Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A.A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

GENERIC DATA MODEL FOR SEMICONDUCTOR MANUFACTURING SUPPLY CHAINS

Georg Laipple

Robert Bosch GmbH
Tübinger Straße 123
Reutlingen, 72762, GERMANY

Stéphane Dauzère-Pérès

Mines Saint-Etienne, Univ. Clermont Auvergne
CNRS, UMR 6158 LIMOS, CMP
Gardanne, F-13541, FRANCE

Thomas Ponsignon

Infineon Technologies AG
Am Campeon 1-15
Neubiberg, 85579, GERMANY

Philippe Vialletelle

STMicroelectronics
850 rue Jean Monnet
Crolles, F-38926, FRANCE

ABSTRACT

Semiconductor manufacturing systems, with several hundreds of different production steps, reentrant loops, cleanroom conditions and a job shop organization, are probably the most complex production systems and so is the manufacturing data structure. This high complexity requires precise planning, modelling and simulation to improve manufacturing transparencies and avoid planning inefficiencies. Although many modelling approaches for semiconductor production systems have been published, much less work has been conducted to define a common data model that supports the industrial complexity. This paper presents a common conceptual data model for semiconductor supply chains, developed in close collaboration by three European semiconductor manufacturers. Described semiconductor manufacturing use cases motivate the proposed conceptual data model. This model includes master entities, defining the structure of semiconductor supply chains with common definitions of key elements, and tracing entities, describing time dependent changes and events of the system. Expected benefits of the conceptual data model are also discussed.

1 INTRODUCTION TO SEMICONDUCTOR MANUFACTURING SYSTEMS

A semiconductor supply chain is an extremely complex and dynamic global system that involves multiple companies and sites. This is the reason semiconductor supply chains are often described with the terms semiconductor supply network or semiconductor supply chain network. In frontend, wafers are structured in clean room job shop wafer fabs. In the backend of the supply chain, structured wafers are divided to microchips and assembled, packaged and tested in flow shop facilities. Within a frontend wafer fab, hundreds of different products are produced on the job shop equipment through over 1000 different process steps. The complexity of material flows increases due to many reentrant loops.

The combination of diverse global pull-oriented flow shop systems in backend and huge push-oriented job shop systems in frontend requires expensive decoupling stocks. The cycle times of the semiconductor supply chains vary from a few weeks to several months. Due to the extreme diversity and complexity of semiconductor manufacturing systems, several specialized semiconductor companies collaborate within the same supply chain. This increases the economies of scale but also the coordination effort. Global optimization of those complex supply chains allows high efficiency improvements. Therefore, as a first step, a generic data model has to be developed in collaboration with the main supply chain partners to unify the description of the system structure, system behavior and system control. A generic semiconductor

978-1-5386-6572-5/18/$31.00 ©2018 IEEE

manufacturing data model enables the application of unified optimization models and planning approaches for the whole collaborative supply chain. Up to now, the effort of data modelling is usually performed by each single participating company independently. Thereby, data models are different throughout the supply chain and only local optimization models can be applied. This limits synchronization and global optimization of the huge, dynamic, collaborative semiconductor supply chains. The essential lack of a generic data model is the motivation for this paper, which proposes a conceptual generic data model describing the structure, behavior of semiconductor supply chains. The model has been developed in close collaboration with Bosch, Infineon Technologies and STMicroelectronics.

In the following section, different industrial use cases from Bosch, Infineon Technologies and STMicroelectronics are introduced, motivating the necessity of the generic data model. Then, in Section 3, a literature review characterizes the lack and the need of a generic data model for semiconductor supply chains. The structure of the proposed data model is then introduced and the single entities are listed in Section 4. The paper concludes with the description of necessary further steps and the invitation for discussion of the conceptual data model in further research.

2 INDUSTRIAL USE CASES FROM THE PROJECT PRODUCTIVE4.0

With the following industrial use cases of three different semiconductor manufacturing companies, the necessity and benefits of a generic data model for semiconductor manufacturing supply chains are highlighted from various views. A generic supply chain data model has to be able to capture requirements from all these use cases and to help every company synchronizing itself with the supply chain to sustainably increase efficiency in dynamic collaborative supply chains.

2.1 Bosch Use Cases

The semiconductor division of the Robert Bosch GmbH is integrated into a manufacturing supply network that is continuously expanding both internally and externally. To be able to efficiently plan, control and optimize this complex dynamic network, it is necessary to synchronize the planning processes of all the different contributors. Therefore, individual planning frameworks have to be unified. The first elementary step towards company overarching supply chain synchronization is the development of a generic data model describing both the structure and the history of the dynamic supply chain.

