

HARTMUT STADTLER
CHRISTOPH KILGER
Editors

Supply Chain Management and Advanced Planning

Concepts, Models, Software
and Case Studies

Third Edition



Springer

Supply Chain Management and Advanced Planning

Hartmut Stadtler · Christoph Kilger
(Eds.)

Supply Chain Management and Advanced Planning

Concepts, Models, Software and Case Studies

Third Edition

With 173 Figures and 56 Tables

Professor Dr. Hartmut Stadtler
FG Produktion und Supply Chain Management
FB Rechts- und Wirtschaftswissenschaften
TU Darmstadt
Hochschulstraße 1
64289 Darmstadt
Germany
stadtler@bwl.tu-darmstadt.de

Dr. Christoph Kilger
j&m Management Consulting AG
Kaiserringforum
Willy-Brandt-Platz 5
68161 Mannheim
Germany
christoph.kilger@jnm.de

Cataloging-in-Publication Data
Library of Congress Control Number: 2004110194

ISBN 3-540-22065-8 Springer Berlin Heidelberg New York
ISBN 3-540-43450-X 2nd edition Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media
springeronline.com

© Springer Berlin · Heidelberg 2000, 2002, 2005
Printed in Germany

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Hardcover-Design: Erich Kirchner, Heidelberg

SPIN 11010463 42/3130-5 4 3 2 1 0 – Printed on acid-free paper

Preface

Hartmut Stadtler¹, Christoph Kilger²

¹ Darmstadt University of Technology, Department of Production & Supply Chain Management, Hochschulstraße 1, 64289 Darmstadt, Germany

² j & m Management Consulting AG, Kaiserringforum, Willy-Brandt-Platz 5, 68161 Mannheim, Germany

Preface to the Third Edition

Four years have passed since the first edition of our book – and still its readership is growing rapidly: You may even be able to buy a Chinese translation soon!

The field of Supply Chain Management (SCM) and Advanced Planning has evolved tremendously since the first edition was published in 2000. SCM concepts have conquered industry – most industry firms appointed supply chain managers and are “managing their supply chain”. Impressive improvements have resulted from the application of SCM concepts and the implementation of Advanced Planning Systems (APS). However, in the last years many SCM projects and APS implementations failed or at least did not fully meet expectations. Many firms are just “floating with the current” and are applying SCM concepts without considering all aspects and fully understanding the preconditions and consequences. This book provides comprehensive insights into the fundamentals of SCM and APS and practical guidance for their application.

What makes this book different from others in the field? Firstly, the material presented is based on our experiences gained by actually using and implementing APS. Furthermore, we have tried to extract the essence from three leading APS and to generalize the results – instead of merely reporting what is possible in a single APS. Secondly, this book is not just a collection of papers from researchers who have come together at a single conference and published the resultant conference proceedings. Instead we have structured the area of SCM and Advanced Planning into those topics relevant for turning APS successfully into practice. Then we have asked prominent researchers, experienced consultants and practitioners from large industry firms involved in SCM to join our group of authors. As a result, this edition (product) should be the most valuable source of knowledge for our readers (customers).

You may have observed that creating our team of authors has much in common with forming a supply chain in industrial practice. This story can be expanded even further: Several authors are also partners (contributors) in other supply chains (author groups). It is the task of the steering committee

(editors) to make our supply chain work and make it profitable for every partner. This model not only worked for the lifetime of a product's life cycle but also twice for its relaunch. We hope that our supply chain will stick together for some time in the future for the best of our customers – YOU!

What is new in this third edition, apart from the usual update of chapters?

- A section on strategic issues in SCM has been added as a subsection of Chap. 1.
- The contents of Chaps. 2 and 3 are restructured with a greater emphasis on Supply Chain Analysis.
- Latest issues and recommendations in Strategic Network Planning now have been prepared by two authors (Chap. 6).
- A new chapter has been added showing how to generate production and purchasing orders for uncritical items by utilizing the well-known MRP logic (Chap. 11).
- The chapters on the Definition of a Supply Chain Project (Chap. 15) and the Selection Process of an APS (Chap. 16) have been rewritten in light of new experiences and research results.
- Demand Fulfilment and ATP (Chap. 9) now is based on several APS and thus presents our findings in a more generalized form.
- There are two new case studies, one from the pharmaceutical industry (Chap. 19) and one from the chemical industry (Chap. 22). Also, all case studies now follow a common structure.

