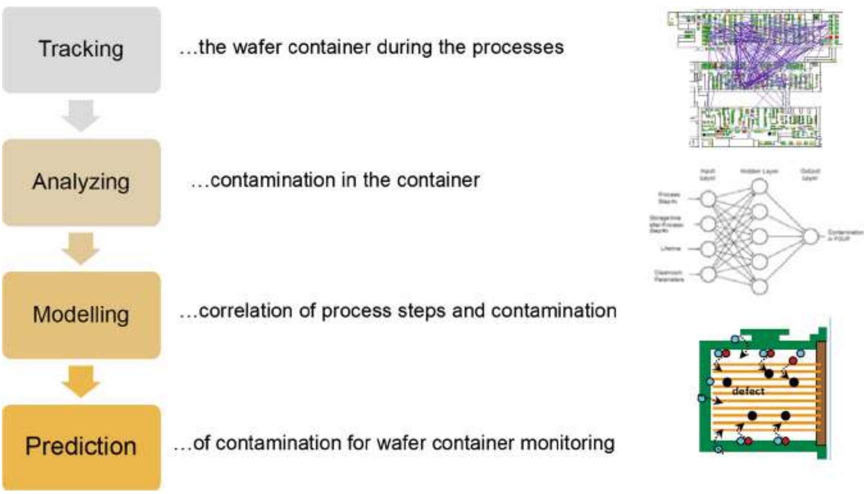

Fig. 5. Digital twin principle for wafer container

4 Summary

A first solution approach towards a smart platform for rapid prototyping of low-volume devices was outlined and actual implementation started. The approach comprises R&D in smart experiments, digital twins, smart logistics and smart organization. The actual development of solutions was preceded by a survey. This procedure is in line with the proven approach described, e.g., by the CRISP-DM model (Shearer 2000): This step is focused on knowledge and understanding of the objective on the one hand, and the alignment of expectations of the stakeholders on the other. Similarly, McKinsey describes an "advanced analytics" approach, which they suggest to apply in semiconductor industry in general and R&D in particular. In contrast to traditional analytics, they also recommend to start from the business problem that needs to be solved and to head for information and respective data needed to derive the correct decisions in respect thereof (Batra 2016).

Project Information: The project "integrated Development 4.0" leads the digital transformation of singular processes towards an integrated digital value chain based on the "digital twin" concept. Development, planning and manufacturing will benefit from the "digital twin" concept in terms of highly digitized virtual processes along the whole product lifecycle.

The project iDev4.0 has received funding from the ECSEL Joint Undertaking under grant agreement No 783163. The JU receives support from the European Union's Horizon 2020 research and innovation program. It is co-funded by the consortium members, grants from Austria, Germany, Belgium, Italy, Spain and Romania.

References

Batra, G.: Improving the semiconductor industry through advanced analytics. McKinsey & Company. Website, March 2016 (2016). https://www.mckinsey.com/industries/semiconductors/our-insights/improving-the-semiconductor-industry-through-advanced-analytics. Accessed 05 Oct 2018

Davenport, T.H.: How to design smart business experiments. Strateg. Dir. **25**(8) (2009). https://doi.org/10.1108/sd.2009.05625had.004

Shearer, C.: The CRISP-DM model: the new blueprint for data mining. J. Data Warehous. **5**(4), 13–32 (2000)

Söderberg, R.: Toward a digital twin for real-time geometry assurance in individualized production. CIRP Ann. **66**(1), 137–140 (2017)

Sturm, R.: Model-based approach for online performance evaluation of automated transportation systems in semiconductor manufacturing. Universität Stuttgart (2006). https://doi.org/10.18419/opus-4090

PCMC DC-DC Converter Development Methodology by Means of dSPACE

Osvaldo Gasparri[1](✉), Roberto Di Lorenzo[1], Paolo Del Croce[2], and Andrea Baschirotto[1]

[1] University of Milan Bicocca, Milan, Italy
o.gasparri@campus.unimib.it
[2] Infineon Technologies Austria AG, Villach, Austria

Abstract. This work describes the development methodology of an electronic device. All the phases, starting from the definition of the specifications up to the fabrication of the device, will be explained using the evolution of a project aimed at developing a high-performance DC-DC converter for LED driver application as a use case. The project involves teams and experts from the technology department, the product development and the fabrication. It is therefore important to have an efficient process flow, ensuring the optimal communications and information transmission between the teams. In addition, the project here presented introduces a novelty regarding the concept phase. The use of a rapid control prototyping tool, called dSPACE, guarantees the feasibility of a concept circuit before proceeding directly with the design phase, avoiding time-wasting. In fact, the mentioned emulator allows the preliminary evaluation of the circuit in ideal conditions and, subsequently, a more real study on hardware.

