TOWARDS A SEMICONDUCTOR SUPPLY CHAIN SIMULATION LIBRARY (SCSC-SIMLIB)

Jingjing Yuan
Thomas Ponsignon
Infineon Technologies AG
Am Campeon 1-12
Neubiberg, 85579, GERMANY

ABSTRACT

Simulation is a widely used technique for analyzing and managing supply chains. Simulation software packages offer standard libraries for selected functions and application areas. However, no commercial or freeware simulation tool proposes building blocks specific to semiconductor manufacturing. Thus, we propose in the present paper a library with a collection of simulation objects that can be used to model supply chain activities of various scales in the semiconductor industry. The library denoted by SCSC-SIMLIB strives for reducing modeling effort. It also enables standardization and benchmarking. We first describe the requirements for such a library and we suggest an architecture. Then, selected objects of the library are presented in more detail. Finally, we demonstrate the benefits of SCSC-SIMLIB, which is implemented by means of the simulation software AnyLogic, with a use case based on a simple example.

1 INTRODUCTION

The production of a microchip consists of thousands of manufacturing processes. One of the most essential steps is wafer fabrication, which is a series of complex physical and chemical processes that imprint layers of integrated circuits onto silicon wafers. There are more than 500 fabrication steps in wafer fabrication facilities (wafer fabs). Some steps may be repeated or changed in orders, depending on the different requirements of the end products. Then, the wafer test and sawing is conducted in the probe facility. Each individual die on the wafer will be tested and the good dies are marked with ink and traced by an electronic map. Next, the good dies on the probed wafers are reconstituted on laminate and encapsulated into a plastic or ceramic package. This happens in assembly facilities. Finally, in test facilities the assembled dies are tested again and packed into tubes, trays or tapes for shipment.

The operations in wafer fabs and probe facilities are referred to as the front-end of semiconductor manufacturing, with a typical cycle time of 40 to 100 days. The assembly and test processes make up the back-end production of a semiconductor, with a typical cycle time of 5 to 20 days. Front-end and back-end operations are decoupled by means of a stocking point for semi-finished products known as the die bank. After back-end operations, finished products are stored in a distribution center before shipment to the customers. The production partners at the front-end are called silicon foundries and subcontractors at the back-end. Front-end processes are complex, capital-intensive and technology-intensive, and thus most of the front-end production is performed in highly industrialized areas. Stable production flow and high capacity utilization are the major concerns at front-end.
Yuan and Ponsignon

In contrast, the back-end production is less complex and more labor-intensive and therefore located mostly in countries with low production costs. Among others, Uzsoy et al. (1992, 1994), Gupta et al. (2006) and Mönch et al. (2012) provide an overview of semiconductor manufacturing operations and challenges.

Due to the different characteristics and global presence of the various manufacturing steps, the supply chain management is an important strategic consideration for the semiconductor industry (Chien et al. 2011). In addition to the complex internal supply chain, the end-to-end supply chain that spans from the supplier’s supplier to the customer’s customer brings more challenges with different sources of uncertainty both on demand and supply side (Lee 2004).

Simulation offers great advantages for analyzing and managing supply chains. A simulation model is able to mimic the behavior of a real-world system with the incorporation of uncertainty. It can be used to analyze multiple performance metrics (Buckley and An 2005). Simulation is also an experiment method. By varying the inputs of a simulation model, it is possible to run what-if scenarios and compare various alternatives. This enables a decision maker to evaluate a project prior to its real execution and thus to reduce associated risks and take better decisions. In semiconductor industry, many supply chain challenges and use cases of different natures need to be addressed from different levels of the organization. The problems cover a broad range of levels of detail, from the factory level operational process to the global supply chain network design and integration. In the literature, most simulation models are built ad-hoc, dedicated to one particular project or application. As such, some commonly used functional modules (e.g., lots, machines, production sites) are often modeled and simulated repeatedly for every simulation project. Hence, a standard and reusable library of simulation objects is desired to save simulation effort and reduce modeling errors.