With a generic semiconductor data model, Bosch expects the following two disruptive advantages for effectiveness and efficiency in semiconductor supply chains:

1. The ability to develop unified optimization models (either from science or from industry) that can be applied to any semiconductor manufacturer without exchanging sensitive production data.
2. The ability to synchronize the dynamic collaborative supply chains within a common planning process basing on this generic data model.

Every supply chain contributor has to be able to maintain this data model. On the basis of this generic data model, a common planning process can be established to enable company- and site-overarching supply chain planning and synchronization to increase service levels with efficient production planning and control. Bosch is on the basis of this generic data model in the public co-funded project Productive4.0 aiming for a hierarchical supply chain simulation-based optimization model to optimize the following mid- to long-term production use cases:

1. Integration and ramp-up of new products into a mature supply chain. Thus, capacities, capabilities and release plans can be adjusted considering given demand and maximization of service levels.
2. Integration and loading of a new fab into the mature supply chain. Therefore, capacity and capability of the new fab can be adjusted and release plans for the supply chain can be determined under given demand considering maximization of service levels and balancing of utilization.

3. Change of product mix in a mature supply chain. Therefore, capacities and new demands are given, just release plans, capabilities and control policies can be adjusted to maximize service levels.

These mid- to long-term use cases all aim for global optimization across the entire supply chain. The common data structure presented in this paper is expected to support first steps in this direction.

2.2 Infineon Technologies Use Cases

Infineon Technologies operates a global production chain that is composed of eighteen owned frontend and backend fabrication sites located in North America, Central and Eastern Europe, and Asia complemented by production partners. The specialization of the fabrication sites in certain technologies and process blocks led to the emergence of complex flows of materials across the fabs. The average production process nowadays spans over at least four countries and three trips around the world. As a result, the scope of operational excellence is no longer confined to single fabs, but it also requires an efficient utilization of the resources available in the global supply chain. Thus, supply chain management became a key element for being able to match the volatile market demand with the production capacities (Ehm et al. 2011). While the complexity of operations in and across the fabs increased, the need for a better understanding of the interactions within the supply chain arose. Simulation offers undoubted advantages for analyzing the impact of planning decisions onto the execution system and vice-versa. Infineon Technologies pursue this approach for several years now where simulation models are built to mimic the behavior of the real-world production chain. However, the availability of data in the right quantity and quality is a major challenge to allow an accurate representation of the reality. So far, different attempts have been made to populate our simulation models with approximated data, called synthetic data (Yuan and Ponsignon 2014; Sutter and Ponsignon 2016). While this approach can be suitable for selected input parameters that require a tedious effort for the extraction and preparation, we believe that real data (or related statistical distributions) should be privileged when used to provide sources of uncertainty in the simulation. Thus, we see the development of this generic data model as an opportunity to provide a cross-validated set of data from which the modeler can pick relevant data to perform his simulation experiments with an accrued accuracy.

In Ewen et al. (2017), a semiconductor supply chain simulation testbed was proposed and made available to researchers and practitioners. However, as discussed in Section 3, the semiconductor manufacturing model relies on the MIMAC data sets of Fowler and Robinson (1995). Hence, with this generic data model, we aim at completing our effort towards an improved representation of the production system. Because this more detailed modeling comes along with an increased computational burden when simulating an entire supply chain, we expect the hierarchical structure of the generic data model to support a proper aggregation of data. One of its anticipated applications will be to enable the analysis of planning decisions taken at the supply chain level (i.e., demand planning, capacity planning, master planning) versus local planning and control decisions taken in each production sites (i.e., production scheduling, and detailed dispatching). In our focus is the top-down disaggregation of plans and schedules, the bottom-up propagation of early warnings, and the stability of decisions in a rolling horizon setting.

2.3 STMicroelectronics Use Cases

STMicroelectronics supply chain is currently undergoing major change, because it targets not only the most aggressive technologies, but also the differentiated technologies whose demand is fueled by the Smart Driving (ADAS) and the Internet of Things (IoT). The objective of STMicroelectronics in the public co-funded project Productive4.0 is then to meet the challenges of Industry4.0 by focusing on the agility and competitiveness of its supply chain.

As they are generally built from history, aggregated capacity models used at enterprise level may severely underestimate the impact of mix changes on actual fab capacity and corresponding product cycle time. Most of the time, process and recipe qualifications (Johnzén et al. 2011) must be considered in details in order not to jeopardize the on-time delivery of products to customers. On the other hand, very detailed models used at factory level for purposes such as dispatching and scheduling or OEE improvement cannot be handled at enterprise level because of their complexity which means explosion of computation time.