Keywords: Methodology · DC-DC Converter · dSpace · Simulink · Concept

1 Motivation

The aim of this work is to make the development process of an industrial microelectronic device that starts from the definition of the specifications and ends with the product fabrication, more efficient. To do this, it is important to operate within a well-defined workflow between the different teams involved, each of which will mainly work on a specific phase of the project. The clear definition of the process phases, together with the clear transmission of the project information/results between all the involved teams, by means of digital twins, guarantees the continuity of the chain, improving efficiency and minimizing troubles. Each step involves the use of dedicated software that will return a digital file in a special format, subsequently used as input for the next step. The result is a digital chain in which the composition rings and junctions are already defined from the beginning.

Each process step is explained here using the evolution of a Buck DC-DC converter as a use case. Since, for the specific microelectronic case, circuit design is usually the longest and most problematic phase, it is convenient to exploit a concept phase where the architecture of the entire circuit is identified and validated in advance. This is what

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 136–144, 2020.
https://doi.org/10.1007/978-3-030-48602-0_13

the environment dSPACE does through the mixed implementation between software (Simulink) and hardware (a dedicated board), returning a block scheme of the prototype circuit to be designed.

The developed methodology intends to be valid for any other device to be realized, improving production efficiency.

2 Methodology Description

The chapter explains the digital phases of the development process one by one. Figure 1 shows a flowchart of the development process leading to the production of a generic circuit. The process has a well-defined chronological sequence. Each step involves experts of specific teams, working on specific software. Furthermore, each block of the chain is connected to both the previous and the next. This connection (arrows, in Fig. 1) could be a simple passage of information, such as a file containing the dataset, or it could be a very complicated file in a special format containing the circuit design/layout. These files are the single source of truth for the teams concerned (also spelled out in Fig. 1). In fact, in each new step, the starting point is the file returned from the previous phase, regardless of what happened before or how. Digital twins are used to efficiently switch from one step to another without discontinuity or troubles.

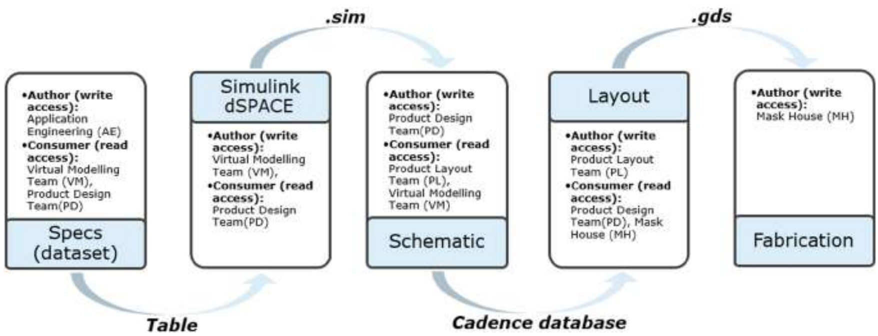

Fig. 1. Development methodology

Once the workflow is clear, it is necessary to define the project and its application. In the automotive industry any electronic device works with the car battery, whose voltages range from 7–8 V to >27 V due to cold-cranking and load dump. However, to avoid damages to such devices, their supply source should be stable and consistent with their datasheets. A Buck DC-DC Converter could be used to reduce the value of a voltage/current supplied as an input. In addition, a control circuit is used to make its output equal to the desired value. The controlled DC-DC Converter is then placed between the battery and the electronic device to ensure correct operation. In the case considered, the device is the LED for street lighting.