Many simulation tools provide standard libraries for selected functions or application areas (among others, Brooks Automation (2001); Kelton et al. (2010); Borshchev (2013)). In a library, lower-level processes or components are documented as objects, which can be reused across different models. These standard sub-processes are hidden in the objects and cannot be changed by the user, which makes the models more extensible and scalable but also easy to use. A simulation library for semiconductor supply chain is not yet available. Thus, we propose in the present paper the first steps towards a library with a collection of simulation objects that can be reused to model and standardize supply chain activities of various scales in the semiconductor industry. Our semiconductor supply chain simulation library is denoted as SCSC-SIMLIB throughout this document.

The paper is organized as follows. In the next section, we discuss related literature. In Section 3, we describe the requirements of a simulation library for semiconductor supply chains and we propose an architecture of SCSC-SIMLIB. Then, selected objects of the library are presented in more detail. In Section 4, we show an application of SCSC-SIMLIB with a practical use case using a simple supply chain model. We conclude in Section 5 and we give some future research directions.

2 RELATED LITERATURE

In the last two decades, simulation has become a popular approach to investigate problems that arise in semiconductor supply chains. Several publications use simulation based approaches to find adequate supply chain designs, to improve supply chain outputs, and to evaluate supply chain performance (among others, Ingalls (1998), Shapiro (2000), Chang and Makatsoris (2001), Persson and Olhager (2002) Fayez et al. (2005)). However, most of the published works rely on ad-hoc simulation models, which cannot be applied to other use cases than those they have been built for. It emphasizes the need for a standardized, modular, and multi-purpose modeling concept.

First steps towards a reference model for semiconductor supply chains are described in Ehm et al. (2011). A reference model is proposed that can be used as a test-bed to compare planning algorithms. A similar approach was followed within the MIMAC project to achieve a set of test-beds for simulating single wafer fabs (MASM 1997). While reference models provide a standardized framework to the research community, their pre-defined, fixed structure does not allow to address all kinds of problems.
Given the broad scope of problems in semiconductor supply chains, the modeling effort as well as the computational burden are of major concern. Several papers seek to achieve an aggregated representation of operations, which performs as good as full detailed simulation models. Among others, Jain et al. (1999) discuss the criticality of detailed modeling for simulating semiconductor supply chains. A reduction method for wafer fab operations is also suggested by Hung and Leachman (1999), where processing steps on non-bottleneck machines are replaced by delays. While the computing times are significantly reduced, the reduction approaches is dependent on the loading scenario. This does not satisfy the criterion for universality that is expected from a simulation library. Duarte et al. (2007) and Ehm et al. (2011) chose a compact representation of single manufacturing site. To a certain extent, we follow this approach for high-level representation of factories in SCSC-SIMLIB. Another stream of research seeks to reduce the simulation effort by means of distributed simulation. Lendermann et al. (2003), Chong et al. (2006) and Gan et al. (2007) uses a High-Level Architecture (HLA). However, due to the complexity and the large modeling effort, we do not follow this method in SCSC-SIMLIB.

We can conclude that none of the above mentioned works fulfills at the same time the criteria for modularity, universality, and standardization. This strengthens the gap for a multi-purpose library, which we propose to fill with SCSC-SIMLIB.

3 DESCRIPTION OF SCSC-SIMLIB

In this section, we provide guidelines for the implementation of a SCSC-SIMLIB by first discussing the requirements and criteria. Then, we suggest a library architecture for the effective organization of the library objects. We also describe in detail a selection of key library objects and the modeling methods.

3.1 Requirements of a Simulation Library for Semiconductor Supply Chains

In advance of modeling the objects, it is important to understand the requirements and requisite features of a quality simulation library. First of all, a clear and effective organization of library objects provides a straightforward framework for both developers and users. The objects are categorized based on the functions they perform and the level of supply chain to be simulated. The same supply chain elements or activities can be modeled as several objects to capture different levels of detail and performance measures. For example, a wafer fab object could be used as one of many manufacturing nodes in an end-to-end supply chain. In other cases, the detailed movements of lots within the wafer fab are of interest to the decision makers, so that it is necessary to model the work stations and machine operations within the factory object. As such, each library object should be defined for a specific purpose and developed for a certain level of detail. In addition, there should be clear organizations of material, information and financial flows within and between the library objects. A more detailed discussion of library structure will be presented in Subsection 3.2.1.

