Towards an Infrastructure Enabling the Internet of Production

Jan Pennekamp*, René Glebke*, Martin Henze*, Tobias Meisen§, Christoph Quix¶, Rihan Hai‖, Lars Gleim‖, Philipp Niemietz†, Maximilian Rudack‡, Simon Knape†, Alexander Epple†, Daniel Trauth†, Uwe Vroomen‡, Thomas Bergs†, Christian Brecher†, Andreas Bührig-Polaczek‡, Matthias Jarke‖, Klaus Wehrle*
*Communication and Distributed Systems, RWTH Aachen University, Germany
{pennekamp, glebke, henze, wehrle}@comsys.rwth-aachen.de
‖Databases and Information Systems, RWTH Aachen University, Germany · {hai, gleim, jarke}@dbis.rwth-aachen.de
‡Foundry Institute, RWTH Aachen University, Germany · {m.rudack, u.vroomen, sekretariat}@gi.rwth-aachen.de
¶Fraunhofer Institute for Applied Information Technology, Germany · christoph.quix@fit.fraunhofer.de
†Machine Tools and Production Engineering, RWTH Aachen University, Germany
{p.niemietz, s.knape, a.epple, d.trauth, t.bergs, c.brecher}@wzl.rwth-aachen.de
§Technologies and Management of Digital Transformation, University of Wuppertal, Germany · meisen@uni-wuppertal.de

***Abstract*—New levels of cross-domain collaboration between manufacturing companies throughout the supply chain are anticipated to bring benefits to both suppliers and consumers of products. Enabling a fine-grained sharing and analysis of data among different stakeholders in an automated manner, such a vision of an *Internet of Production (IoP)* introduces demanding challenges to the communication, storage, and computation infrastructure in production environments. In this work, we present three example cases that would benefit from an IoP (a fine blanking line, a high pressure die casting process, and a connected job shop) and derive requirements that cannot be met by today's infrastructure. In particular, we identify three orthogonal research objectives: (i) *real-time control of tightly integrated production processes* to offer seamless low-latency analysis and execution, (ii) *storing and processing heterogeneous production data* to support scalable data stream processing and storage, and (iii) *secure privacy-aware collaboration in production* to provide a basis for secure industrial collaboration. Based on a discussion of state-of-the-art approaches for these three objectives, we create a blueprint for an infrastructure acting as an enabler for an IoP.**

***Index Terms*—Internet of Production; Cyber-Physical Systems; Data Processing; Low Latency; Secure Industrial Collaboration**

I. Introduction

Over the last decades, supply chains in manufacturing have evolved into highly complex and tightly interconnected structures encompassing several different stakeholders, ranging from suppliers of raw materials and refining industries to manufacturers of sub components, whole systems and, ultimately, the consumers. All these entities can contribute to the success of a product by *exchanging knowledge*, e.g., when consumers communicate desired and perceived properties of the product to the manufacturer or when companies within the supply chains share characteristics and specifications of their products among their respective customers and suppliers. The potential benefits of collaboration throughout the supply chains, with service providers, and between the different stakeholders are tremendous; the McKinsey Global Institute, e.g., argues that by increasing the usage of big data analytics, crowdsourcing, and advanced product life cycle management, companies could reduce product development costs and time to market by between 25 % and 50 % [1]. Gains in profit margins by 2 % to 3 % through increased digitization are similarly predicted [1], as are general improvements in product quality [2].

So far, however, the anticipated potentials have not yet been unleashed. Manufacturing companies are traditionally reluctant to share more details on their production processes than absolutely necessary in fear of accidentally disclosing trade secrets while getting only meager benefits in return [1]. Additionally, even if a company chose to open up to potential collaborations, modern production systems already create orders of magnitude more data than they produce goods [3], and the infrastructure currently present in production environments has not been designed with a sharing of data or the incorporation of foreign data into local processes in mind. Incentives and technical foundations for cross-domain collaborations are thus missing, and knowledge is predominantly retained locally.

Concepts such as the Industrial Internet of Things (IIoT) and Industry 4.0 have already targeted the collection and processing of manufacturing data to automate production processes as well as the development and deployment of autonomous control mechanisms to improve system safety and reliability (see e.g., [4] for an overview of existing approaches). Yet, the concepts proposed so far mainly aim at specific locally-constraint advances in single areas, such as food supply chains, transportation/logistics, or healthcare [5]. In contrast, to enable real cross-domain and inter-company collaboration, semantically adequate and context-aware data from production, development and usage need to be made available to (and potentially reacted to by) interested parties in real time, at a reasonable level of granularity, and at a potentially global scale. The isolation of data produced by (and available to) production machinery on the shop floor needs to be overcome: Machines and production sites all around the world should be enabled to exchange information with each other via a single global infrastructure, essentially creating an *Internet of Production*.

Author manuscript.