Every project starts with the definition of the specifications: this happens e.g. in a meeting with all the experts from the teams involved. The list of specifications, or dataset, contains the key features of the circuit to be designed. These can be circuit technical specifications, such as the value of the DC-DC output current, or costumer requirements such as economic or temporal aspects. Note that the product price influences the technical specifications. Since the silicon prize is 0.01–0.02 €/mm^2 (depending on the technology), the area of the chip must be minimized. Consequently, the most complex circuit topologies could already be discarded because they would be too bulky. In the meantime, it is important to guarantee high efficiency of the circuit: the best solution will be chosen in terms of efficiency/price ratio.

The specification list will then be placed in a database or in a read-only file stored on a corporate server. Only employees within the project have the right of access to open the file and read the dataset at one time. In the event that a specification changes (e.g. if the concept circuit test through dSPACE shows a discrepancy between a specification and a technology limit), the data will be modified or updated. So all the colleagues will only see the updated data. This avoids problems related to bad communications between teams. Table 1 shows the DC-DC Converter dataset. The parameters are initially listed and are no longer changed. Only their values can change.

Table 1. List of specifications (dataset)

Parameter	Measure unit	Value
Switching frequency	kHz	<400
Input voltage range	V	[7,26]
Inductance	μH	<40
LED peak current	A	<3.5
LED ripple current	A	[2.6,3.4]
LED mean current	A	3 ± 5%

This first phase returns a string of numbers/ranges obtainable from the Table, namely:

$$\{400, [7, 26], 40, 3.5, [2.6, 3.4], [2.85, 3.15]\}$$

This string will be the input for the concept phase of the main-block-circuit. Note that the product price is not directly part of that dataset. However, some specifications may indirectly bring this information: for example, the inductance value is limited to 40 μH. A higher value would result in a more expensive coil and a higher area occupation. On the other hand, a higher inductance value would lead to a more accurate led driver: in fact the inductance is inversely proportional to the current ripple (Everett 1999; Zainal 2003). The lower the ripple, the closer the output current to the mean value desired by the LED, allowing it to operate safely. Finally, technology also influences the specifications: the higher the switching frequency, the higher the power consumption but, at the same time, the higher the accuracy. A compromise must be found.

The concept phase, aimed at finding the optimal solution, begins now. First of all, researchers and product development experts study the state of the art for a first analysis of all possible solutions. At the same time, the rules that describe the circuit operation must be deduced. For the reader, the DC-DC Converter can be represented by a black box (consisting of a transistor, an inductor etc.). (Mike 2011) provides the proper schematic. Its role is simple: whatever changes from the output, the circuit must guarantee a fixed current to the output load (the LED). This is done by turning a switch (the transistor) on and off repeatedly (at 400 kHz). In this way, the current supplies to the output rises and falls very quickly, obtaining an almost constant value (Everett 1999). However, a DC-DC Converter alone cannot be insensitive to external perturbations: it needs a "control loop". By reading the output current value, the control loop is able to determine when it is time to turn the switch on and off, making the current as close as possible to the desired value. Here is an example of how the control loop might work: reads the output current; compares the current with the peak current specification; turns the switch off if the output current exceed this value; waits 1/400 kHz before turning the switch back on. This is how the so-called "Peak Current Mode Control (PCMC) DC-DC Converter" works.

Obviously, there are many other control loop solutions (Lloyd 1985, Frank 2007). The concept phase must filter the best solutions, simulate them and compare their results. This is made with a first high-level simulation in ideal conditions of the solutions found. The high-level concept circuit is created using the MATLAB's Simulink software. Here a circuit appears like a set of blocks connected together. The blocks can be predefined (logical operators, counters, registers, etc.) or can be programmed using MATLAB to perform the desired function. It is therefore possible to simulate any type of circuit. The PCMC DC-DC Converter has the block representation of Fig. 2.

