Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

A DATA MODEL FOR PLANNING IN THE SEMICONDUCTOR SUPPLY CHAIN

Irfan Ovacik

Intel Corporation
5000 W. Chandler Blvd.
Chandler, AZ 85226, USA

ABSTRACT

A recent gathering of academic and industry researchers has identified the need for a reference model for planning and control in semiconductor supply chains. The purpose of this model is to provide a common language for researchers working on different aspects of modeling and analysis of the semiconductor manufacturing supply chain, facilitate better communication and provide a common starting point for performance assessment across different analysis approaches to the planning and control problems. This paper introduces a data model to advance the discussion of this reference model. The data model is generic in that it is not specific to semiconductor manufacturing, but has been used in practical settings to drive analysis and application development to serve several planning functions in a major semiconductor manufacturing company.

1 INTRODUCTION

The Dagstuhl Seminar on *Modeling and Analysis of Semiconductor Supply Chains* (Chien et al. 2016) brought together a number of researchers from academia and industry to discuss the current state of modeling and analysis in the semiconductor supply chain and to develop future research topics that are interesting both from a research and practice perspective. One of the expected outcomes of the seminar was "developing a significant draft of a conceptual reference model for planning and control of a supply chain in the semiconductor industry that can be used for analysis and performance assessment purposes and to foster a common understanding in the research community both in academia and industry". The discussions at the seminar on this topic are the main motivation for this paper. Based on more than 20 years experience building commercially-available planning solutions for a software company and in-house custom solutions for a major semiconductor company, we hope to share our experience and contribute to efforts to define a reference model for the planning and control problems in the semiconductor supply chain.

We begin by acknowledging that the components of the model described here are not new. There are many references to these components, such as a *Bill of Material*, across the vast literature on various supply chain topics. These components are also part of the language used in practical settings across multiple business processes, applications and tools used in managing the supply chain. The key contribution of this paper is in presenting a compact set of generic components which are sufficiently versatile to model supply chain planning and execution problems across a wide range of industries, including the various complexities encountered in the semiconductor supply chain.

While models of this nature are not new to practitioners, there has been very little effort to make them available in the open literature. All software solution providers have a version of this model which provides data to the algorithms embedded in their software. These models are considered proprietary and only available to paying customers. The same is true for any of the custom solutions used by many companies, whether built in-house or outsourced to another software development house. There are also a number of books and papers that describe similar models but focus on the facts that need to be represented by the model rather than the structure of the model, mostly in the context of the particular problem that the book or the paper is addressing. These tend to be restrictive in scope, either too prescriptive or too general to be of any use from a general supply chain modeling perspective. For example, Mönch, Fowler and Mason (2013) describe a number of components that are needed to model a semiconductor fabrication facility and parameters associated with those components in the context of using discrete-event simulation techniques in the face of stochastic decision problems. The same components and parameters would be relevant from an overall supply chain perspective and would apply to other techniques described in the same chapter. While the need for a model as described here is a key component of effective communication of models and comparison of approaches among the researchers, the incentives to write about such models are not necessarily present in the current environment whose focus is more on specific problems, either matching solution methods to new problems or developing new and improved solutions to problems previously defined in the literature.

The model described here is not specific to the semiconductor supply chain or any specific problem in the supply chain. It is analogous to a frame or skeleton using which a given supply chain can be represented by adding the relevant data. It is also not necessary to use or instantiate all its components . A fabless semiconductor company where all manufacturing is outsourced to external suppliers would require very little detail of the manufacturing operation, leaving the Capacity Planning and Production Planning and Scheduling tasks to the supplier. In contrast, if the company owns its own manufacturing facilities, it would need to model the manufacturing operations in much more detail to be able to capture the data to perform the required planning functions. A retail or distribution company would only model its distribution network and would not need to model any manufacturing facilities. Although these examples cover different types of supply chains, the components discussed in this paper can accommodate most, if not all, of these variations.

Given that the purpose of the paper is to provide a basic set of components with which supply chains can be described, we deliberately avoid prescribing what particular data elements will be needed to describe a semiconductor supply chain, but will give examples throughout the paper as we describe the components of the data model.