INTERNET OF PRODUCTION (IOP): Based on the envisioned advances, our proposed IoP poses a number of significant challenges on a potential underlying infrastructure. On the one hand, the infrastructure needs to support the control of production processes in real time: Decisions must be made as quickly as possible to account for the (diverse) individual process requirements. On the other hand, the infrastructure needs to be able to deal with vast amounts of heterogeneous production data, parts of which are processed right away, while other fractions have to be (permanently) stored for future analyses. The concept of collaborating companies furthermore necessitates secure networking and processing paradigms, an objective which has so far usually been out of scope due to the focus on locally-constrained scenarios.

The remainder of this paper is organized as follows. In Section II, we first motivate the vision underlying an IoP along three real-world use cases: (i) a *fine blanking* line producing up to 6.2 Gbit/s worth of measurement data, (ii) a horizontal cold-chamber *high pressure die casting* process with measurement data on different time scales, and (iii) a *connected job shop* with significant uncertainties of machining operations, components, and tools. From these use cases, we derive challenges and requirements that cannot be met by today's communication, storage, and processing infrastructure. In Section III, we then lay out a blueprint for an infrastructure that can act as an enabler for an IoP. To this end, we propose and discuss three orthogonal research objectives that need to be tackled to realize an IoP-supporting infrastructure. We conclude our paper and provide an overview of further research challenges in Section IV.

II. POTENTIALS OF AN INTERNET OF PRODUCTION

To showcase the envisioned significant benefits of an IoP, but also the resulting key challenges for the infrastructure which are not sufficiently addressed by the current state of research, we selected three distinct real-world use cases: a fine blanking line, a die casting process, and a connected job shop. In the following, we discuss these use cases in more detail, identify potential benefits an IoP may bring, and derive requirements for the underlying infrastructure of an IoP.

A. High-Throughput Data Processing in Fine Blanking

Our first use case, fine blanking, is a precision forming process in which large numbers of identical work pieces are cut from a coil of raw material in a press [6]. The raw materials can vary in thickness (reaching down to less than 1 mm, with potential fluctuations over the course of the coil), are influenced by environmental conditions such as heat, and are subject to physical stress such as uncoiling and punching within the press. Produced work pieces may hence vary greatly in the quality of their surfaces and edges and may even exhibit fine cracks [7], [8], resulting in some of the produced pieces to be rejected. The blanking punches used to cut the raw material within the press furthermore are cast of high-strength materials in complex processes and are thus very valuable [9]. When subjected to excessive forces (e.g., too high punching speeds or punching depths that exceed the thickness of the raw materials), the punches can easily wear off, requiring replacements with long down times of the respective production line.

Tightly integrating the various process parameters, including properties of the raw material, as well as forces, temperatures, and vibrations within the press, would thus enable an ad-hoc adaption of the process, resulting in a reduction of rejected work pieces through better models for quality prediction [10]. The longevity of the machinery of the line may at the same time be increased using advanced predictive maintenance techniques [11]. The sources of the needed parameters are not restricted to the fine blanking line; the supplier of the raw material coil, e.g., could make a digitalized production log available along with the coil, which may include data that helps configuring the fine blanking line accordingly. In turn, parameters gathered during the fine blanking process could also be used in processes further downstream, enabling a traceability of the goods and their production conditions throughout the entire supply chain.

The high process speed of fine blanking (up to 140 punches per minute in our line) and the general complexity of fine blanking lines (over 100 separately controllable components in the press alone), however, render a tight integration of data gathered during the various stages of the supply chain challenging. In our line, data exchanges occur at frequencies between 2.5 and 1000 kHz, resulting in up to 6.2 Gbit/s worth of data that need to be integrated for real-time control decisions [3], as well as potentially stored for downstream use. Considering the likely scenario of multiple such lines to be operated at the same site, such numbers will quickly lead to an overload of the network and processing infrastructure available in most factory settings, rendering a centralized processing of the data infeasible. To guarantee the needed high data exchange frequencies and data rates, finding new solutions which facilitate a (pre-)processing of data close to its origin is imperative. Furthermore, exchanging raw data between different companies can enable the involved organizations (or, in cases of data leaks, even third parties) to reverse-engineer process optimizations and, hence, uncover trade secrets. Mechanisms protecting against the undesired use of confidential data hence need to be integrated in such a way that their functionality is provided throughout the entire infrastructure realizing an IoP.