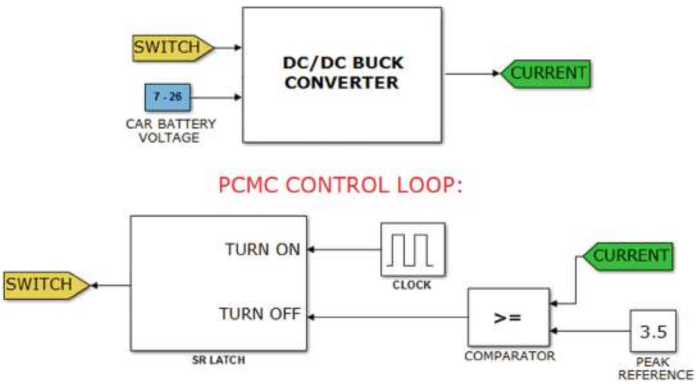

Fig. 2. Simulink block approach of the controlled DC-DC Converter

Simulation's plots and data are the feedback to understand the suitability of the solutions. If unsatisfactory, the circuit can be easily adjusted or upgraded. In fact, from the PCMC loop (Fig. 2, lower side), many other solutions can be obtained (Osvaldo 2019). For example, the clock that decides when to turn on the switch can be replaced with another comparator: if the current drops below a lower value (e.g. 2.6 A), the switch is turned on. The Hysteretic Current Mode Control (HCMC) is thus obtained which accurately controls the current ripple. While, for better control of the average current value the Average Current Mode Control (ACMC) can be obtained (Lloyd 1999).

After a first screening of the solutions and their simulations, the test on hardware is the next step: no longer in ideal condition, but this time also considering all the parasitic elements that could affect the efficiency/operation of the circuit. The rapid control prototyping tool called dSPACE comes into play (dSpace 2011). dSPACE takes the Simulink file as an input. In fact, this emulator contains an FPGA module (programmable via Simulink) on which the Simulink circuit is downloaded, after having been automatically translated and coded in VHDL. So the emulator dSPACE becomes the desired circuit to be tested. For example, it can be used to test all the DC-DC control loops once at a time. The DC-DC Converter (Fig. 2, upper side) is instead incorporated in a Printed Circuit Board (PCB). Connected to the PCB output there is the LED. In this way, the test takes into account the actual behavior of the switch and the LED. Figure 3 shows the prototype.

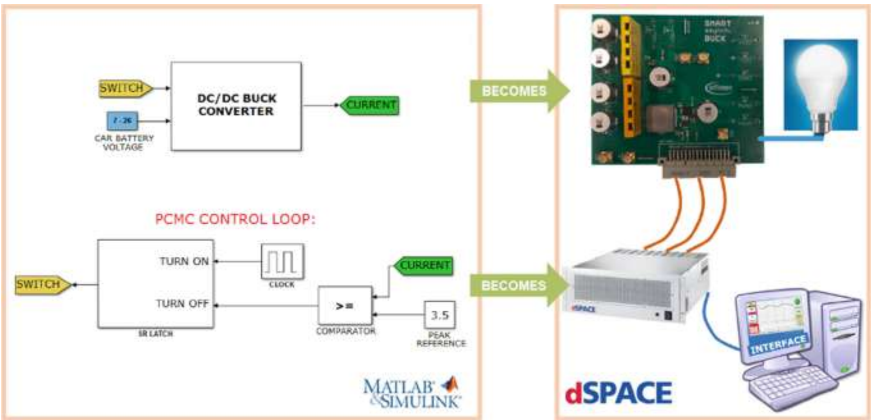

Fig. 3. Set-up. From Simulink (left) to dSPACE (right)

dSPACE contains a PowerPC that deals with little time-demanding calculations and interfacing with the rest of the modules. It also provides real time communication with the host PC. This means that, both before and during the test, the user can manage everything with a Graphic User Interface (GUI), easily created using the dedicated software Control Desk. Here, the user can directly enter specifications string returned from the first phase (from Table 1) and run the circuit testing. Figure 4 shows the scheme of the laboratory environment, relating to the DC/DC use case. The results obtained in the concept phase are then presented to the AE/PD experts, who will decide among the chosen solutions the optimal one with which to proceed to the design phase.