In the next section, we introduce the building blocks that make up the data model. We describe each entity and how each entity relates to the others in the data model. In Section 3, we discuss some of the practical applications of this model in research, and wrap up with conclusions and next steps in the final section.

2 BASIC BUILDING BLOCKS OF THE DATA MODEL

The data model consists of a number of core entities and a number of derived entities formed by combining core and other derived entities. Each entity is described by a type and a set of attributes as illustrated by Figure 1. Typically the set of attributes that define the entity is a function of the type of the entity. Each entity can then be associated with a set of facts or measures, which can be time-phased – values changing over the planning horizon - or static in nature where the data is either constant over the planning horizon or does not have a time component.

For example, a location can represent a factory, a work center within a factory, a warehouse or a storage location within a warehouse. A location with type *Warehouse* can have *Address* as one of its attributes whereas a location with type *Storage Location* can have an attribute *Parent Location* which could be populated with the warehouse location which contains the storage location and therefore inheriting *Address*.

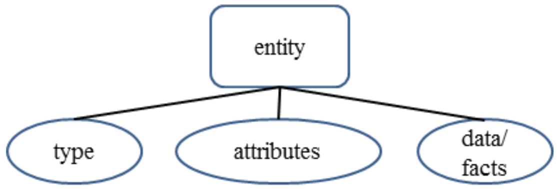

Figure 1: An entity in the data model.

Entities can be grouped as necessary and, in turn, be associated with attributes and facts. For example, all factories belonging to an outsourced supplier can be grouped together to form a *Location Group* with which data at the supplier level, such as minimum contractual volumes, can then be associated.

Entities can be combined to form new entities. A common example is the entity *Item-Location* which is a combination of the core entities *Item* and *Location* to represent the locations at which an item can exist in the supply chain. An item-location is defined when an item is consumed, manufactured, and stored at a given location. There are typically a number of supply chain related facts that can be characterized at the Item-Location level, such as inventory, safety stock targets, schedules, and projected inventory positions.

We now describe the key core and derived entities that make up the supply chain model and provide definitions and examples for each.

- Item: Raw materials, semi-finished products and finished products that exist across the supply chain.
- Location: Places where items can be consumed, manufactured or stored.
- Item-Location: Combinations of Items and Locations where a given item is either consumed, manufactured or stored.
- Bill of Material (BOM): Represents the relationship between a set of Items that are consumed (Input Items) and a set of Items that are manufactured (Output Items).

For supply chain modeling purposes, a Location is also associated with the BOM, since each transformation from a set of input items to a set of output items needs to take place at a Location. Multiple input items can be consumed to make an output item, and each input item can have one or more alternate items – this is the case in most industries that involve an assembly process where different manufactured or purchased components are brought together to create a new finished or semi-finished product. Multiple output items in a given Bill of Material are also possible.

Item, Location, Item-Location, and BOM describe the overall structure of the supply chain. Figure 2 shows a simple example illustrating how these entities relate to each other. In this case, *location1* represents a manufacturing location while *location2* and *location3* make up the distribution network. Items *item3* and *item4* are finished goods items that are produced using Items *item1* and *item2*. The finished good items are also stored in the distribution network at *location2* and *location3*.

We now define the entities that describe manufacturing and transportation processes.

- Operation: Represents each step that the products must undergo in order to transform the Input Items into the Output Items as defined by the BOM.
- Manufacturing Route: Represents the set and sequence of operations that products need to go through in order to transform the input items to the output items as defined by the BOM.
- Resource: Represents the machines, tools and operators needed to perform an Operation.
- Transportation Route: Represents the origin-destination location pairs on which Items need to travel.