B. Global-Scale Data Analysis in Die Casting

The horizontal cold-chamber high pressure die casting (HPDC) process is representative of permanent mold based near-net-shape production technologies and serves as our second use case. The HPDC process is capable of mass producing high quality, complex geometry castings from non-ferrous alloys for mechanically demanding applications such as crash relevant parts for car body components [12]. Variances in quality and productivity are intrinsic to the process, due to complex thermal, mechanical, and chemical interdependencies. The highly individual production parts and their corresponding die designs have significantly differing requirements for the machine's optimal operating point. The shot curve, intensification pressure, spraying process, dosing temperature, vacuum levels, dwell time in the shot sleeve and in the cavity, and the overall cycle time all impact the part quality and productivity. However, their interdependencies have not yet been fully resolved in a general model since the individuality of the cavities, machines, process designs, and the production environments cannot be accounted for. A stable HPDC process can still cause scrap rates in a range of 5 %–10 %. Common defects include pre-solidified material from the shot sleeve, entrained air, flash, and sticking related defects [13].

Real time controlled horizontal cold chamber HPDC machines generate significant amounts of process data, but even for the experienced expert, finding the root cause of a defect retrospectively from the available information can be difficult. One reason is that the HPDC process generates data on different time scales. The fast shot and therefore the injection process, typically lasts anywhere between 20 ms and 200 ms, the spraying cycle takes around 10 s depending on the spraying strategy, and the overall cycle time is usually around 60 s. Identification of an instable sub-process of the overall cycle is therefore not trivial; especially if an extensive log of previous production data as a comparison base is missing. Nevertheless, first investigations using mathematically driven models have already shown the potential of analyzing data generated within HPDC production environments [14].

In general, most process-related defects in HPDC are caused by a complex combination of factors. In real production environments, eliminating certain process variables to trace the defect origin is often infeasible. Therefore, extracting all machine and auxiliary system signals is desirable to make them accessible for adequate models within an IoP.

For production use, several issues still need to be addressed regarding the data transfer, storage within an appropriate infrastructure, and access for external entities while ensuring stakeholder confidentiality. The process data provided by the HPDC machine and its auxiliary systems need to be prepared in a coherent, structured form with appropriate sampling rates for every sub-system of the HPDC cell. The HPDC machine, its thermal regulation units, the spraying system and the dosing furnace should ideally comply with commonly deployed standards, such as the OPC-UA [15], to enable timely target actual comparisons, appropriate subsequent data storage, and analysis by suitable models within an IoP.

C. Real-Time Connected Job Shop

Machining operations are key technologies in most industries to produce goods for the global market. Due to the high variety of machine tool designs even very complex workpiece geometries can be manufactured. With increasing complexity of the workpiece, choosing the right process parameters for an optimal, fast, and reliable process becomes more challenging. Even with ideal parameters, all machining processes are subject to continuous changes following, e.g., the wear of machine tool components and tools. Therefore, uncertainties in terms of process stability and quality occur, resulting in an increased lead time or even in defect workpieces. In our third use case, data from the shop floor could help, e.g., to generate real-time production insights, adjust processes in real-time, optimize existing processes automatically, and predict failures of the process or the machines before they occur.

Machine tools provide a constant stream of internal sensor data and process commands. Recording this information and combining it with other sources, for example the ERP, MES, or quality system, enables the creation of a digital snapshot of the production. By applying model-based analytics, a digital twin of the workpiece and the machine tool can be generated. In previous research, a process parallel material removal application and a machine health monitor have been created [16]. The material removal application calculates process force estimations parallel to the process and models the machine behavior wrt. tool deflection and geometrical axis errors. From this information, it generates a digital version of the manufactured workpiece, which is then virtually measured to assess the quality of the real work piece. The machine health monitor records the load on different components of the machine tool and based on this history, it predicts the remaining life of each of the components.

In the future, all information can be used to monitor, adjust, and optimize existing machining processes or even help selecting optimal process parameters for new processes. Furthermore, the data allows an automatic detection of cutter wear, as well as to schedule maintenance intervals based on the predicted remaining life of the machine component. In a collaborative environment, also external suppliers or service providers could utilize this information for their own processes. For example, such a scenario would enable a machine tool manufacturer to predict, which machine needs new components and when, by receiving information about the load on machine tools. Hence, the manufacturer can stock up these components and ship them to customers just in time.

Overall, machine-process interactions are very complex and not constant over time. Due to the high uncertainties, a broad amount of process data from different processes is necessary to create and validate more advanced models. Depending on the desired output of the models, different data sources are required as well. To predict the remaining life of a component data from multiple machines with these components are required to optimize the prediction accuracy. Therefore, key aspects of an IoP are, who will be the owner of the data, where the data will be stored, and who has access and the resources to process it. Furthermore, a concept must be developed on how to make data accessible (locally, externally, or in the cloud [17]) and how to stream the data of large machine pools scalable through the network.