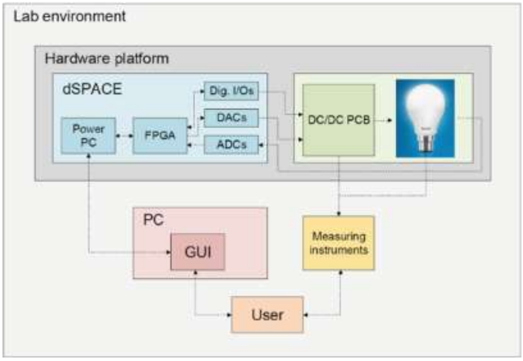

Fig. 4. Scheme of the Laboratory environment

The .sim file from the concept phase is the input for the design phase in which the Simulink blocks will be designed at transistor level. This is accomplished by creating a schematic in Cadence, software of Virtuoso. This is the most critical phase. Many difficulties can arise. Every single component is simulated by varying the process, temperature etc., to ensure the correct circuit operation even in the worst situation. Once the circuit is designed, the schematic file is given to the layouter who will draw the layout. The.gds file returned will then be the input for the fabrication phase. Figure 5 shows these latter phases. However, before obtaining the final product, the chips must be tested in laboratory to verify their operation, robustness etc. Sometimes there are errors or inaccuracies that make it necessary to modify the schematic. The process restarts from the design phase by creating a loop. It will take a long time for the chip to work properly, in line with the specifications and ready to be sold.

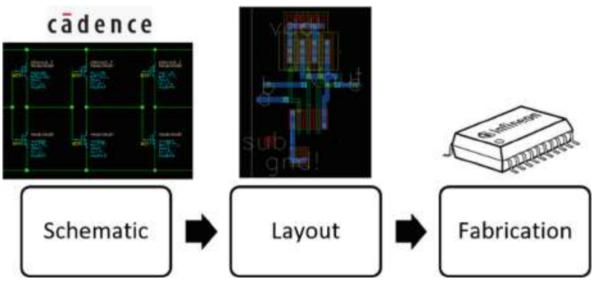

Fig. 5. Final development phases

3 Results

This section is reserved to show the results of the concept phase for the PCMC DC/DC converter LED driver, both during the Simulink simulation and the dSPACE test, using the set-up of Fig. 3. Figure 6 shows the circuit operation during a Simulink simulation.

At one point, an abrupt change in the voltage of the car battery was simulated: the output current was changed accordingly. Worst conditions simulations allow to understand the robustness of the circuit.

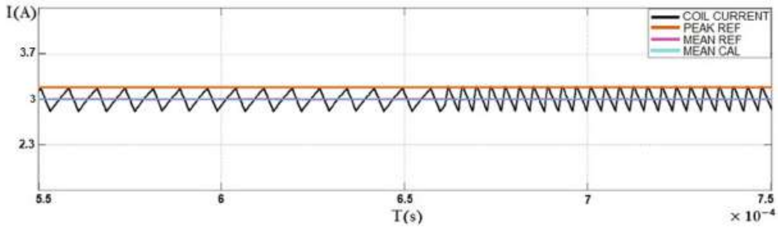

Fig. 6. Output current waveform from Simulink simulation

After a first screening of the simulated solutions, dSPACE is used to test their behavior no longer in ideal condition. As shown in Fig. 3, the basic DC-DC Converter circuit becomes a PCB, while the control loops are downloaded one by one into the FPGA of dSPACE.

Through a GUI, Fig. 7, users controls the operation of the system. In particular, some circuit specifications can be set. It is also possible to set an upper limit for voltage and current against over-current and over-voltage. Waveforms, e.g. the output current, can be directly displayed on the computer together with other parameters, such as the input voltage from the car battery. However, for better precision, the oscilloscope is used. Figure 8 shows all waveforms of interest from scope, which also provides the peak, the ripple, the mean and the switching frequency values. Note that, in this case, the output current sensing is made through a 0.5 Ω resistance. Therefore, all those values in the low bar of Fig. 8 relating to a current signal, but displayed in Volts on the scope, should be divided by 0.5 Ω, to obtain the respective in Amps. For example: I_{max} = 1,65 V/0,5 Ω = 3,3 A. The results from Simulink and dSPACE can be then compared to evaluate the impact of the parasitic elements and the presence of the LED on the response of the system.