1. *Building a factory meta-model for improving supply chain visibility*. The first use case for STMicroelectronics is to develop a modelling approach to automatically derive aggregated fab models usable at supply chain level. These meta-models would be extrapolated from the detailed ones by providing a "mix hypothesis" (typical, min, max by product type) as input. They could then be used as black boxes, provided the mix hypothesis is kept valid. In a later evolution, the validity domain of these models could be extended by addressing the modelling of product cycle time, i.e. the impact of equipment qualification, loading and variability.
2. *Improving Supply Chain effectiveness through better execution*. The second use case aims at modelling the demand in terms of "probability" in order to improve the adequacy and the robustness of process qualifications. STMicroelectronics today uses a tool to manage qualifications (Johnzén et al. 2011), which determines the best qualifications (tool, recipe) to be performed to optimize the toolset workload balancing and robustness over a given horizon. This tool is today taking as input an "average demand" over the considered horizon, not considering peaks that happen every day on specific recipes. Taking random demands into account in qualification decisions, will enable better anticipation of (long or heavy) qualifications, thus reducing variability and improving cycle times and factory agility.

Working with the other partners in the public co-funded project "Productive4.0", STMicroelectronics aims at developing a generic model to facilitate the exchanges between the various stages or levels of the global supply chain. An additional objective is to minimize the effort needed to maintain the various models needed at each level, thus focusing resources on tasks with real added value.

3 SHORT LITERATURE REVIEW IN SEMICONDUCTOR DATA MODELLING

The need of data to model modern semiconductor manufacturing facilities has recently been emphasized by Hassoun and Kalir (2017). According to Hassoun and Kalir (2017), most of the research has been conducted on the MIMAC data sets of Fowler and Robinson (1995), which should be renewed for multiple reasons. In particular, more recent features such as Queue Time Constraints (QCTs), also called Time Constraint Tunnels, are ignored. They are becoming critical (see for instance Lima et al. 2017) since the latest product technologies require more and more QCTs, i.e. time constraints between two non-consecutive operations in a product route. Another important reason to revise the MIMAC data sets is their limited size, e.g. the maximum number of process flows is 21 (and usually much smaller) whereas hundreds of process flows can be found in the European fabs of Bosch, Infineon Technologies and STMicroelectronics. Rose (2000) demonstrates why using more complex models, such as the MIMAC ones at the time, is more relevant than using simpler models. The same argument supports the fact that more complex models than the MIMAC ones are now necessary to study modern semiconductor manufacturing facilities and supply chains.

Supply chain models using Discrete Event System Specification (DEVS) and Model Predictive Control (MPC) approaches are proposed in Huang et al. (2009). However, the supply chain is for a high-volume production structure and remains quite simple. Testbeds for much more complex supply chains are proposed in Ewen et al. (2017), together with reduction techniques that are assessed by comparing a detailed model and the reduced models. The building blocks of the reference data sets are the base system and process, the customers and demand generation, and the planning and control system. An architecture was suggested with information and control flows. A particular attention was given to achieving a realistic depiction of the supply chain's behavior. However, in Ewen et al. (2017), the wafer manufacturing facilities are limited to the MIMAC data sets. Hence, the proposed testbed lacks details about modern semiconductor manufacturing.

To answer these needs, this paper aims at formalizing a generic data model that encompasses the most relevant characteristics of semiconductor manufacturing supply chains. This work can be seen as supporting the goal of Hassoun and Kalir (2017) to provide a structure for more meaningful industrial data and of Ewen et al. (2017) to support testbeds at the supply chain level. Our goal is to provide the main building blocks of the reference data sets in Ewen et al. (2017), but with more complexity at the factory level.

4 THE CONCEPTUAL DATA MODEL

4.1 Structure of the Conceptual Data Model

Semiconductor supply chains can be described as multi-level hierarchical networks. To be generic, a semiconductor data model must be able to consolidate the different aggregation levels. The data model must not be restricted to single production systems or machine types but should provide a description as general as possible without losing detail or accepting redundancy. From the same data model, flow shop systems from backend as well as job shop systems from frontend have to be captured. This enables users from different industrial use cases to describe models of different aggregated levels with the same data model. The generic data model for semiconductor manufacturing supply chains thereby contains the central master and tracing entities. While master entities describe the system, tracing entities in this context store all time dependent system changes and events and have to be continuously updated. The master entities will serve as reference entities for the tracing entities. Tracing entities refer via Foreign Key (FK) relations to the descriptive master entities. Described objects in the master entities can be identified with Primary Key (PK) relations. This strict segmentation reduces the size of the tracing tables and avoids redundant data. Additionally, to initiate the models, various snapshot entities are integrated to the generic data model. Snapshot entities are updated event-based from the underlying databases and describe the actual system state concerning work in process (WIP), equipment state and actual demand. Specific further entities in the data model describe strategical parameters such as target product cycle time and target product costs. These entities are also updated event-based or on demand.