III. An Infrastructure for the IoP

To allow manufacturers to embrace the full potential offered by an IoP, we require a fundamental shift away from today's isolated infrastructure to a fully interconnected production landscape. Most importantly, as identified in our analysis of

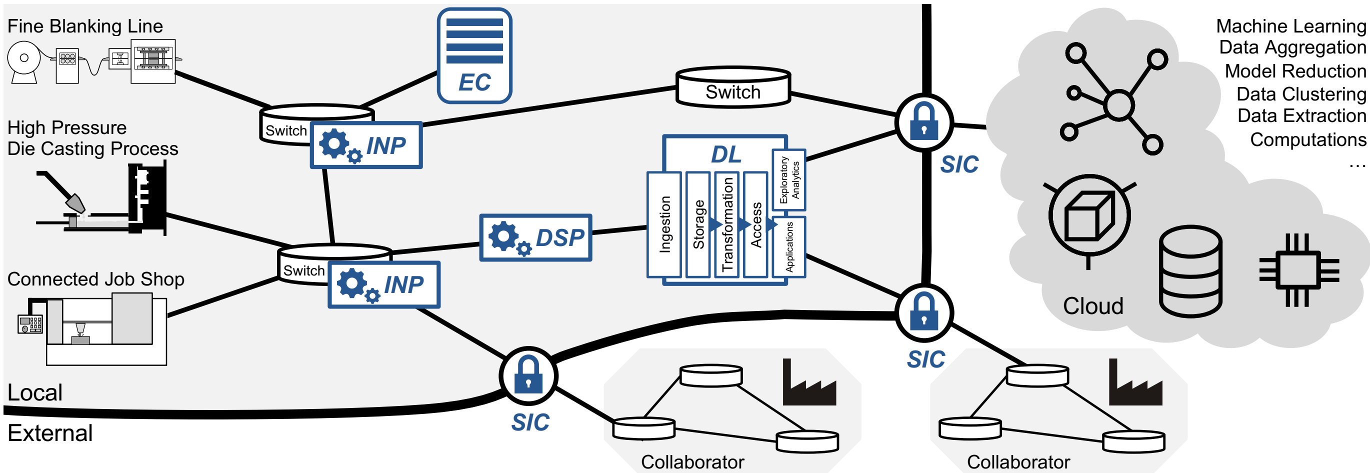

Fig. 1. The local infrastructure (*light gray background*) of the Internet of Production incorporates machinery-close sensors to improve today's manufacturing processes (as indicated by our three use cases on the left). Our overall design features three key components: (i) in-network processing (INP) and edge clusters (EC) for *seamless low-latency analysis and execution*, (ii) data stream processing (DSP) and data lakes (DL) for *scalable data processing and storage*, and (iii) and mechanisms for *secure industrial collaboration* (SIC) with external collaborators (bottom). Additional external processing power and more sophisticated models in the cloud (*dark gray background* on the right) complement these local resources.

the three real-world use cases, such an infrastructure needs to provide solutions for the following three distinct challenges: (i) *real-time control of tightly integrated production processes*, (ii) *storing and processing heterogeneous production data*, and (iii) *secure privacy-aware collaboration in production*. By tackling these challenges, we pave the way towards realizing the vision of an IoP and thus allow production systems to benefit from adequate cross-domain knowledge transfers.

We propose an infrastructure for an IoP and provide a roadmap to turn this infrastructure into reality. As shown in Figure 1, our envisioned infrastructure consists of three orthogonal approaches (marked in blue) that address the challenges derived from the use cases: (i) *seamless low-latency analysis and execution* to cope with huge bandwidth and real-time processing demands, (ii) *scalable data processing and storage* to allow for flexible data mining and reuse of measured and pre-processed manufacturing data, and (iii) *secure industrial collaboration* to create a global interconnection of production systems securely processing data of different collaborators.

In this infrastructure, production data flows from the machinery (represented by our three use cases on the left) through a local production site network to external resources in the cloud (right) for offloading data aggregation, complex data extractions, and compute-intensive machine learning algorithms as well as other production sites for collaborations (bottom). We observe that both mechanisms for *research objective (i)* (e.g., realized using in-network processing (INP) and edge computing (EC)) as well as *research objective (ii)* (e.g., achieved using data stream processing (DSP) and data lakes (DL)) are ideally realized *within* the local production site network. In contrast, the mechanisms for *research objective (iii)* are enforced at the border of the local production site network (SIC), i.e., before data is sent to external entities or other production site networks.

Subsuming our use cases, we identify common characteristics that an infrastructure for an IoP must exhibit to support the vision of highly-configurable, precise, and interconnected production sites. In the following, we discuss each of the three approaches in more detail, briefly refer to related work, and highlight the remaining challenges and research efforts required for turning the infrastructure for an IoP into reality.