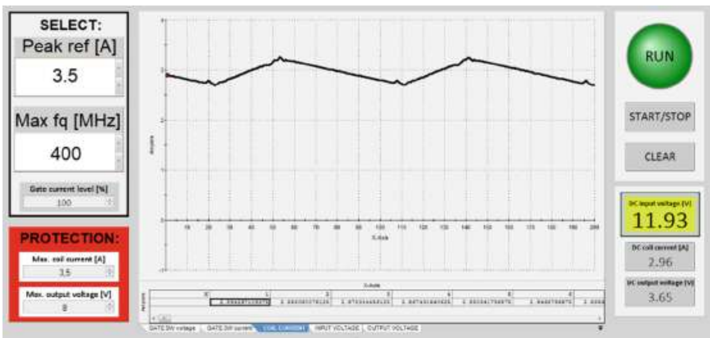

Fig. 7. Graphic User Interphase (GUI) during circuit testing

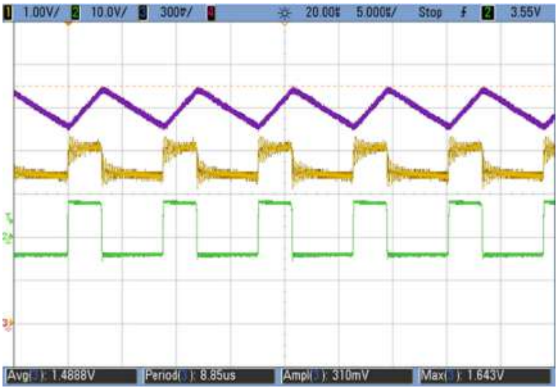

Fig. 8. From scope: current (violet); output voltage (yellow); switch activity (green)

4 Discussion/Implications

The results of the simulated and tested solutions will be presented and discussed before the experts. A table resuming the pros and cons of all the topologies will be the tip of the scale to choose the optimal solution on which to continue further development. Table 2 is the example in the case of the DC-DC Converter, in which many control loops have been tested and compared. The PCMC was preferable over the HCMC and ACMC. Subsequent improvements led to the PCMC with Slope Compensation and the final optimal concept renamed Variable off Time PCMC (Osvaldo 2019).

Table 2. Pros and cons of the screened solutions

Parameter	HCMC	ACMC	PCMC	PCMC + SC	VoT-PCMC
Loop simplicity	++	−	++	+	+
Regulation Precision	−	+	−	−	++
Noise immunity	−	++	−	−	++
Reference perturbation	−	++	+	+	+
Line perturbation	+	+	+	+	++
Rapidity	++	−	++	++	++
Sampling advantages	−	−	++	++	++
Stability	++	++	−	++	++
Improvement available	−	−	++	−	−

The VoT-PCMC Simulink file becomes the input of the delicate design phase. During Simulink's ideal simulations no problems arise (since the limiting factor is the computer hardware). While, during the dSPACE test phase, the efficiency of the circuit could be significantly reduced. In fact, the dSPACE platform has intrinsic delays that may lead to systematic errors. The source of these delays lies in the Analog to Digital Converter (ADC). This is used for instance for reading the output current (analog signal) and for translating it into a digital signal to be used later to make comparisons/turn the switch on and off. Delays could lead to wrong system states. However, once the error source is known, problems can be avoided.

References

dSpace: Hardware Reference, FPGA DS5203, PPC DS1005, datasheet (2011)

Rogers, E.: Understanding Buck Power Stages in Switchmode Power Supplies, Texas Instruments (1999)

Xi, F.: Design of DC-DC Converters, October 2007, IEEE SSCS Dallas Chapter (2007)

Dixon Jr., L.H.: Closing the feedback loop. Texas Instruments (1985)

Dixon, L.: Average Current Mode Control of Switching Power Supplies. Uni-trode (1999)

Wens, M., Steyaert, M.: Design and Implementation of Fully-Integrated Inductive DC-DC Converters in Standard CMOS, Chapter 2, 2011th edn. Springer, Heidelberg (2011)

Gasparri, O.: Variable off-Time Peak Current Mode Control (VoT-PCMC) as method for average current regulation in Buck Converter Drivers, November 2019. Genova, ICECS (2019)

Salam, Z.: Power Electronics and Drives, Chapter 3, UTM-JB (2003)

Author Index

S. Keil et al. (Eds.): EADTC 2018/2019, LNEE 670, pp. 145–146, 2020.
https://doi.org/10.1007/978-3-030-48602-0

S

V

Z