This edition would not have been possible without the advice from industry partners and software vendors. Many thanks to all of them for their most valuable help. This is also the last edition, where Jens Rohde has administered all the papers and prepared the files to be sent to the publisher. Thank you very much, Jens, for this great and perfect service and all the best for the future!

Hartmut Stadtler
Christoph Kilger

Darmstadt, April 2004
Mannheim, April 2004

Preface to the Second Edition

Success stimulates!

This also holds true when the first edition of a book is sold out quickly. So, we have created this second edition of our book with great enthusiasm.

Attentive readers of the first edition will have realized an obvious gap between the scope of Supply Chain Management (SCM), namely integrating legally separated companies along the supply chain and the focus of Advanced Planning Systems (APS) which, due to the principles of hierarchical planning, are best suited for coordinating intra-organizational flows. Now, collaborative planning is a new feature of APS which aims at bridging this gap. Consequently, this new topic is the most apparent addition to the second edition (Chap. 14).

But there are also many other additions which are the result of greater experience of the authors – both in industrial practice and research – as well as latest APS software developments. Examples of new materials included are:

- The different types of inventories and its analysis are presented in Chap. 2.
- The description of the SCOR-model and the supply chain typology have been enlarged and now form a separate chapter (Chap. 3).
- There is now a comparison of planning tasks and planning concepts for the consumer goods and computer assembly industry (Chap. 4).
- New developments in distribution and transport planning have been added (Chap. 12).
- Enterprise Application Integration is explained in Chap. 13.
- Chapter 17 now presents implementation issues of APS in greater detail.
- Some case studies have been updated and extended (Part IV).
- Rules of thumb have been introduced to allow users and consultants to better estimate and control computational times for solving their decision models (Part VI).

Like in the first edition we have concentrated on the three most popular APS because we have realized that keeping up-to-date with its latest developments is a very time consuming and challenging task.

SCM continues to be a top management theme, thus we expect our readers to profit from this update and wish them great success when implementing their SCM solution.

Many thanks to all who contributed to the first and second edition!

Hartmut Stadtler
Christoph Kilger

Darmstadt, January 2002
Mannheim, January 2002

Preface to the First Edition

During the late 80s and throughout the 90s information technology changed modern manufacturing organizations dramatically. Enterprise Resource Planning (ERP) systems became the major backbone technology for nearly every type of transaction. Customer orders, purchase orders, receipts, invoices etc. are maintained and processed by ERP systems provided by software vendors – like Baan, J. D. Edwards, Oracle, SAP AG and many more. ERP systems integrate many processes, even those that span multiple functional areas in an organization, and provide a consistent database for corporate wide data. By that ERP systems help to integrate internal processes in an organization.

Mid of the 90s it became apparent that focussing on the integration of internal processes alone does not lead to a drastic improvement of business performance. While ERP systems are supporting the *standard* business workflows, the biggest impact on business performance is created by *exceptions* and *variability*, e.g. customers order more than expected, suppliers deliver later than promised, production capacity is reduced by an unforeseen breakdown of equipment etc. The correct reaction to exceptions like these can save a lot of money and increase the service level and will help to improve sales and profits. Furthermore, state-of-the-art *planning procedures* – for planning sales, internal operations and supply from the vendors well in advance – reduce the amount of exceptional situations, helping to keep business in a standard mode of operation and turning out to be more profitable than constantly dealing with exceptional situations.

This functionality – powerful planning procedures and methodologies as well as quick reactions to exceptions and variability – is provided by *Advanced Planning Systems*. An Advanced Planning System (APS) exploits the consistent database and integrated standard workflows provided by ERP systems

to leverage high velocity in industry. Due to these recent developments, software vendors of APS boost a major breakthrough in enterprise wide planning and even collaborative planning between the partners along a supply chain.

Do APS hold the promises? What are the concepts underlying these new planning systems? How do APS and ERP systems interact, and how do APS supplement ERP systems? What are the current limits of APS and what is required to introduce an APS in a manufacturing organization successfully?

These were the questions we asked ourselves when we started our project on “Supply Chain Management and Advanced Planning” in summer 1998. Since we realized that there were many more interested in this new challenging field, the idea of publishing this book was born.