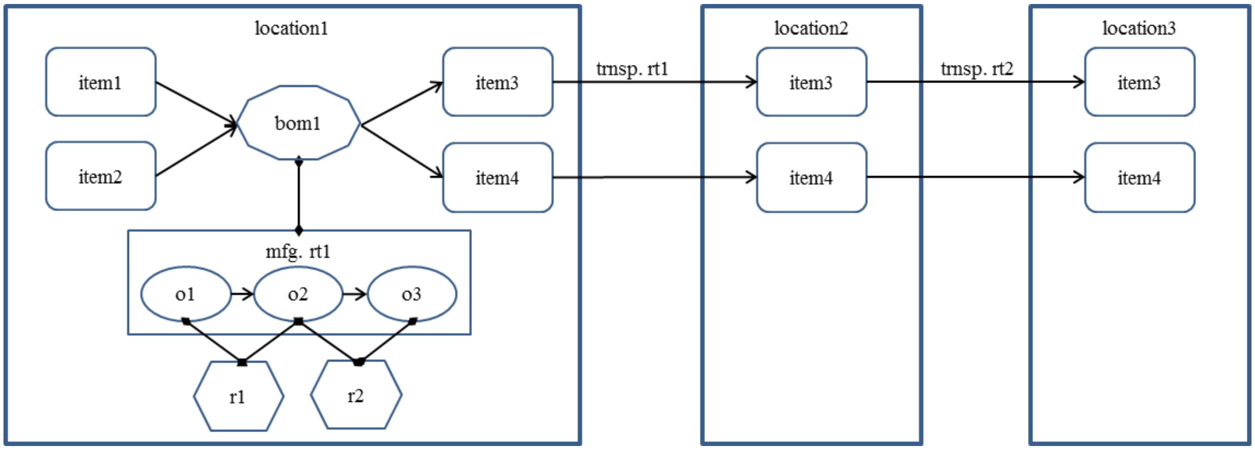

Figure 2: Relationships between entities.

Manufacturing Route, Operation and Resource describe the manufacturing process in the supply chain. In Figure 2, *mfg.rt1* represents the manufacturing process. It consists of three operations *o1*, *o2*, and *o3*. Operation *o1* uses Resource *r1*, Operation *o3* uses Resource *r2* whereas Operation *o2* uses both Resources *r1* and *r2*.

We would need to create an association between Item/BOM and Manufacturing Route, an association between the Manufacturing Route and Operations, and finally an association between the Operations and the Resources in order to be able to fully define the relationship of the products with the manufacturing process. Note that the reason for these entities, core or derived, is to allow us to describe the supply chain and the relationships between different elements in the supply chain while also adding relevant facts and data elements to those entities. For example, the number of a particular resource available and the amount of time each resource is available for manufacturing would be data elements attached to a Resource, but the Run Rate (how many pieces or how much volume the resource can process in a given time period) would be attached to a derived BOM/Item-Route-Operation-Resource entity, noting that the Run Rate could differ based on what product the Resource is processing at which Operation. Figure 3 shows the core and derived entities to model the manufacturing process.

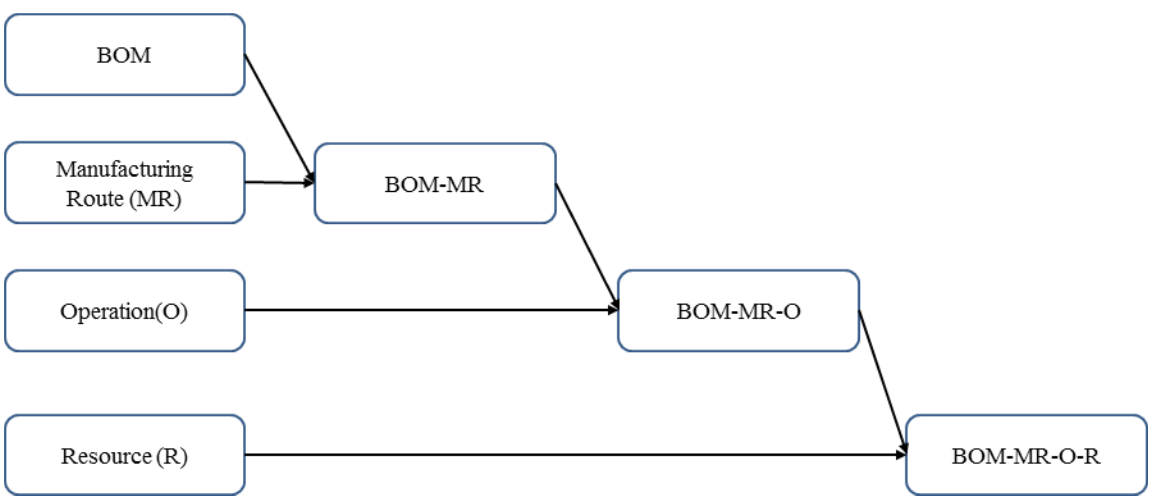

Figure 3: Entities for the manufacturing process.