4.2 Master Entities

The tables below detail the different master entities of the generic semiconductor data model. The entity “Supply Chain” describes the supply chain from the topmost level with all possible material flow edges, referring on the FabIDs. The “Plant” entity describes the plant as a physical location of one or more fabs. The Plant belongs to a supply chain and can either be internal or external. The “Fab” entity depicts the fab as a technological and physical production unit. The overall organizational working structures within the fabs is described in the entity “Workshop”. Each workshop has a certain technical and operational workforce organized in shifts described in the “Shift” entity.

Table 1. Master Entities: “Supply Chain”, “Plant”, “Fab”, “Workshop” and “Shift”.

Supply Chain	**Data Type**	**Description**
ID	String	PK; Name of the supply chain
From FabID	String	FK; Source ID of supply chain edge
To FabID	String	FK; Sink ID of supply chain edge
Plant	**Data Type**	**Description**
ID	String	PK; Will generally refer to the geographical location of the plant
Supply Chain ID	String	FK
Internal	Boolean	Does the plant belong to the company or to an external partner
Fab	**Data Type**	**Description**
ID	String	PK; Name of the production unit
Plant Id.	String	FK; Geographical location

Table 2. Master Entities: “Supply Chain”, “Plant”, “Fab”, “Workshop” and “Shift”.

Type	String	Front End, Wafer Testing, Back End Assembly, Back End Final Test, …
Internal	Boolean	Does the fab belong to the company or to an external partner
MaxWIP	Integer	Hard capacity restriction of process units (NULL if no restriction)
Process unit	String	Smallest process unit in the fab (Wafer / Chip/ Device)
Workshop	**Data Type**	**Description**
ID	String	PK; Name of the workshop
FabID	String	FK; Name of the production unit
Type	String	Functional, divisional
Shift Model	String	
Internal	Boolean	Does the Workshop belong to the company or to a production partner
Working days per week	Integer	Days productive per week
Working hours per day	Time	Hours productive per day
Shift	**Data Type**	**Description**
ID	String	PK; Name of shift
Workshop ID	String	FK; name of workshop the shift is assigned to
Shift duration	Time	Duration of a shift in the workshop in hours
Workforce operational	Integer	Overall number of operational workers assigned to the shift
Workforce technical	Integer	Overall number of technical workers assigned to the shift

The following entities describe the behavior of every machine in semiconductor facilities. Each machine is allocated to a group of machines doing similar processes. The equipment groups are described and allocated in the master entity “Equipment Group”. The entity “Equipment” describes the semiconductor equipment of any equipment model. It allocates it to higher level entities and describes the behavior of the equipment. MTTR and MTBF are thereby pre-calculated values. Deviations can be calculated from historical data from the tracing entities. The setup matrix in the “Setup” entity describes the duration of sequence dependent setups for every single machine. In parallel, a material transfer matrix estimates the duration of material from on to another equipment. The entity “Maintenance” includes all different maintenance types and their durations. A maintenance can be described several times and can vary slightly between the equipment. The maintenance scheduling determines the way a maintenance is scheduled (counter- or time-based), the maintenance frequency determines the “distance” between the maintenances either in completed units or in machine hours.

Table 3. Master Entities: “Equipment Group”, “Equipment”, “Setup” and “Maintenance”.

Equipment Group	Data Type	Description
ID	String	PK; Name of the Workshop
Fab ID	String	FK
Workshop ID	String	FK; organizational unit responsible for the equipment group
Type	String	Refers to the technology used for the process tool: Lithography, Plasma Etching, Diffusion, Tester, Prober, etc.
Equipment	**Data Type**	**Description**
ID	String	PK; Name of the machine
Workshop ID	String	FK; Operational Workshop ID
Equipment Group ID	String	FK; technical equipment group affiliation; FK to FabID
Location	String	Ideally (X,Y) coordinates but can be the name of a building, floor or area
Maximum Batch Size	Integer	Physical limit (can be overridden at recipe / product / step level)
Number of load ports	Integer	
Maximum Parallelization	Integer	Number of parallel used process chambers or max parallel used load ports
Number of clustered Steps	Integer	If more than one step in a sequence is operated by same equipment, it is called cluster equipment with a number of clustered steps >1; else 1
SOP	Date	Date of first production

Table 4. Master Entities: “Equipment Group”, “Equipment”, “Setup” and “Maintenance”.