A. Seamless Low-latency Analysis and Execution

The amount of data that needs to be exchanged not only between different parts of one production line, but also between lines and their respective controlling units and further superordinate systems, is expected to rise steeply in terms of frequency, numerosity, and size. Traditional factory floor control and interconnection architectures based on more or less centralized systems such as (networks of) PLCs are expected to be too limited in their scalability to keep up with the growing demands in the mid to long term [3]. New, more flexible architectures that enable a horizontal scaling and thus a distribution of the processing load within the network itself hence have be developed. Moreover, reconfigurations of production lines, including the physical movement of sub-systems (possibly even during active production steps), introduce uncertainties in the connectivities: A sub-system or controller may become temporarily unavailable during reconfiguration, or may move too far away from the process it was controlled by or used to control, so that continued proper operation cannot be guaranteed anymore due to losses or latencies.

Different efforts towards enabling inherently more decentralized and low-latency processing schemes have been made, although not specifically targeting industrial applications. For one, software-defined networking (SDN), with OpenFlow [18] and P4 [19] as the most prominent representatives of currently employed protocols and programming languages, allows pushing low-latency processing rules to in-network devices (INDs) such as routers and switches to facilitate packet filtering or the enforcement of forwarding policies. Not originally developed for the execution of intricate algorithms, INDs are however limited in their processing power, and the rules that can currently be executed on them are limited. Yet, first projects that combine SDN with more expressive languages such as eBPF have shown that INDs can be used to execute rudimentary control algorithms [20]. The extent to which INDs may be used for the execution of control, safety, or further IoP-related functionality, however, is a research question that needs to be addressed when developing of an infrastructure for an IoP. Further, to guarantee seamless operation in reconfiguration phases, methods for the coordinated distribution and activation of functionalities among INDs need to be developed.

For cases in which the capabilities of INDs do not suffice, the concept of edge computing (EC), i.e., offloading functionality to more powerful machines such as PCs or clusters situated close to the controlled processes, can be employed [21]. EC devices, while boasting large memory and powerful computation units, are however often equipped with general-purpose operating systems, whose complex network stacks incur delays due to processing overheads. Solutions aiming at either bypassing the stacks [22] or at partially offloading application logic into the operating system kernel [23] may be employed to lower these delays. Being situated farther away from the controlled processes than INDs, EC devices further likely cannot be used for processes such as controlling machine tools in job shops, which require latencies below 1 ms [24], which are hard to reach when not directly connected to the process. Thus, how to establish a precise interplay between higher-level, medium-latency EC devices and more low-level, low-latency processing on INDs is an open research question for realizing an infrastructure for an IoP.

At last, even when some IoP infrastructure has eventually been deployed, the inherently decentralized nature of this infrastructure will necessarily incur a re-evaluation of algorithms that have been used in the more centralized industrial scenarios of today. For example, whether feasible distributed control laws exist for all problems remains an open challenge [25]. Hence, when building an infrastructure for an IoP, questions of algorithmic complexity and tractability, as well as methods ensuring the co-location or confluence of certain pieces of information, also need to be considered.

B. Scalable Data Processing and Storage

With the ongoing change to use production data not only for control and management, new challenges arise for existing production networks. They are neither designed for transfers of large amounts of raw data, nor do they offer sufficient speed as soon as the data transfer takes place outside of the field level. However, many application scenarios, such as our use cases, require the results of analyses and decision making processes to be generated in real time (less than 10 ms), but even secondary tasks, e.g., data aggregation and compaction, must often be completed under rigorous time constraints to avoid overloading networking and storage resources. Lastly, integrating data from a variety of sources traditionally requires a large amount of manual effort. Within an IoP, these challenges can be addressed by employing *data lakes* in conjunction with *data stream processing* and semantic enrichment.

Data stream processing (DSP) provides a promising basis for the time-sensitive provision of processing results while distributing computations and reducing the network load. To prevent the network from being overloaded, algorithms for the selection, compression, pre-transformation, and aggregation of data on the edges must be distributed automatically, considering technical and functional requirements. Further challenges result from the design as a learning "system of systems".

Models for prediction or classification, e.g., must first be trained on a sufficient database. To use the learned models, the data sources have to be distributed and integrated into the data processing pipelines. Thereby, requirements for response speed and availability as well as consistency must be maintained. Integrated architectures, such as the Lambda architecture [26], are frequently used in such contexts and combine batch and stream processing in a parallel fashion, resulting in complex systems that are difficult to develop and maintain. Research and development of simplified approaches followed, such as the Kappa architecture [27]. Modern and advanced approaches, as implemented by Samza [28], take up these ideas and build on four main principles [29]: (i) *Everything is a stream*: Batch operations become a subset of streaming operations. Hence, everything can be treated as a stream. (ii) *Immutable data sources*: Raw data is persisted and views are derived, but a state can always be recomputed as the initial record is never changed. (iii) *Single analytics framework*: Keep it short and simple (KISS) principle. A single analytics engine is required. Code, maintenance, and upgrades are considerably reduced. (iv) *Replay functionality*: Computations and results can evolve by replaying the historical data from a stream.