Transportation Route is associated with the transportation process. In Figure 2, there are two Transportation Routes, *trnsp. rt1* and *trnsp. rt2*, representing the transportation lanes from *location1* to *location2* and from *location2* to *location3*, respectively. Item-Transportation Route is the other key entity in this process. Both *item3* and *item4* are associated with each of the Transportation Routes. In the case of transportation, any data elements that are independent of the item being transported would be associated with the Transportation Route. A typical example is the time it takes to transport a product from its origin to its destination. Other data elements that are dependent on both the item and the route would be captured under the Item-Transportation Route entity. A typical example would be the transportation cost of an item on a given route. Figure 4 shows the core and derived entities required to model the transportation process.

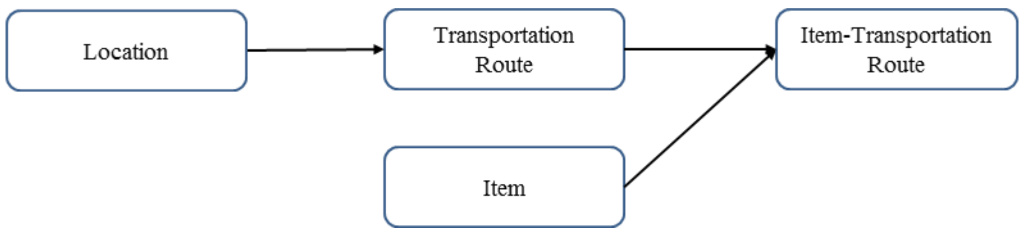

Figure 4: Entities for the transportation process.

Another key advantage of this model is its ability to allow abstractions. To support multiple planning processes in the supply chain across Capacity Planning, Inventory Planning, Master Planning and Production Planning and Scheduling, the model would need to be instantiated to a fairly high level of detail, especially to support Capacity Planning and Production Scheduling. However, this level of detail may not be necessary for Inventory Planning or Master Planning, but can be abstracted from the more detailed representation.

To illustrate this point, let's take two data elements, Throughput Time (TPT) and Yield, that are key to any supply chain analysis. For a detailed analysis involving Production Scheduling, we may choose to capture Yield at the operation level and the TPT as a function of the volume on a particular operation-resource combination and the run rate of the resource at that operation. For Master Planning or Inventory Planning, it may suffice to use aggregate values at the Manufacturing Route level with no need for a close link with the lower level values at the Operation level. In fact, it is often preferable to use aggregate, but conservative, values for Master Planning and Inventory Planning as opposed to deriving these from the lower level data used in the Production Scheduling process.

As indicated earlier, the data elements attached to an entity could be static - a single value that applies to the entire planning horizon - or time-phased to handle cases where its values change over time. The latter is more typical in practical situations where TPT and Yield values improve over time as the manufacturing processes mature. Whether static or time-phased, the values described so far are deterministic in that they do not reflect any uncertainty in the process or any impact due to the load on the system. The same idea can also be extended to those situations. Entities to which TPT and Yield are attached could still be used, but now to attach the parameters of statistical distribution, such as mean, variance, and the type of the distribution. Similarly, we could attach the parameters of a clearing function that describes the relationship between the load and TPT (Asmundsson, Rardin and Uzsoy 2009).

3 PRACTICAL APPLICATIONS IN RESEARCH

The entities described here provide the basic structure to describe a supply chain. The data attached to the entities, in turn, describes a particular supply chain. The data can be classified into three categories:

- Structural data, referring to the building block of the supply chain in terms of products that flow through the supply chain and the potential paths that they can take.

- Behavioral data, referring to how the supply chain behaves as products move through it, such the TPT and yield through a certain stage or the run rate of a machine at a given step.
- Transactional data, referring to the state of the supply chain in terms of where the products are (WIP, inventory), what the expected products movements are (plans/schedules), and what demand the supply chain is trying to satisfy.

