Chapter 2

Demand Fulfillment in Make-to-Stock Manufacturing

In this chapter, we provide a description of the current state of demand fulfillment in make-to-stock manufacturing. In order to do this, we start with an introduction of basic concepts as the customer order decoupling point. Subsequently, we discuss the structure of advanced planning systems and show the activities involved in an exemplary demand fulfillment process in make-to-stock manufacturing. Additionally, we explain the notion of available-to-promise.

2.1 Make-to-Stock and the Customer Order Decoupling Point

In the work of Fleischmann and Meyr (2004) it is shown that demand fulfillment is strongly related to the position of the customer order decoupling point (CODP). Therefore—before we explain demand fulfillment and advanced planning systems in more detail—we give a short explanation of the CODP. As seen in Fig. 2.1, the CODP divides the supply chain into forecast-driven and order-driven processes (Sharman, 1984, Hoekstra and Romme, 1992).

The CODP holds the inventory that is needed to hedge against forecast errors and replenishment uncertainty. The CODP plays a pivotal role in our analysis since many decisions upstream of the CODP are dependent on the available inventory and on future replenishment orders. For a more detailed discussion of the CODP concept and its impact in different production environments (including make-to-order (MTO), assemble-to-order (ATO) and
make-to-stock (MTS)) we refer to Fleischmann and Meyr (2004). Note that the supply chain may contain additional production processes downstream of the CODP, but this is not the case for MTS production, since the CODP holds already the final product.

2.2 Structure of Advanced Planning Systems

Tasks of demand management are well-established in make-to-stock manufacturing companies and are usually supported by modern information systems, in particular advanced planning systems. These systems provide support for the entire planning tasks along the supply chain, from long-term strategic decision making to short-term operational decisions. Rohde et al. (2000) classify the supply chain decisions in a two-dimensional matrix (Fig. 2.2). This matrix is vertically structured according to the planning horizon and horizontally structured according to the sequence of planning tasks—from upstream to downstream in the supply chain.

![Figure 2.2: Structure of Advanced Planning Systems (Rohde et al., 2000)](image)

The planning tasks in the supply chain extend from procurement of raw materials over production and distribution towards selling of final products. Since all these tasks are interrelated, a consecutive processing will not lead to optimal plans. For instance, selling of products can only be done with information about production and distribution in order to generate reliable due dates. On the other hand, the production task requires reliable demand forecasts in order to decide about lot-sizes and working times. To be able to support all these tasks, APS usually have several modules structured according to planning horizon and forecast accuracy. In order to cope with the interdependencies, the modules are organized in a hierarchical order. Long-term decisions based on
low forecast accuracy provide the limits of the lower tasks which are done on the basis of better information. In order to do so, the upper tasks anticipate decision making in lower levels (Schneeweiss, 2003, Sect. 2.1). From top to down, the decisions become more accurate. Fig. 2.2 illustrates that on the top level, the strategic network planning is responsible for coordinating the entire supply chain. One the levels below, the tasks become more specialized.

The planning tasks associated with this thesis are located in the lower-right corner and are supported from the so-called demand fulfillment module. In order to illustrate how demand fulfillment for make-to-stock manufacturing is supported by advanced planning systems, Fig. 2.3 illustrates the involved supply chain activities and information flow in an exemplary consumer goods MTS supply chain.

In the mid-term, the demand planning activity is responsible for generating aggregate demand forecasts which are handed over to the master planning activity. It receives deterministic demand forecasts and prices as inputs from the demand planning (Kilger and Wagner, 2008) and then determines the best combinations of sales, production and replenishment quantities and the corresponding inventories under given capacity constraints. The planning horizon usually ranges from a few weeks up to several months. Therefore, data accuracy is low and the planning can only be done on the basis of aggregated data, i.e., products and customers are aggregated to groups or classes, respectively. In order to be able to balance supply with demand, the master planning activity basically determines the extent of seasonal stocks and possible changes of working times. A detailed description of the master planning activity can be found in the work of Rohde and Wagner (2008).
The generated master plan is handed over to the short-term activities. As demand forecasts usually become more accurate on the short-term, the planned stocks (from the master plan) and the current inventory are matched with the short-term demand forecasts. The result of this planning step is the net demand, which serves as input for the lot-sizing and scheduling task. This step additionally considers the working time restrictions from the master plan and computes the lot-sizes of the anticipated production quantities. Note, as this is a make-to-stock environment, all planning tasks up to here are performed on the basis of forecasts.

The last activity—which is the essential one in the scope of this work—is demand fulfillment. This activity takes into account the planned production quantities and quotes incoming customer orders according to their desired delivery date. This activity is the only one considered so far that is not based on forecasts but on the actual customer orders. In traditional material resource planning systems (MRP), order quotes are generated on the basis of available inventory. If there is no stock on-hand, the orders are quoted against the production lead-time. Kilger and Meyr (2008, p. 182) give a simple example that illustrates the weakness of such an order promising mechanism: it is not guaranteed that capacity constraints are not violated and a feasible plan is generated. Therefore, modern APS make use of more sophisticated methods. The inner logic of these methods is based on the notion of available-to-promise (ATP). In the following section, we introduce this key functionality of modern APS demand fulfillment solutions.