Data lake (DL) architectures extend the concept of directly processing information by collecting and provisioning historic production data, i.e., older process data can be taken into account to adaptively improve the deployed manufacturing process. Consequentially, an IoP infrastructure further poses new challenging requirements on data storage and analytics. That is, the system should be able to handle the huge amount of heterogeneous data from different production machines in a "storage first, query later" manner. The infrastructure may not be able to process the massive collected real-time data and extract all the necessary information during data ingestion but rather at a later phase. Moreover, the specific data items of the production process and their utility may appear vague in the design phase, which only unravels after a deeper analysis or more human expert input. Traditional data warehouses cannot fulfill these aspects since they require a fixed schema, which only allows a specific and mostly predefined set of analysis.

For the required data processing needs in an IoP infrastructure, data lakes offer a big data repository which allows data to be ingested in their raw format. Thus, in real-time production environments, expensive overhead of transforming data for storage can be significantly reduced. Moreover, different from traditional data warehouses, data lakes support the concept of "schema-on-read". Data can be imported and stored in data lake first, and users can process the data, and analyze the complete history of a production process later.

Though there are existing data lake systems [30]–[32], these are mostly crafted for the specific requirements of an organization or focus on only certain DL-specific tasks, e.g., metadata management [33], data quality [34], user data analytics [35], or data preparation [36]. Additionally, existing systems are lacking semantic enrichment of data which allows for it to be shared and reused across application, enterprise, and community boundaries. A generic and extensible DL infrastructure for handling production data in an IoP is missing.

Foremost, the DL infrastructure for an IoP should be able to ingest, store, integrate, and query the data according to the specific process requirements within the network. Moreover, the data lake should be able to store heterogeneous ingested production data (structured data, semi-structured, unstructured data) in appropriate storages to reduce the transformation cost and facilitate easy information extraction later on. Furthermore, the data lake should not only satisfy the known production requirements, it should also be generic and extensible for possible future needs. Lastly, semantic enrichment, e.g., using Semantic Web technologies [37], should be employed to facilitate integration across organizational boundaries, different content, information applications, systems, and domains and enable the creation of machine-actionable knowledge.

To tackle these challenges, *data lake* technologies, such as *Constance* [33], [38], should be combined with *stream processing* capabilities and semantic enrichment [39], enabling scalable data processing and storage for production and opening up further opportunities for continuous and adaptive improvement of manufacturing processes.

C. Secure Industrial Collaboration

The vision of an IoP highly encourages industrial collaboration to maximize the benefits of cross-domain knowledge transfers. These collaborations can exist between companies along one supply chain or even between companies that implement similar processes (e.g., processing of the same material or operating related machinery): They exchange information to increase their revenue by adapting and improving their individual process. Due to confidentiality requirements, these entities should only interact over secured dedicated communication channels as their individual process knowledge is very valuable. Furthermore, they might be interested in masking their identity to hide their implemented processes from other entities or even competitors (e.g., to hinder conclusions about their individual process progress and to conceal themselves as a target for reverse-engineering attacks). However, depending on the granularity of exchanged information, de-anonymization is still possible regardless of any obfuscation.

To provide data security when utilizing, processing, and comparing confidential data, today, mainly two privacy-preserving approaches exist. First, factories can securely offload their computations to an (even untrusted) cloud without revealing their data [40], i.e., any operations are conducted on the encrypted data and, therefore, the (untrusted) third party is unable to access the processed data. Alternatively, collaborators can rely on secure multi-party computation to jointly compute a result without making individual inputs public [41]. Using such approaches, manufactures could, e.g., perform anonymous comparisons of the efficiency of their production processes [42]. To protect the privacy of different stakeholders, an IoP can initially protect participating entities through pseudonyms. In conjunction with digital signatures, such a system can also ensure data integrity [43] and, hence, allow for a secure industrial collaboration.

To address more advanced challenges in secure industrial collaborations, additional effort is required. First, the intended collaboration of an IoP might exceed today's privacy-preserving processing capabilities [44] and therefore render current implementations infeasible as the computations and models within an IoP are significantly more complex. Therefore, research must propose approaches that also support industrial models and data streams with vast amounts of heterogeneous information, potentially stored in data lakes, while still providing data security and privacy [45]. Second, accountability along the supply chain is an important aspect to certify the quality of goods. To this end, the data security aspect of an IoP also refers to new mechanisms that provide accountability and to conduct plausibility checks to protect companies from fake data. Third, on a different level, apart from tit-for-tat incentives for data sharing (as successfully applied in e.g., the BitTorrent protocol [46]), monetary transactions can motivate an information exchange of private (local) data. Here, the purchasing entity might be able to link these monetary flows to a stakeholder pseudonym, i.e., they can degrade the payee's privacy. Consequently, an IoP is in need of a secure privacy-sensitive system to handle smart payments. Fourth, a secure bootstrapping of collaborations (without a centralized third-party) between previously unaffiliated companies remains unsolved. Previously non-cooperating entities face the challenge of identifying collaborators that could potentially help them to benefit from cross-domain collaboration. For example, a company operating a particular machine is usually unaware of other companies using the same machine. Ideally, such a design limits the influence of potentially "more powerful" entities, such as the machine manufacturers, who might have an incentive to interfere with such a direct communication between their respective customers.