To facilitate the simulation of semiconductor supply chains in a full-scale, the library should include a variety of objects that span from customers, to planning systems, factories and data storage with the use of proper modelling approaches. Many manufacturing and planning activities can be captured by discrete-event simulation, while agent based simulation is a popular option for modeling human behaviors (Gilbert 2008).

The library objects should be scalable and adaptive to different problem scales. A typical product technology group may contain hundreds of base products, which are disaggregated to intermediate products and eventually to stock keeping units along the manufacturing process. The objects should be easy to modify to cope with different loading levels and process complexity.

A selection of key performance indicators (KPIs) is employed at each production or planning stage to evaluate and compare different proposals. Each object should measure and collect the relevant data for the calculation of predefined KPIs. The gathered data should be stored effectively with proper format and order, and can be easily retrieved later during the result analysis. The performance measures of interest vary depending on the project objectives and different levels of decision making. In this paper, the KPIs are chosen based on the study by Gunasekarana et al. (2001).
It is important to note that enabling the connection of objects with external data sources escalates the modeling power of the library objects. On one hand, it alleviates the effort to input and output data to and from the simulation model. It allows users to vary the input parameters and collect the results without spending too much time to modify the source code of the model. On the other hand, it permits the automation of online simulation experiments by allowing inter-operability with enterprise resource planning (ERP) systems and manufacturing execution systems (MES).

The library objects need to be generic enough to address different use cases, embracing a certain degree of flexibility. Some functionalities and input information are set as compulsory, and others are optional. Features and functions of all objects should be well documented at each development stage to ensure the ease of use for both library users and developers. It has to be clearly indicated what is the required information type and whether a function is requisite or optional.

Lastly, the library should be software independent and portable to different simulation platforms. Theoretically, the simulation library can be developed using any programming language or software. The selection of simulation platform is based on its effectiveness and practicability in application.

3.2 Architecture of SCSC-SIMLIB

In this subsection, we discuss the architecture of the proposed library. We start by describing the simulation levels, which SCSC-SIMLIB addresses. Then, we briefly outline how material and information flows are represented.

3.2.1 Simulation Levels and System Modeling

We define the four levels of simulation used in semiconductor supply chains according to which the library is organized. These levels are – ranked from low to high level – equipment, single production site, company supply chain and end-to-end supply chain. While simulation models on the lowest level focus on single machines, the second level encompasses the entire factory’s shop-floor and includes production planning and control. The third level models the company supply chain as a whole including all subcontractors and silicon foundries in its full width. Lastly, the fourth level integrates all players of the semiconductor supply chain including tier one customers, end customers, primary suppliers and raw material suppliers. Each lower level is included within the next consecutive level. Each level comprehends a set of library objects that are pertinent for the defined scope. Note that the same component can be modeled as several objects, distinguished by the level of detail, to serve different modeling purposes. Table 1 summarizes the simulation levels.

<table>
<thead>
<tr>
<th>Level</th>
<th>Scope</th>
<th>Selected use cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 4</td>
<td>End-to-end supply chain</td>
<td>Study interactions between customers, manufacturers, suppliers</td>
</tr>
<tr>
<td>Level 3</td>
<td>Company supply chain</td>
<td>Optimize the design of the internal supply chain network</td>
</tr>
<tr>
<td>Level 2</td>
<td>Single production site</td>
<td>Predict the factory output and performance</td>
</tr>
<tr>
<td>Level 1</td>
<td>Equipment</td>
<td>Assess the performance of dispatching rules for a workcenter</td>
</tr>
</tbody>
</table>

In addition to organizing SCSC-SIMLIB according to the model scope, we distinguish objects with regard to their function. We define three categories of function: Plan, Make, and Master Data. Objects that belong to Plan refer to planning and control decisions. Make function refers to objects that execute the planning decisions and mimic the production processes. Master Data contains invariant objects, which serve as a reference of basic data for Plan and Make.
3.2.2 Modeling of Material and Information Flows

The communication between library objects in SCSC-SIMLIB is categorized according to material and information flows. Material flows refer to movement of entities in and between Make objects. The movements of material are determined by the planning and control decisions taken by Plan objects, which lead to information flows between Plan and Make objects. Note that interactions between Plan objects require information flows as well. Furthermore, invariant information is stored in Master Data objects and are provided to Plan and Make objects via information flows. Figure 1 shows the modeling of material and information flows in SCSC-SIMLIB.