Delivery mode	String	Refers to the transportation mean that is used to deliver products on the considered machine (i.e. AMHS, Operator, AGV, etc.)
Dispatch mode	String	Refers to the system used to allocate products to the machine (can be Full-Auto, Auto-planning, etc.)
Internal Buffer Size	String	Used for furnaces (hidden loading/unloading)
MTTR	Time	Mean time to repair
MTBF	Time	Mean time between failure
Setup	**Data Type**	**Description**
From Recipe ID	String	FK; last setup operation state
To Recipe ID	String	FK; target operation state
Equipment ID	String	FK; Name of equipment
Duration	Time	Duration of setup
Maintenance	**Data Type**	**Description**
ID	String	PK; Name of Maintenance
Duration	Integer	Duration in minutes
Maintenance Scheduling	String	Counter-based, time-based
Maintenance Frequency	Integer	Number of units between maintenance (if maintenance scheduling is counter-based, in units are wafers completed, otherwise units are machine hours)
Technician Requirement	String	Number of technical workers required for the maintenance
Equipment ID	String	PK; Name of equipment needing this maintenance

The “Worker Qualification” entity describes the qualification state of technical and operational workforce of a shift. The attribute maintenance ID is only necessary for technical workers. If maintenance is “NULL”, the necessary workforce specifies the number of operational workers necessary to run the equipment.

Table 5. Master Entities: “Worker Qualification” and “Material Transfer”.

Worker Qualification	Data Type	Description
Shift ID	String	FK; Name of shift
Maintenance ID	String	FK; Name of maintenance
Equipment ID	String	FK; Name of equipment
Workforce	Integer	Number of workers necessary to maintain / run equipment
Material Transfer	**Data Type**	**Description**
From Equipment ID	String	FK; start equipment of material transfer
To Equipment ID	String	FK; target equipment of transfer
Material Transfer Time	Time	Duration of transfer from on to the other equipment
Transfer mode	String	Automation, AGV, milkrun or manual

Another cluster of three master entities are the workflow descriptive master entities. The master entity “Operation” depicts the generic behavior of a single process step. Each operation has a foreign key relation to equipment and to route which means that the operation is required by the specific routes and requires a specific equipment. The master entity “Route” describes the structure of a workflow out of operations. Every product uses a specific route in its value stream, i.e. the complete sequence of operations with all mandatory and optional operations necessary to complete the product. A route makes use of the single operations. Additionally, the master entity “Time constraints” is necessary. As single operations can set the lot to a time-critical state, time constraints can ensure that the operation setting it to an uncritical state is not delayed. Also the other case where certain time has to be spent between two operations can be covered with this master entity.

Table 6. Master Entities: “Operation”, “Route” and “Time Constraints”.

Operation	**Data Type**	**Description**
ID	String	PK; Name of operation / process step
Next cluster operation	String	If operation is part of a cluster sequence, name of next cluster operation
PTime per unit	Time	Capacity consuming / relevant equipment processing time
CTime per unit	Time	Overall cycle time spent on the tool (incl handling etc.)
Variance PTime	Time	Variance of capacity relevant operation processing time
Variance CTime	Time	Variance of overall operation cycle time
Sampling rate	Double	Execution rate of operation If operation is mandatory sampling rate is 1 (1-sampling rate = skipping rate of optional operation)
Recipe ID	String	Recipe name of the operation
Yield Probability	Double	Yield rate of the operation in percent
Route	**Data Type**	**Description**
ID	String	PK; Name of route (usually equals product names requiring the route)
Sequence Number	Integer	PK; Sequential number of operations along the route
Operation ID	String	FK; Name of operation
Block	String	Name of block / section / stage / layer route operation belongs to
Time Constraints	**Data Type**	**Description**
ID	String	PK; Name of time constraint
Route ID	String	FK; Name of Route Time constraint belongs to
Start Sequence Number	Integer	FK; sequence number in the route time window starts
End Sequence Number	Integer	FK; sequence number in the route time window ends
Maximum Time	Time	Maximum Duration
Minimum Time	Time	Minimum Duration

To describe products in the generic data model, there are several hierarchical aggregation levels 1 to 4. “Product Level 4” is the descriptive level with the highest granularity. The product levels 3 to 1 are aggregated levels that are based on information from product level 4. The assignment of level 4 to aggregated levels is m:1. The static bottlenecks are determined in the “Product Level 3” master entity because the mid capacity planning is done on this first aggregated product level. There is no big difference in capacity from the aggregated level 3 product to its child level 4 products. As along the supply chain, secondary products are assigned to primary products, the “BOM” master entity characterizes the possible assignment.

Table 7. Master Entities: “Product Level 4” to “Technology”, “Route” and “Bill of Materials”.