IV. Conclusion and Future Research Challenges

Through cooperation, manufacturing companies can create a new level of cross-domain collaboration as envisioned by an Internet of Production (IoP) to reduce product development costs, increase gains in profit margins, as well as generally improve product quality and safety. In this paper, we presented three distinct use cases (a fine blanking line, a high pressure die casting process, and a connected job shop) and based on their requirements, we derived a set of research goals for an underlying infrastructure: (i) *real-time control of tightly integrated production processes*, (ii) *storing and processing heterogeneous production data*, and (iii) *secure privacy-aware* *collaboration in production*. These factors are crucial to unleash the full potential of a global interconnection of production systems with the purpose of cross-domain knowledge transfers. We surveyed existing approaches that might help with our identified research goals. For example, in-network processing and edge clusters as well as data stream processing and data lakes have the potential to implement today's vision of an IoP. However, so far, no existing concept can solve all derived challenges as past research did not envision a global collaboration between manufacturing companies.

We proposed three objectives that are key factors when developing a new infrastructure that enables an IoP: (i) *seamless low-latency analysis of data*, (ii) *scalable data processing and storage*, and (iii) *secure industrial collaboration*. These approaches open up a completely new field of exciting future research challenges. We especially expect significant challenges wrt. interoperability when integrating today's devices and machines into a global infrastructure as well as adaptability when considering future devices and machines with yet unforeseen properties and demands. Hence, the overall infrastructure must support (secure) adaptability even after an initial (limited or local) rollout, especially in collaborative scenarios at global scale. Furthermore, increasing computational and storage resources will enable new production use cases based on complex operations that are not supported by today's networks, models, and available resources. Consequently, an infrastructure for the IoP must also be able to adapt to novel use cases. To conclude, realizing the infrastructure for a globally interconnected IoP is a challenging tasks that requires a tremendous interdisciplinary effort.

ACKNOWLEDGMENTS: The authors would like to thank the German Research Foundation (DFG) for the kind support within the Cluster of Excellence "Internet of Production" (IoP) under project ID 390621612.