![Diagram of Material and Information Flows in SCSC-SIMLIB](image)

Figure 1: Modeling of Material and Information Flows in SCSC-SIMLIB.

3.3 Selected Objects of SCSC-SIMLIB

3.3.1 Lot Object

Lot objects are generated and released to the production network when a new production order is executed. They flow through a series of Make objects and are discarded when their processing is completed. The lot objects change forms at different stages of the supply chain. For example, front-end lots are un-batched and re-batched to back-end lots at the die bank.

A lot object should carry the features of a production lot. For example, a front-end lot object is characterized by its lot size, lot type, number of dies per lot, priority and allocated fabrication route. In addition, it needs to record important information and statistics of this lot through the production, such as the number of good dies per lot, due date, planned start week, scheduled start date, real start date, and real completion date. This enables the analysis of performance measurements such as production yield and cycle time. Note that some data are mandatory to activate the lot object and enable basic functions, such as the lot type (i.e., front-end or back-end lot), while others can be optional based on the research problem. The features of back-end lots are very different from the front-end lots due to the nature of semiconductor production. As a matter of fact, the production unit in back-end is a chip, whereas that in front-end is a wafer. These differences should be reflected in the lot objects.
3.3.2 Factory Object

Similar to the lot object, the factory object is also distinguished between front-end and back-end. A front-end site receives material from the wafer storage or the preceding front-end site or supplier in the upstream supply chain and delivers the processed lots to the succeeding front-end site or the die bank. On the contrary, a back-end site receives material from the preceding back-end site or the die bank and delivers the processed lots to the succeeding back-end site or the distribution center. Both have similar structure and functionalities but different configurations of parameters and KPIs.

As a result of its complexity and significant role in the supply chain, the factory object is a set of several different sub-objects. The sub-objects belong to different categories (Make, Plan and Master Data) and they may contain further subsidiary objects. In the literature, a variety of full detailed models and reduced models are researched to depict the operations (cf. Section 2). For the purpose of universal application of the objects and the saving of modeling effort, a high-level aggregated representation of the manufacturing process is applied in this paper.

The production unit and local production planner are the main components of a factory object. Just as the name suggests, the production unit processes the lots with uncertainties in consideration. Meanwhile, it records important statistics and performs the calculation of manufacturing KPIs, such as capacity utilization, level of work in progress (WIP), throughput and flow factor (FF). Effective presentations of the KPIs are critical for interpreting the results and identifying potential problems and opportunities. The statistics can be accessed by other objects for further analysis. The factory object receives every week the production requests from an overarching planning logic on corporate level. The local production planner then schedules the production that is assigned to this factory according to the order priority and resource availability. Through the local production planner object, the weekly production orders are allocated to every weekday with the aim of smoothing the production and satisfying the requests.

3.3.3 Customer Object

A supply chain network is driven by customer demand. In the semiconductor industry, the demand picture usually changes over time. Customers send forecasts to reserve the supply capacity and place orders as time approaches the delivery date. The nearer to the delivery date, the more accurate shall be the demand picture. In addition, the ordering behavior of an organization or an individual changes with the perception of the market. A customer will interact with other elements in the market, such as his own end customers, competitors and suppliers. Over-ordering occurs when the market appears to be more promising than it later turns out to be. Customers also tend to over-order when they expect a shortage of supply. Under-ordering occurs in the opposite cases.

As such, a promising approach for building a realistic customer object is a model that not only takes into account the volatile demand forecasts but also incorporates customer ordering behavior, i.e., by means of agent based modelling. The customer object is able to model customers from different tiers, including the distributors, original equipment manufacturers as well as the end customers. Each customer object contains a number of inherent states to represent different customer behaviors. Historical demand patterns and experience of customer managers can be used to parameterize the customer objects.

3.3.4 Planning Object

The planning and control logics strive for fulfilling customer orders in an effective and efficient manner. Each object may model a particular planning function, such as matching demand with supply, sequencing the production orders and determining the optimal stock target. Interactions between planning objects and production objects create a feedback loop and allow constant update of the planning outcome.