This book is the result of collaborative work done by members of four consultancy companies – aconis, j & m Management Consulting, KPMG and PRTM – and three universities – University of Augsburg, Darmstadt University of Technology and Georgia Institute of Technology. Our experiences stem from insights gained by utilizing, testing and implementing several modules of APS from i2 Technologies, J.D. Edwards and SAP AG. Tests and evaluations of modules have been conducted within several projects including students conducting their final thesis.

On the other hand, some members of the working group have been (and still are) involved in actual APS implementation projects in several European enterprises. The real-world experience gained from these projects has been merged with the results from the internal evaluation projects and provided valuable insights into the current performance of APS as well as guidelines how to setup and conduct an APS implementation project.

Since summer 1998 our group has spent much time gaining insights into this new fascinating field, working closely together with colleagues from academic research, vendors of APS and customers of APS vendors. However, we are aware of the fact that APS vendors are constantly improving their systems, that new areas come into focus – like supplier collaboration, Internet fulfilment, customer relationship management – and that, because of the speed of developments, a *final* documentation will not be possible. Hence, we decided to publish this book as a report on the current state of APS, based on our current knowledge and findings, covering the major principles and concepts underlying state-of-the-art APS.

This book will be a valuable source for managers and consultants alike, initiating and conducting projects aiming at introducing an APS in industry. Furthermore, it will help actual users of an APS to understand and broaden their view of how an APS really works. Also, students attending postgraduate courses in Supply Chain Management and related fields will profit from the material provided.

Many people have contributed to this book. In fact, it is a “Joint Venture” of the academic world and consultancy firms, both being at the forefront of APS technology. Hans Kühn gave valuable input to Chap. 2, especially to the

section on the SCOR-model. Daniel Fischer was involved in the writing of Chap. 9 on Demand Fulfilment and ATP. The ideas of the KPI profile and the Enabler-KPI-Value Network, described in Chap. 15, were strongly influenced by many discussions with Dr. Rupert Deger. Dr. Hans-Christian Humprecht and Christian Manß were so kind as to review our view of software modules of APS (Chap. 18). Dr. Uli Kalex was the main contributor to the design of the project solutions, on which the computer assembly case study (Chap. 21) and the semiconductor case study (Chap. 23) are based. Marja Blomqvist, Dr. Susanne Gröner, Bindu Kochugovindan, Helle Skott and Heinz Korbelius read parts of the book and helped to improve the style and contents. Furthermore, we profited a lot from several unnamed students who prepared their master thesis in the area of APS – most of them now being employed by companies implementing APS. Last but not least, we would like to mention Ulrich Höfling as well as the authors Jens Rohde and Christopher Sürie who took care of assembling the 24 chapters and preparing the index in a tireless effort throughout this project.

Many thanks to all!

We wish our readers a profitable reading and all the best for applying Advanced Planning Systems in practice successfully.