Product Level 4	**Data Type**	**Description**
ID	String	PK; Name of the product in highest granularity
Route ID	String	FK; Name of the route assigned for the product
Supply Chain ID	String	FK; Name of the supply chain assigned for the product
Product level 3 ID	String	FK; Assignment of product to first product aggregation level
Chips per Wafer	Integer	Number of chips on a wafer
Product Level 3	**Data Type**	**Description**
ID	String	PK; Name of the product in level 3
Product level 2 ID	String	FK; Assignment of product to second product aggregation level
Bottleneck Equipment Group ID	String	FK; Name of equipment group that limits material flow according to capacity planning
Max units per week	Integer	Maximum of unit completes (Wafers / chips) per week on the bottleneck equipment group and thereby on the product
Product Level 2	**Data Type**	**Description**
ID	String	PK; Name of the product in level 2
Technology ID	String	FK; Assignment of product to highest product aggregation level (technology)

Table 8. Master Entities: “Product Level 4” to “Technology”, “Route” and “Bill of Materials”.

Technology	**Data Type**	**Description**
ID	String	PK; Name of the technology
Supply Chain ID	String	FK; Assignment of technology to supply chain
Bill of Materials (BOM)	**Data Type**	**Description**
Edge ID	String	PK; Name of assignment
From Product Level 4	String	FK; Name of secondary product level 4
To Product Level 4	String	FK; Name of primary product level 4
BOM Factor	Integer	Number of secondary product units assembled to one primary product unit

The master entities build the descriptive part of the generic data model. From these entities, all necessary information about the structure of the supply chain and the production system can be extracted.

4.3 Tracing Entities

Within the tracing entities, all historical data can be stored making the structural framework described in the master entities above vivid. In the “Event Calendar” tracing entity, certain events like planned facility shutdowns, etc. can be scheduled. The “Lot State” tracing entity covers all historical process starts and ends of a lot in any facility. The “Lot State” tracing entity provides the full history of operations of a specific lot. Additionally, the “Lot Event” tracing entity gives the history of lot events (e.g. Hold events, Split events, Merge events, etc.). Equivalently to the lots, the equipment also has to be historically tracked. The “Equipment State” tracing entity tracks all dates of equipment state changes. In case the state is changed to maintenance, the attribute maintenance ID refers to the master entity maintenance.

Table 9. Tracing entities: “Event Calendar”, “Lot State”, “Lot Event”, “Equipment State” and “Equipment capability”.

Event Calendar	**Data Type**	**Description**
Workshop ID	String	FK; Name of the workshop
Fab ID	String	FK; Name of the facility
Event Start	Time	Start Date
Duration	Time	Event Duration in hours
Capacity loss	Double	Facility capacity loss in percent
Frequency	String	Unique, annual, quarterly, monthly, daily,
Lot State	**Data Type**	**Description**
Lot ID	String	PK; Name of the lot
Operation ID	String	PK; Name of operation
Sequence Number	Integer	FK; Sequential number of the operation lot has entered
Product level 4	String	FK; Name of product level 4 lot belongs to
Route ID	String	FK; Name of Route lot belongs to
Parent Lot ID	String	FK; Name of ancestor lot (important to rebuild the lot history)
Customer Due Date	Date	Date customer needs the lot delivered (DD.MM.YYYY hh:mm:ss)
Operational Due Date	Date	Optional date production planning has set to guarantee fulfillment of customer due date (DD.MM.YYYY hh:mm:ss)
Enter State	Time	Time state is entered
Leave State	Time	Time state is finished
Enter Quantity	Integer	Amount of units in the lot when entered
Lot Event	**Data Type**	**Description**
Lot ID	String	PK; Name of the lot
Operation ID	String	FK; Name of operation
Sequence Number	Integer	FK; Sequential number of the operation lot has entered
Event ID	String	PK; Name of event that occurred to lot (Hold, Split, Merge, Scrap, etc.)
Parent Lot ID	String	FK; Name of ancestor lot (important to rebuild the lot history)

Table 10. Tracing entities: “Event Calendar”, “Lot State”, “Lot Event”, “Equipment State” and “Equipment capability”.

Enter Main Quantity	Integer	Number of units in the lot (wafers / chips) when entered the event
Leave Main Quantity	Integer	Number of units in the lot (wafers / chips) when left the event (Important for splits, merges and scraps)
Event Date	Date	Date event occurred (DD.MM.YYYY hh:mm:ss)
Equipment State	**Data Type**	**Description**
Equipment ID	String	FK; Name of the equipment
Last State	String	Name of last state (e.g. productive, standby, scheduled down, etc.)
State	String	Name of equipment state (e.g. productive, standby, scheduled down, etc.)
Change DT	Date	Date of state change from last state to state (DD.MM.YYYY hh:mm:ss)
Maintenance ID	String	FK; Name of maintenance in case state is maintenance
Equipment Capability	**Data Type**	**Description**
Operation ID	String	FK; Name of operation
Equipment ID	String	FK; Name of the equipment
Qualification DT	Date	Date operation is qualified on equipment (DD.MM.YYYY hh:mm:ss)
Disqualification DT	Date	Date operation is disqualified on equipment (DD.MM.YYYY hh:mm:ss)
Residual Qualification Period	Time	Time left until disqualification of operation (in case operations have to be frequently checked and requalified)

4.4 Snapshot Entities

Snapshot entities are introduced to initiate models. The information stored in these entities can also be extracted from the tracing entities. But as it can be quite complex and may require high redundant calculation times, it is useful to integrate frequently updated entities to provide the initial information. Thereby, for calculation time reasons, snapshot data are accepted to contain partially redundant data. The “Current WIP” entity describes the actual material distribution in the supply chain. The “Demand” entity stores redundant demand at every product level.