The constrained planning object is one of the key planning objects, which is modeled to determine the production requests while considering demand and capacity. The information inputs are the demand signal from customer objects, the target safety stock level from the inventory control object and the
inventory information from the Make objects, including die banks, factories and distribution centers. The production has to cover forecasted demand, planned stock increase and order backlog. The production that starts at the current time will fulfill the demand in the future with the consideration of a planned cycle time. The planned cycle time is estimated with regard to the production flow of the considered product.

However, non-capacitated production requests may not be feasible due to the limited resources. For this reason, a feasible production plan is obtained by constraining demand with available capacity. When there is a competition for the same production resources, the capacity is allocated according to the order priority. The priority is decided based on the demand types. External demand from customer orders is given a higher priority compared to internal demand for fulfilling additional buffer stock increase.

In addition to determining the quantity and schedule of production requests, the planning logic is also able to assign the adequate supply chain to achieve the production target. Products with different process steps are likely to require different technologies and processing times. Depending on the required production technology and resource availability, these product families will follow different supply chains. The planning logic analyzes the product specifications and facility capabilities and compute the best suitable supply chain for each production order.

4 APPLICATION OF SCSC-SIMLIB

In this paper, a SCSC-SIMLIB is implemented using the Java based simulation software AnyLogic. The library objects and models can be exported as standalone Java applications, which are also executable on other simulation platforms. The library objects are organized and named according to the four levels of simulation and categories described in Subsection 3.2.1. In Master Data objects, we propose a set of default values to facilitate the modeling process so that users do not necessarily need to configure each parameter in the objects. Table 2 shows selected objects of a SCSC-SIMLIB implemented in AnyLogic.

<table>
<thead>
<tr>
<th>Plan</th>
<th>Make</th>
<th>Master Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 4</td>
<td>Market demand generator</td>
<td>Silicon foundry</td>
</tr>
<tr>
<td></td>
<td>Subcontractor</td>
<td>End customer</td>
</tr>
<tr>
<td>Level 3</td>
<td>CONWIP</td>
<td>Front-end site</td>
</tr>
<tr>
<td>Constrained planning</td>
<td>Back-end site</td>
<td>Finished product</td>
</tr>
<tr>
<td>Supply chain planner</td>
<td>Die bank</td>
<td></td>
</tr>
<tr>
<td>Order management</td>
<td>Distribution center</td>
<td></td>
</tr>
<tr>
<td>Level 2</td>
<td>Lot Release</td>
<td>Workcenter</td>
</tr>
<tr>
<td>Dispatching</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 1</td>
<td>Scheduling</td>
<td>Lot</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Machine</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Operator</td>
</tr>
</tbody>
</table>

With the library objects, we are able to describe a supply chain with a certain level of complexity. In the following section, we demonstrate the use of library objects by building a semiconductor supply chain model that evaluates the impact of demand forecasting policies on the performance of the supply chain.

For the purpose of fast implementation with a relatively good accuracy on an aggregated level, a compact model is developed to mimic the production process in the factory object. The processes in the factory are separated in two groups: One is value-added (production) and another is non-value-added (waiting). The production is represented as a time delay modeled by a single machine and the waiting is modeled as a queue in front of the machine. In every period of time, a certain number of lots is inserted
into the queue object as controlled by the Lot Release object. A lot remains in the queue until it can be processed by the machine. While entering the machine object, a process time is being assigned to the lot that is calculated from the raw process time of the product under consideration and an additional stochastic time-lag, which models the process uncertainty (e.g., machine downtime). The machine object is able to process several lots simultaneously. The capacity of the machine object is related to the assumed maximum throughput of the factory object in one period of time while considering a target flow factor of FF=2.5. The lot sequencing follows a general First-In-First-Out rule, but overtaking is allowed by the differences among the process times. The lots share the same resources and compete with each other for the production capacity. The more lots in the factory object, the longer the time of the lots within the factory object. The amount of WIP of different product types is recorded so that the logic units can make use of this information for production planning and control, and the decision makers can monitor the factory performance. The compact model is easy to implement, but rough assumptions are made on a few stochastic parameters to keep the model simple.