Hartmut Stadtler
Christoph Kilger

Darmstadt, June 2000
Mannheim, June 2000

Contents

Preface	V
<i>Hartmut Stadtler, Christoph Kilger</i>	

Introduction	1
---------------------------	---

Hartmut Stadtler

References	5
-------------------------	---

Part I. Basics of Supply Chain Management

1 Supply Chain Management – An Overview	9
--	---

Hartmut Stadtler

1.1 Definitions	9
-----------------------	---

1.2 Building Blocks	11
---------------------------	----

1.3 Origins	24
-------------------	----

References	33
------------------	----

2 Supply Chain Analysis	37
--------------------------------------	----

Christopher Sürie, Michael Wagner

2.1 Motivation and Goals	37
--------------------------------	----

2.2 Process Modelling	39
-----------------------------	----

2.3 Performance Measurement	49
-----------------------------------	----

2.4 Inventory Analysis	56
------------------------------	----

References	62
------------------	----

3 Types of Supply Chains	65
---------------------------------------	----

Herbert Meyr, Hartmut Stadtler

3.1 Motivation and Basics	65
---------------------------------	----

3.2 Functional Attributes	66
---------------------------------	----

3.3 Structural Attributes	69
---------------------------------	----

3.4 Example for the Consumer Goods Industry	71
---	----

3.5 Example for Computer Assembly	76
---	----

References	80
------------------	----

4 Advanced Planning	81
----------------------------------	----

Bernhard Fleischmann, Herbert Meyr, Michael Wagner

4.1 What is Planning?	81
-----------------------------	----

4.2 Planning Tasks Along the Supply Chain	86
---	----

4.3 Examples of Type-Specific Planning Tasks and Planning Concepts	92
--	----

References	106
------------------	-----

Part II. Concepts of Advanced Planning Systems

5 Structure of Advanced Planning Systems	109
<i>Herbert Meyr, Michael Wagner, Jens Rohde</i>	
References	115
6 Strategic Network Planning	117
<i>Mark Goetschalckx, Bernhard Fleischmann</i>	
6.1 Components of the Strategic Network Design Problem	121
6.2 Review of Models in the Literature	127
6.3 Modelling Strategic Supply Chain Design	128
6.4 SNP Modules in Advanced Planning Systems	133
6.5 Conclusions	134
References	136
7 Demand Planning	139
<i>Michael Wagner</i>	
7.1 A Demand Planning Framework	139
7.2 Statistical Forecasting Techniques	143
7.3 Incorporation of Judgmental Factors	148
7.4 Additional Features	150
References	157
8 Master Planning	159
<i>Jens Rohde, Michael Wagner</i>	
8.1 The Decision Situation	160
8.2 Model Building	166
8.3 Generating a Plan	174
References	177
9 Demand Fulfilment and ATP	179
<i>Christoph Kilger, Lorenz Schneeweiss</i>	
9.1 Available-to-Promise (ATP)	180
9.2 Allocated ATP	186
9.3 Order Promising	191
References	195
10 Production Planning and Scheduling	197
<i>Hartmut Stadtler</i>	
10.1 Description of the Decision Situation	197
10.2 How to Proceed from a Model to a Production Schedule	198
10.3 Model Building	201
10.4 Updating Production Schedules	207
10.5 Number of Planning Levels and Limitations	209
References	213

11 Purchasing and Material Requirements Planning	215
<i>Hartmut Stadtler</i>	
11.1 Basics of Material Requirements Planning	215
11.2 Generation and Timing of Uncritical Orders	217
11.3 Quantity Discounts and Supplier Selection	222
References	226
12 Distribution and Transport Planning	229
<i>Bernhard Fleischmann</i>	
12.1 Planning Situations	229
12.2 Models	235
References	243
13 Coordination and Integration	245
<i>Jens Rohde</i>	
13.1 Coordination of APS Modules	245
13.2 Integration of APS	248
13.3 Supply Chain Event Management	256
References	257
14 Collaborative Planning	259
<i>Christoph Kilger, Boris Reuter</i>	
14.1 Introduction	260
14.2 Types of Collaborations	262
14.3 A Generic Collaboration Process	270
14.4 Software Support	276
References	278

Part III. Implementing Advanced Planning Systems

15 The Definition of a Supply Chain Project	281
<i>Christoph Kilger</i>	
15.1 Supply Chain Evaluation	283
15.2 Supply Chain Potential Analysis	291
15.3 Project Roadmap	297
References	300
16 The Selection Process	303
<i>Christoph Kilger, Ulrich Wetterauer</i>	
16.1 Creation of a Short List	304
16.2 APS Requirements	309
16.3 Implementation and Integration	311
16.4 Post-implementation Effort and Support Model	315
References	316