2.3 Available-to-Promise

2.3.1 Definition

The notion of available-to-promise is strongly related to advanced planning systems, and, with the success of APS solutions in the past years, is facing increasing attention. Nevertheless, it is not a very new concept. Fischer (2001) and Kilger and Meyr (2008) mention the work of Schwendinger (1979) as the earliest reference to available-to-promise. Since there is already a detailed analysis and literature review of ATP (see Fischer, 2001), we will not cite any references before the year 2001 and put an emphasis on the years thereafter.

Kilger and Meyr (2008) define ATP as “... the current and future availability of supply and capacity that can be used to accept new customer orders”. A different definition of ATP is provided in the work of Ball et al. (2004), in which ATP is defined as “... a business function [...] directly linking customer orders with enterprise resources”. The commercial software vendor SAP defines ATP quantity as the “quantity available to MRP for new sales orders” and
2.3 AVAILABLE-TO-PROMISE

ATP check as a "... function used to check [...] if a product can be confirmed" (SAP Help Portal, 2008).

Throughout this work, we follow the definition of Kilger and Meyr (2008) when referring to the term ATP and see it as a quantity rather than a functionality. It is important to note that a positive ATP quantity does not mean that there is stock on-hand, because ATP takes also future available quantities into account that are still to be produced. In the work of Fischer (2001), four different functions associated with the notion of ATP are identified:

- Availability check of products and evaluation of alternative solutions
- Order confirmation and due date assignment
- Steps taken in case of temporary inability to deliver
- Due date monitoring and order repromising

In the first step, the ATP quantities are used to check the availability of products. In cases when the product is currently not available, the next time the product will be available can directly be derived from the ATP quantities (as future supplies are considered). The availability check allows for a direct confirmation of customer orders including the determination of due dates. These two functions are essential for this work. The other two functions involve more activities as demand fulfillment and are beyond the scope of this work. For example, in case of temporary inability to deliver, a re-planning of the production (lot-sizing and scheduling) has to be done.

2.3.2 Dimensions of ATP

As seen before, the master planning activity calculates aggregated plans structured according to certain categories, e.g., product or customer groups. As the master plan is generally the basis for calculating ATP quantities, ATP exhibits a similar structure. Kilger and Meyr (2008, p. 184) mention several alternative possibilities to structure ATP quantities, e.g., location, sourcing type, region, market etc. but refer to product, time and customer as the most important ones. In the following, we give a short description of these three dimensions.

The level of detail of the ATP quantities in the product dimension corresponds to the location of the CODP. In case of MTO/ATO, products are customer specific and production/assembly is driven by incoming customer orders. The current and future supply and capacity that is taken into account to accept new customer orders is represented by the production capacity after the CODP and the inventory of semi-finished goods stored at the CODP. In the case of MTS manufacturing, the production is entirely based on forecasts. The current and future supply taken into account when the customer order arrives is hence represented by finished goods.
The master plan is generally created for a few weeks up to a few months. Accordingly, the granularity of the time dimension ranges from days to months, mostly depending on the forecast accuracy. The ATP quantities are therefore also structured according to certain time buckets, from which the customer orders are fulfilled.

The third important dimension is the customer dimension. Kilger and Meyr (2008) distinguish this dimension according to a supply- or demand constraint mode. Since the motivation of this thesis is driven by the problem to decide which customer to fulfill first when capacity is scarce, the supply constraint mode is the important one in this thesis. In this mode, not all orders can be fulfilled and the ATP functionality has to provide means to decide about customer priorities. Thus, APS assign customers to certain classes in order to have a customer hierarchy.

The allocation of the quantities from the master plan to customer hierarchies is usually done in APS on the basis of simple rules. Kilger and Meyr (2008, Sect. 9.4.3) distinguishes three important rules:

- **Rank based**: Allocation of quantities according to predefined ranks. The customer with the highest rank gets the forecasted quantities, the lower ranks the remaining quantities.
- **Per committed**: The quantities are allocated proportionally to the forecasts of the different customers. If customer A has forecasted 100, and B 200, then B gets twice as much as A irrespective of the actual available quantities.
- **Fixed split**: This rule is independent of the demand forecasts. Quantities are allocated according to a predefined fixed ratio, e.g., customer A gets 60%, B gets 40%.

The process of determining the available quantities is called allocation planning (AP).

To find a reliable due date for a customer order, it is searched through demand fulfillment alternatives in the mentioned dimensions. This means, e.g., to search in the time dimension, checking for ATP back- or forwards in time, in the product dimension, checking for substitute products, and in the customer dimension, checking for availability in other priority classes (Kilger and Meyr, 2008). This process is called order promising. Usually, simple rules are defined as search strategies for the different dimensions (Meyr et al., 2008a, Sect. 18.3.1). Profitability of different fulfillment alternatives is generally not taken into account during this search. However, recent software systems do not only consider available-to-promise (ATP) quantities (available inventory) or capable-to-promise (CTP) quantities (available capacity), but also follow a profitable-to-promise (PTP) logic that enables them to compare customer orders and fulfillment alternatives according to their priority.
are defined as “... the remaining capacity of the assembly lines, if this capacity is a potential bottleneck” (Fleischmann and Meyr, 2004). As they are only applicable in MTO and ATO production systems, CTP is not relevant here.