Table 11. Snapshot entities: “Current Machine State”, “Current WIP” and “Demand”.

Current Machine State	Data Type	Description
Equipment ID	String	FK; Name of equipment
Current state	String	Name of actual state
Prev. Change DT	Date	Change date from previous to current state (DD.MM.YYYY hh:mm:ss)
Next State	String	Potential next state
Expected Change DT Mean	Date	Expected mean date for change to next state (DD.MM.YYYY hh:mm:ss)
Expected Change DT Variance	Time	Expected variance date for change to next state
Current WIP	**Data Type**	**Description**
Operation ID	String	FK; Name of operation
Equipment ID	String	FK; Name of equipment
Product level 4 ID	String	FK; Name of product level 4
Lot ID	Sting	FK; Name of Lot
Current State	String	Actual state of Lot ID
Time in State	Time	Time lot already spent in current state
Priority	Integer	Priority of the lot
Current Main Quantity	Integer	Current amount of units in the lot
Residual Time in State	Time	Expected residual time lot spends in current state
Demand	**Data Type**	**Description**
Type of Demand	String	Demand forecast or real demand
Product Level 2 Quantity	Integer	Quantity of demand for product level 2 for due date
Product Level 3 Quantity	Integer	Quantity of demand for product level 3 for due date
Product Level 4 Quantity	Integer	Quantity of demand for product level 4 for due date
Product Level 2 Due Date	Date	Due date for demand product level 2 (DD.MM.YYYY hh:mm:ss)

Table 12. Snapshot entities: "Current Machine State", "Current WIP" and "Demand".

Product Level 3 Due Date	Date	Due date for demand product level 3 (DD.MM.YYYY hh:mm:ss)
Product Level 4 Due Date	Date	Due date for demand product level 4 (DD.MM.YYYY hh:mm:ss)
Current Main Quantity	Integer	Current amount of units in the lot

4.5 Strategy Entities

The strategy entities provide information on the system control targets and parameters. As explained earlier, these entities do not necessarily have to be maintained, as certain information may not exist in a structured way or are underlying highest confidentiality requirements. This is the reason these entities will only be described shortly in this section.

The "Release Plan" entity shows the quantity planned to be released to a fab or facility. The "Cycle Time Commitment" entity helps controlling the WIP via the release plan. The "Delivery Commitment" directly corresponds to the release plans and cycle time commitments introduced in the corresponding entities. The "Target Release Level" entity is an additional target that depends from the release and commitment policy. The service level targets directly affect the delivery commitment and the release plan.

5 CONCLUSIONS AND PERSPECTIVES

Based on the joint effort of three semiconductor manufacturers within the European project Productive 4.0, a conceptual data model for semiconductor manufacturing supply chains is proposed in this paper. The model includes 33 different entities, decomposed into 19 master (structural) entities and 14 tracing, snapshot and strategy entities. The entities are referring each other via key relations. The overview can be globally taken from the conceptual ER-model, that the interested reader can get access to by contacting the authors. As this generic data model still is in the conceptual state, it will be further elaborated towards a logic data model and be implemented in the manufacturing partner's infrastructure to provide a generic data basis for the development of generally applicable simulation and optimization models.

We believe this work helps to provide common definitions to characterize semiconductor manufacturing systems and to synchronize planning in collaborative supply chains, but also to specify important challenges that the industry is facing nowadays through the description of use cases. We hope that they will help to foster new research.

In the future, we would like to get feedback from the academic and industrial semiconductor manufacturing community on the data model to complete it and refine it. The model will support the development of common planning and optimization models. Also a set of real and anonymized data can be shared between partners of the Productive4.0 project, but also potentially with interested parties outside the project. New relevant testbeds should then be made available, which is recognized as a need in the literature.

ACKNOWLEDGMENTS

This work has received funding from the Electronic Component Systems for European Leadership Joint Undertaking under grant agreement No 737459 (project Productive4.0). This Joint Undertaking receives support from the European Union Horizon 2020 research and innovation program and Germany, Austria, France, Czech Republic, Netherlands, Belgium, Spain, Greece, Sweden, Italy, Ireland, Poland, Hungary, Portugal, Denmark, Finland, Luxembourg, Norway, Turkey.