17 The Implementation Process	317
<i>Ulrich Wetterauer, Herbert Meyr</i>	
17.1 The APS Implementation Project	317
17.2 Modelling Phases of an APS-Project	332
References	337
<hr/>	
Part IV. Actual APS and Case Studies	
<hr/>	
18 Architecture of Selected APS	341
<i>Herbert Meyr, Jens Rohde, Michael Wagner, Ulrich Wetterauer</i>	
18.1 i2 Technologies – i2 Six.One	341
18.2 Peoplesoft – EnterpriseOne TM Supply Chain Planning.....	346
18.3 SAP – APO	350
References	353
19 SCM in a Pharmaceutical Company.....	355
<i>Frank Altrichter, Tanguy Caillet</i>	
19.1 Case Description	355
19.2 Objectives of Project	358
19.3 Planning Processes	361
19.4 Results and Lessons Learned	369
20 Food and Beverages	371
<i>Michael Wagner, Herbert Meyr</i>	
20.1 Case Description	371
20.2 Aim of the Project	377
20.3 Model Building in Peoplesoft Strategic Network Optimization ..	377
20.4 Implementing the Master Planning model	381
20.5 Results and Lessons Learned	387
References	388
21 Computer Assembly	389
<i>Christoph Kilger, Lorenz Schneeweiss</i>	
21.1 Description of the Computer Assembly Case	389
21.2 Scope and Objectives	393
21.3 Planning Processes in Detail	395
21.4 Summary and Lessons Learned	405
References	406
22 Demand Planning of Styrene Plastics	407
<i>Boris Reuter</i>	
22.1 Description of the Supply Chain	407
22.2 The Architecture of the Planning System	408
22.3 Model Building with SAP APO Demand Planning	410
22.4 The Demand Planning Process of the Styrene Plastics Division ..	415

22.5 Concluding Remarks	419
References	420
23 Semiconductor Manufacturing	421
<i>Lorenz Schneeweiss, Ulrich Wetterauer</i>	
23.1 Case Description	421
23.2 Objectives of Project	425
23.3 Model Building with the i2 Factory Planner	427
23.4 Lessons Learned	434
24 Scheduling of Synthetic Granulate	435
<i>Marco Richter, Volker Stockrahm</i>	
24.1 Case Description	435
24.2 Objectives	437
24.3 Modelling the Production Process in APO PP/DS	437
24.4 Planning Process	448
24.5 Results and Lessons Learned	449

Part V. Conclusions and Outlook

25 Conclusions and Outlook	455
<i>Hartmut Stadtler</i>	
References	458

Part VI. Supplement

26 Forecast Methods	461
<i>Herbert Meyr</i>	
26.1 Forecasting for Seasonality and Trend	461
26.2 Initialization of Trend and Seasonal Coefficients	467
References	471
27 Linear and Mixed Integer Programming	473
<i>Hartmut Stadtler</i>	
27.1 Linear Programming	473
27.2 Pure Integer and Mixed Integer Programming	477
27.3 Remarks and Recommendations	481
References	483
28 Genetic Algorithms	485
<i>Robert Klein</i>	
28.1 General Idea	485
28.2 Populations and Individuals	486
28.3 Evaluation and Selection of Individuals	488
28.4 Recombination and Mutation	490

XVI Contents

28.5 Conclusions	491
References	492
29 Constraint Programming	493
<i>Robert Klein</i>	
29.1 Overview and General Idea	493
29.2 Constraint Satisfaction Problems	494
29.3 Constraint Propagation	495
29.4 Search Algorithms	497
29.5 Concluding Remarks	498
References	499
Index	501
About Contributors	509

Introduction

Hartmut Stadtler

Darmstadt University of Technology, Department of Production & Supply Chain Management, Hochschulstraße 1, 64289 Darmstadt, Germany

Supply Chain Management – just another shortlived management philosophy? The gains that have been realized when adopting *Supply Chain Management* (SCM) and Advanced Planning are impressive:

- Hewlett-Packard cut deskjet printer supply costs by 25% with the help of inventory models analyzing the effect of different locations of inventories within its supply chain. This analysis convinced Hewlett-Packard to adopt a modular design and postponement for its deskjet printers (Lee and Billington, 1995).
- Campbell Soup reduced retailer inventories on average by 66% while maintaining or increasing average fill rates by improving forecasts and introducing simple inventory management rules (Cachon and Fisher, 1998).
- IBM applied its Asset Management Tool, consisting of analytical performance optimization and simulation, to its personal systems division, saving material costs and price-protection expenses of more than \$750 million in 1998 (Lin et al., 2000).
- BASF introduced vendor managed inventory with five key customers in its textile colours division. With the help of an Advanced Planning System it has been possible to raise the fill rate of its customers' inventory to almost 100%. Customers profited from eliminating safety stocks while it allowed BASF to generate less costly transportation and production schedules (Grupp, 1998).

These impressive gains show the potential of coordinating organizational units and integrating information flows and planning efforts along a supply chain.