Each planning problem in the supply chain requires a set of structural, behavioral and transactional data elements as its starting point. These data elements can span the entire supply chain data or just subsets of it depending on the nature of the planning problem. While there may be multiple ways to solve each problem, the core supply chain elements that are input to the problem remain the same. As long as the various approaches can consume the relevant data from the data model and put the results back into the data model, we have a way of comparing and contrasting different approaches to the same problem all using the same starting point.

For example, an Inventory Planning process would extract the distribution network (excluding the manufacturing network), statistical distributions that define the variability in demand, and the desired service levels from the data model. The output of the Inventory Planning analysis would also be stored as part of the data model in terms of safety stock targets at each point in the distribution network. There could be multiple ways of obtaining the output, spanning from simple heuristics to optimization techniques, to novel methods not yet explored. As long as the solution method can use the same input and produce the same output we would have a sound way of assessing the "goodness" of each approach. As an example, the results could feed a simulation model that models the entire supply chain (using the same data that the Inventory Planning process used) to assess which method actually performs better in terms of meeting customer needs and balancing inventory costs.

Note that in this case, the data model also provides a decoupling point between how the output is generated and how it is evaluated. This also opens up opportunities for sharing and collaboration where the simulation model can come from one set of researchers and the algorithms to calculate safety stocks from another group as long as the solutions developed by each groups can interface with the supply chain data model.

4 NEXT STEPS AND CONCLUSION

As mentioned in the introduction, the goal of this paper is to contribute to the discussion among researchers towards defining a reference model for the semiconductor supply chain. While the reference model definition is a super-set of the data model described here and will contain other components to describe the behavior, business processes and actors in the supply chain, we believe the data model provides a good starting point for modeling the planning and control problems. The data model described here has been used to model real problems in industry across various planning functions, so it provides a good foundation for any research activity in this domain. That said, we do not want to claim that it is complete and can support all planning and control problems in the supply chain. To be of any value to the research community, it would need to be exercised and improved over time while keeping its basic principles intact. Most of the next steps would involve proving that the model is robust enough to support the data that is available to the research community and promoting its use and reuse.

One area of future work would be to start looking at the problems that researchers are actively working on and start representing the test data that they are using in terms of the model described here. This would create a set of sample data sets that can be leveraged by other researchers working the same or similar problems.

There are also a number of datasets in the public domain such as the MIMAC datasets (Fowler and Robinson 1995) originating from industry. Another future activity would be to analyze these datasets and map them to this data model.

To conclude, we believe that this paper provides value by addressing a key gap in the academic and practitioner community. The common data model described here saws the seeds that could evolve into a common language among the researchers, allowing them to effectively collaborate by sharing data and effectively compare and contrast solution approaches by using a common starting point.

ACKNOWLEDGEMENTS

I would like to thank the organizers of Dagstuhl Seminar on Modeling and Analysis of Semiconductor Supply Chains (2016) which provided the motivation for this paper. I would also like to thank Professor Reha Uzsoy for the initial review of this paper and suggestions for improvement.

REFERENCES

Asmundsson, J., R. L. Rardin, and R. Uzsoy. 2009. "Production Planning Models for Resources Subject to Congestion." *Naval Research Logistics* 56: 142-157.

Chien, C-F., H. Ehm, J. Fowler, and L. Mönch. 2016. "Modeling and Analysis of Semiconductor Supply Chains." Dagstuhl Seminar 16062. http://www.dagstuhl.de/en/program/calendar/semhp/?semnr=16062.

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

Mönch, L., J. W. Fowler, and S. J. Mason. 2013. "Modeling and Analysis Tools." In *Production Planning and Control for Semiconductor Wafer Fabrication Facilities*, 29-64. New York: Springer-Verlag.

AUTHOR BIOGRAPHY

IRFAN OVACIK is a Principal Engineer at Intel Corporation. He holds a PhD in Industrial Engineering from Purdue University. His professional interests span all aspects of supply chain management, specifically in the semiconductor manufacturing industry. His email address is irfan.m.ovacik@intel.com.