REFERENCES

Ewen, H., L. Mönch, H. Ehm, T. Ponsignon, J. W. Fowler, and L. Forstner. 2017. "A Testbed for Simulating Semiconductor Supply Chains". *IEEE Transactions on Semiconductor Manufacturing* 30(3):293-305.

Fowler, J. and J. Robinson. 1995. *Measurement and Improvement of Manufacturing Capacities (MIMAC): Final Report.* Technical Report 95062861A-TR, Austin, TX: SEMATECH.

Galvez, C. and F. Moya-Anegón. 2007. "Standardizing Formats of Corporate Source Data". *Scientometrics* 70(1):3–26.
Hassoun, M. and A. Kalir. 2017. "Towards a New Simulation Testbed for Semiconductor Manufacturing". In *Proceedings of the 2017 Winter Simulation Conference*, edited by W. K. V. Chan et al., 3612-3623. Piscataway, New Jersey: IEEE.
Huang, D., H. S. Sarjoughian, W. Wang, G. Godding, D. E. Rivera, K. G. Kempf, and H. Mittelmann. 2009. "Simulation of Semiconductor Manufacturing Supply-Chain Systems with DEVS, MPC, and KIB". *IEEE Transactions on Semiconductor Manufacturing* 22(1):164-174.
Johnzén, C., S. Dauzère-Pérès, and P. Vialletelle. 2011. "Flexibility Measures for Qualification Management in Wafer Fabs". *Production Planning and Control* 22(1)81-90.
Lima, A., V. Borodin, S. Dauzère-Pérès, and P. Vialletelle. 2017. "Analyzing Different Dispatching Policies for Probability Estimation in Time Constraint Tunnels in Semiconductor Manufacturing". In *Proceedings of the 2017 Winter Simulation Conference*, edited by W. K. V. Chan et al., 3543-3554. Piscataway, New Jersey: IEEE.
Rose, O. 2000. "Why Do Simple Wafer Fab Models Fail in Certain Scenarios?". In *Proceedings of the 2000 Winter Simulation Conference,* edited by J. A. Joines et al., 1481-1490. Piscataway, New Jersey: IEEE.
Steiner, R. 2014. *Grundkurs Relationale Datenbanken.* Wiesbaden: Springer Fachmedien.
Sutter, M. and T. Ponsignon. 2016. "SimData Toolbox: Simplifying the Analysis and Preparation of Input Data for Supply Chain Simulation." In Proceedings of the *6th International Conference on Information Systems*, Logistics and Supply Chain, edited by C. Temponi and N. J. Vandaele, 1-8. Berlin Heidelberg: Springer International Publishing.
Yuan, J. and T. Ponsignon. 2014. "Towards a Semiconductor Supply Chain Simulation Library." In *Proceedings of the 2014 Winter Simulation Conference,* edited by A. Tolk et al., 2522-2532. Piscataway, New Jersey: IEEE.

AUTHOR BIOGRAPHIES

GEORG LAIPPLE is Ph.D researcher in supply chain management at Karlsruhe Institute of Technology (KIT). His supervising professor is Prof. Dr. Kai Furmans. He is project manager of the public co-funded project Productive4.0 and member of semiconductor manufacturing engineering department in the Robert Bosch GmbH in Reutlingen. His e-mail address is georg.laipple@de.bosch.com.

STÉPHANE DAUZÈRE-PÉRÈS is Professor at Mines Saint-Etienne in France and Adjunct Professor at BI Norwegian Business School in Norway. He received the Ph.D. degree in 1992 and the H.D.R. in 1998. His research interests broadly include modeling and optimization of operations at various decision levels (from real-time to strategic) in manufacturing and logistics, with a special emphasis on semiconductor manufacturing. He has published more than 65 papers in international journals. He has coordinated many academic and industrial research projects. He won the Best Applied Paper of the Winter Simulation Conference in 2013. His email address is dauzere-peres@emse.fr.

THOMAS PONSIGNON works as a Senior Staff Engineer in the Corporate Supply Chain organization of Infineon Technologies AG in Munich, Germany. He holds a Ph.D. degree in Mathematics and Computer Science from the University of Hagen, Germany. His email address is thomas.ponsignon@infineon.com.

PHILIPPE VIALLETELLE is a Senior Member of Technical Staff at STMicroelectronics Crolles, France. He received an Engineering degree in Physics from the INSA in 1989. He is responsible for the definition of advanced methodologies and tools for the Crolles 300mm production line of STMicroelectronics. Expert in Manufacturing Sciences for Process and Production Control items, he is in charge of driving collaborative projects at European level. His email address is philippe.vialletelle@st.com.