Which manager can afford not to present such substantial gains in improving competitiveness? Nowadays, these gains cannot be achieved by one company alone, because companies have attempted to concentrate their business on those activities which they know best – their core competencies. As a result, all other activities have been outsourced to other firms, when possible. Consequently, the characteristics and the quality of a product or service sold to a customer largely depend on several firms involved in its creation. This brought about new challenges for the integration of legally separated firms and the coordination of materials, information and financial flows not experienced in this magnitude before. A new managerial philosophy was needed – Supply Chain Management.

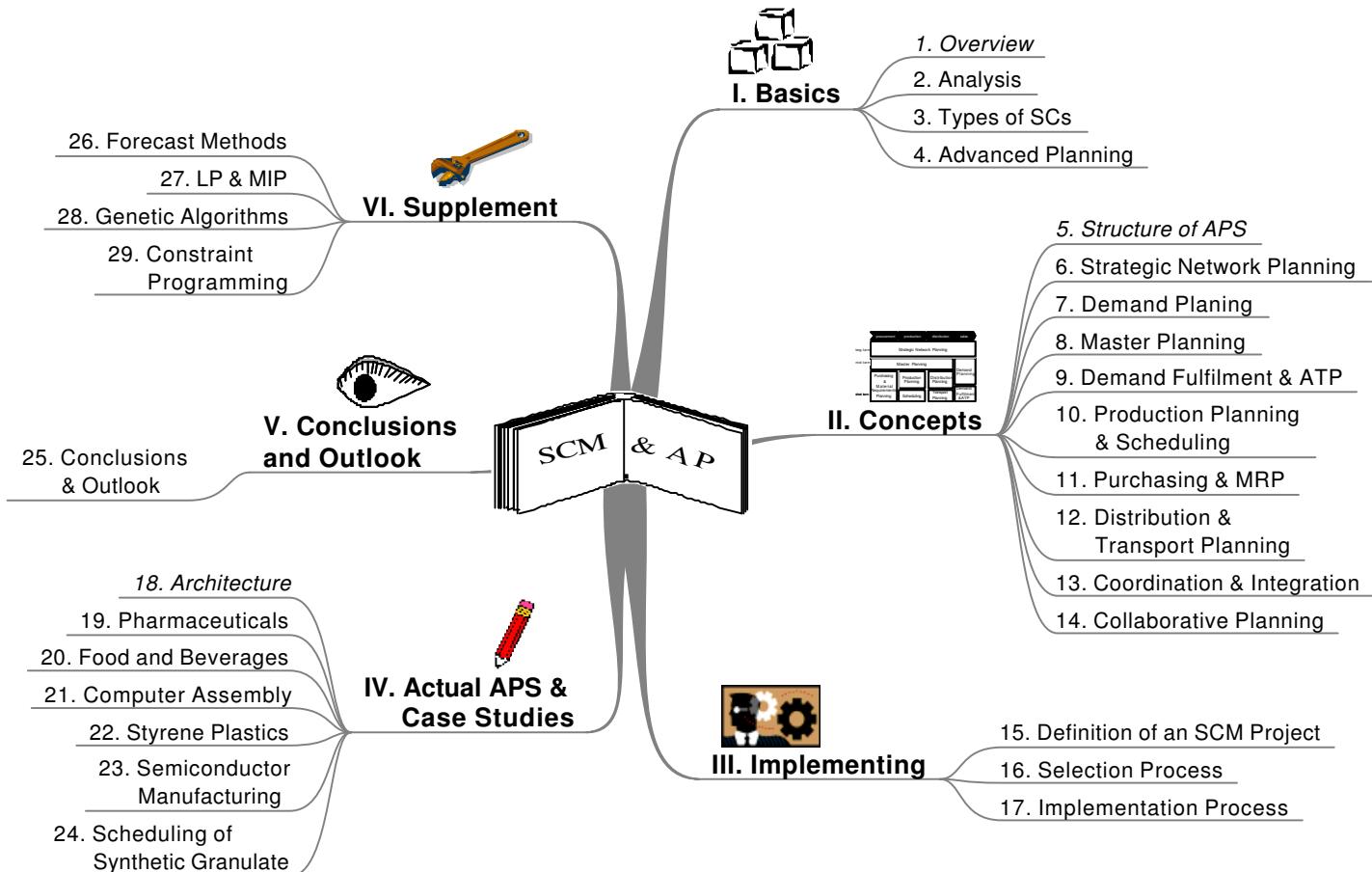


Fig. 1. Structure of this book

As with many management philosophies, impressive gains reported from pilot studies are promised. Often a few principles build the main body of such a new management philosophy. Since there are usually many more facets involved in managing a company successfully, some neglected factors may give rise to improvements achievable by the next management philosophy highlighted a few years later. Still, each management philosophy usually contains some building blocks that are advantageous and will survive over a longer period of time.