REFERENCES

[1] J. Manyika *et al.*, "Big data: The next frontier for innovation, competition, and productivity," https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation, 2011 (accessed May 23, 2018).
[2] K. Liere-Netheler, S. Packmohr, and K. Vogelsang, "Drivers of Digital Transformation in Manufacturing," in *HICSS*, 2018.
[3] R. Glebke *et al.*, "A Case for Integrated Data Processing in Large-Scale Cyber-Physical Systems," in *HICSS*, 2019.
[4] S. Jeschke *et al.*, *Industrial Internet of Things and Cyber Manufacturing Systems*. Springer, 2017.
[5] L. Da Xu, W. He, and S. Li, "Internet of Things in Industries: A Survey," *IEEE Trans. Ind. Informat.*, vol. 10, no. 4, 2014.
[6] F. Klocke and A. Kuchle, *Manufacturing Processes*. Springer, 2009.
[7] R.-A. Schmidt *et al.*, *Cold Forming and Fineblanking*. Hanser, 2007.
[8] H. Voigts *et al.*, "Dependencies of the die-roll height during fine blanking of case hardening steel 16MnCr5 without V-ring using a nesting strategy," *Int. J. Adv. Manuf. Technol.*, vol. 95, no. 5, 2018.
[9] F. Birzer, *Forming and Fineblanking: Cost-effective manufacture of accurate sheet metal parts*. Verlag Moderne Industrie, 1997.
[10] J. Stanke *et al.*, "A predictive model for die roll height in fine blanking using machine learning methods," *Proc. Manuf.*, vol. 15, 2018.
[11] L. Spendla *et al.*, "Concept of Predictive Maintenance of Production Systems in Accordance with Industry 4.0," in *IEEE SAMI*, 2017.
[12] S. Cecchel *et al.*, "Lightweight of a cross beam for commercial vehicles: Development, testing and validation," *Mat. & Design*, vol. 149, 2018.
[13] F. Bonollo, N. Gramegna, and G. Timelli, "High-Pressure Die-Casting: Contradictions and Challenges," *JOM*, vol. 67, no. 5, 2015.
[14] N. Gramegna and F. Bonollo, "Smart Control and Cognitive System Applied to the HPDC Foundry 4.0: A Robust and Competitive Methodology Developed Under EU-FP7 Music Project," Tech. Rep., 2016.
[15] M. Rix *et al.*, "An Agile Information Processing Framework for High Pressure Die Casting Applications in Modern Manufacturing Systems," *Procedia CIRP*, vol. 41, 2016.
[16] M. Königs and C. Brecher, "Process-parallel virtual quality evaluation for metal cutting in series production," *Proc. Manuf.*, vol. 26, 2018.
[17] M. Henze *et al.*, *Network Security and Privacy for Cyber-Physical Systems*. Wiley, 2017.
[18] N. McKeown *et al.*, "OpenFlow: Enabling Innovation in Campus Networks," *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 2, 2008.
[19] P. Bosshart *et al.*, "P4: Programming Protocol-Independent Packet Processors," *ACM SIGCOMM CCR*, vol. 44, no. 3, 2014.
[20] J. Rüth *et al.*, "Towards In-Network Industrial Feedback Control," in *NetCompute*, 2018.
[21] W. Shi *et al.*, "Edge Computing: Vision and Challenges," *IEEE Internet Things J.*, vol. 3, no. 5, 2016.
[22] L. Rizzo, "Netmap: A Novel Framework for Fast Packet I/O," in *USENIX ATC*, 2012.
[23] F. Schmidt *et al.*, "Santa: Faster Packet Delivery for Commonly Wished Replies," in *ACM SIGCOMM*, 2015.
[24] A. Frotzscher *et al.*, "Requirements and current solutions of wireless communication in industrial automation," in *IEEE ICC*, 2014.
[25] L. Lessard and S. Lall, "Quadratic Invariance is Necessary and Sufficient for Convexity," in *ACC*, 2011.
[26] N. Marz and J. Warren, *Big Data: Principles and Best Practices of Scalable Realtime Data Systems*, 1st ed. Manning, 2015.
[27] J. Kreps, "Questioning the Lambda Architecture," https://www.oreilly.com/ideas/questioning-the-lambda-architecture, 2014 (accessed November 22, 2018).
[28] S. A. Noghabi *et al.*, "Samza: Stateful Scalable Stream Processing at LinkedIn," *VLDB Endowment*, vol. 10, 2017.
[29] N. Seyvet and I. M. Viela. Applying the kappa architecture in the telco industry.
[30] I. Terrizzano *et al.*, "Data Wrangling: The Challenging Journey from the Wild to the Lake," in *CIDR*, 2015.
[31] A. Y. Halevy *et al.*, "Managing Google's data lake: an overview of the Goods system," *IEEE Data Eng. Bull.*, vol. 39, no. 3, 2016.
[32] R. Ramakrishnan *et al.*, "Azure Data Lake Store: A Hyperscale Distributed File Service for Big Data Analytics," in *ACM SIGMOD*, 2017.
[33] R. Hai, S. Geisler, and C. Quix, "Constance: An Intelligent Data Lake System," in *ACM SIGMOD*, 2016.
[34] M. H. Farid *et al.*, "CLAMS: Bringing Quality to Data Lakes," in *ACM SIGMOD*, 2016.
[35] H. Alrehamy and C. Walker, "Personal Data Lake With Data Gravity Pull," in *IEEE BDCloud*, 2015.
[36] A. Maccioni and R. Torlone, "Crossing the finish line faster when paddling the Data Lake with Kayak," *PVLDB*, vol. 10, no. 12, 2017.
[37] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," *Sci. Am.*, vol. 284, no. 5, 2001.
[38] R. Hai, C. Quix, and C. Zhou, "Query Rewriting for Heterogeneous Data Lakes," in *ADBIS*, 2018.
[39] A. Pomp *et al.*, "Applying Semantics to Reduce the Time to Analytics within Complex Heterogeneous Infrastructures," *Technol.*, vol. 6, no. 3, 2018.
[40] X. Chen, "Introduction to Secure Outsourcing Computation," *Synthesis Lectures on Information Security, Privacy, and Trust*, vol. 8, no. 2, 2016.
[41] Y. Lindell, *Secure Multiparty Computation for Privacy-Preserving Data Mining*. IGI Global, 2005, ch. 189.
[42] M. Henze *et al.*, "Privacy-preserving Comparison of Cloud Exposure Induced by Mobile Apps," in *EAI MobiQuitous*, 2017.
[43] E. Gaetani *et al.*, "Blockchain-based Database to Ensure Data Integrity in Cloud Computing Environments," in *ITASEC*, 2017.
[44] D. Demmler, T. Schneider, and M. Zohner, "ABY – A Framework for Efficient Mixed-Protocol Secure Two-Party Computation," in *NDSS*, 2015.
[45] L. C. Gleim *et al.*, "Schema Extraction for Privacy Preserving Processing of Sensitive Data," in *MEPDaW-SeWeBMeDA-SWeTI*, 2018.
[46] B. Cohen, "Incentives Build Robustness in BitTorrent," in *P2PEcon*, 2003.