In the present example, the supply chain model includes five Make objects (a Lot Release, a Front-End factory, a Die Bank, a Back-End factory and a Distribution Center object), four Plan objects (two independent Customers, a Supply Chain Planner and a Constrained Planning object) and a Master Data object. The Master Data object contains default invariant information of two intermediate product types and their three corresponding finished product types. Upon each action, the Make and Plan objects retrieve the product-dependent production data from the Master Data object.

The two customers order the three types of finished products with different quantities. The supply chain planner then aggregates the orders depending on the product family groups and provides his
Yuan and Ponsignon

forecasted demand. Different policies can be implemented, which influence the forecasting decision taken by the supply chain planner. In this example, we perform two simulation experiments to show the impact of different demand forecasting policies on the utilization of back-end factory and its variance. In the first case, the planner simply forwards the actual orders as received from the customers. In the second case, the planner smooths the demand using a moving average over the last five periods. Figure 2 shows an overview of model.

The model can be built quickly by drag-and-drops of the pre-determined standard objects from the library to the workplace without much modeling effort. The user merely needs to connect the objects and set up the parameters if the default values are not applicable. Furthermore, the pre-defined graphical representations and animations are associated with the objects to visualize material movements and KPIs.

After having performed both experiments, we can compare the simulation outputs. When no demand smoothing is used, the average utilization is 70.7% with a standard deviation of 2.0%. When a moving average based demand smoothing is applied, the average utilization remains the same while the standard deviation decreases to 0.5%. As a consequence, the variation of facility utilization is about four times bigger when no demand smoothing is applied. This experiment suggests that fluctuations in the order picture increase the volatility on the shop-floor, and thus the supply chain performance, when the external demand is considered as-is for driving the supply chain network. Figure 3 shows a screenshot of the simulation results when the policy “moving average” is being used.

5 CONCLUSIONS AND FUTURE RESEARCH

A simulation library that provides pre-defined standard objects is desired to facilitate and standardize the supply chain modeling in the semiconductor industry. In this paper, the basic requirements and framework of the library are presented to guide the organization and implementation of a SCSC-SIMLIB.
Then, the characteristics of selected objects are discussed. We also built a functioning supply chain model using the library objects without much additional modeling effort. With the use of library objects, the modelers can save modeling time, reduce errors and enable more extensible and scalable models.

More time and effort need to be devoted to the library to bring it to full inter-operability and to complete the range of objects desired for all simulation levels and functions. Including agent features in some library objects, which model human decision and behavior characteristics, will be an interesting topic for future research. A platform for gathering user feedback is suggested, where users can share their experiences and feedbacks with both developers and other users. In addition, the library objects can be further developed and used as a reference model. We intend to make it available to the researcher and practitioner community in the future.

ACKNOWLEDGMENTS

The authors of the present paper would like to thank Hans Ehm and Stefan Heilmayer for their help to design the library as well as Sebastian Eirich and Andreas Melzer for their contribution, respectively, to the implementation of SCSC-SIMLIB in AnyLogic and the documentation of the library.

REFERENCES

Gilbert, N. 2008 *Agent-Based Models*, SAGE.
Yuan and Ponsignon


AUTHOR BIOGRAPHIES

JINGJING YUAN is currently enrolled in the Concurrent Double Masters Programme in Supply Chain Engineering at the Georgia Institute of Technology and the National University of Singapore. During various internships and industry projects, she gathered experience in the semiconductor, automotive and health care industry. Her research interests include network design and supply chain management problems in manufacturing, warehousing and transportation, with the use of effective modeling and simulation methods. Her email address is <yuan_jing_jing@hotmail.com>.

THOMAS PONSIGNON works as a Senior Staff Engineer in the Corporate Supply Chain organization of Infineon Technologies AG in Munich, Germany. He obtained master's degrees in Industrial Engineering from the EPF-Ecole d’Ingénieurs, Sceaux, France and the University of Applied Sciences, Munich, Germany and a Ph.D. in Mathematics and Computer Science from the University of Hagen, Germany. His research interests include production planning and simulation of semiconductor manufacturing networks. His email address is <Thomas.Ponsignon@infineon.com>.