No great phantasies are needed to forecast that SCM will not be the ultimate managerial philosophy, although in our opinion it has many more facets than most of its predecessors. Since there are several facets to look at, SCM is difficult to grasp as a whole. While being aware of the broad area covered by SCM, this book will concentrate on recent developments in coordinating materials and information flows by means of the latest software products – called *Advanced Planning Systems* (APS). During the past ten years progress in information technology – like powerful database management systems – communication means – like *electronic data interchange* (EDI) via the Internet – as well as solution methods to solve large quantitative models – e. g. by mathematical programming – opened up new perspectives for planning and controlling flows along a supply chain. A customer's order, demand forecasts or market trends may be exploded into required activities and sent to all parties in the supply chain immediately. Accurate schedules are generated, which secure order fulfilment in time. Roughly speaking this is the task of APS. Unlike traditional *Enterprise Resource Planning* (ERP) these new systems try to find feasible, (near) optimal plans across the supply chain as a whole, while potential bottlenecks are considered explicitly.

It is our intention to provide insights into the principles and concepts underlying APS. In order to better understand and remember the structure of our book a mind-map has been created (Fig. 1). Part I of the book introduces the basics of SCM starting with a definition of SCM and its building blocks. The origins of SCM can be traced back into the fifties, when Forrester (1958) studied the dynamics of industrial production-distribution systems (see Chap. 1).

As a first step of introducing APS in industry it seems wise to document and analyze the current state of the supply chain and its elements (Chap. 2). A suitable tool for analyzing a supply chain are (key) performance indicators. They can provide valuable insights and guidance for setting targets for an SCM project. A well-known tool for analyzing a supply chain – the SCOR-model – provides a very valuable graphical representation with different levels of aggregation supplemented by performance indicators. Often, inventories at different locations in the supply chain are in the centre of interest of management. Hence, we discuss potential reasons for the existence of inventories.

Although APS are designed to be applicable for a number of industries, decision problems may vary widely. A typology of supply chains (Chap. 3)

will help the reader to identify which characteristics of a specific APS match the requirements of the supply chain at hand, and which do not, thereby guiding the selection process of an APS. Examples from industry illustrate different types of supply chains. Chapter 4 introduces the basics of advanced planning by applying the principles of hierarchical planning and explains the planning tasks along the supply chain by means of the supply chain planning (SCP) matrix.

Part II describes the general structure of APS (Chap. 5) and its modules in greater detail following the SCP matrix. Part II, however, will not only concentrate on functions and modelling features currently available in APS, but it will also describe ideas we regard to be good Advanced Planning and thus should be included in future releases of an APS. The presentation of concepts underlying these modules starts with strategic network planning (Chap. 6) followed by operational planning tasks for procurement, production and distribution. The quality of decision support provided by an APS largely depends on an adequate model of the elements of a supply chain, the algorithms used for its solution and the coordination of modules involved. Chapters 7 to 12 describe the many modelling features and mention solution procedures available to tackle different planning tasks without explicitly referring to specific APS. Although several modules have been identified, software vendors claim to offer a coherent, integrated software suite with close links to ERP systems. These linkages are the topic of Chap. 13.

In case a supply chain consists of several legally separated organizations, planning functions (usually) will not be controlled by a single, centralized APS. Instead, each partner will perform its own decentralized planning functions supported by an individual APS. Here, collaborative planning comes into play (Chap. 14) in order to agree on the exchange of data and the planning process. The overall objective is that the supply chain works in the most effective manner, i.e. ideally without interrupting the flow of information, materials and financial funds.

Part III is devoted to the implementation of an APS within a firm or supply chain. Obviously this requires a lot more than modelling. Often a consultancy company is hired to provide the expertise and manpower needed to introduce new, more efficient processes, to customize the APS and to train personnel. Hence, we describe the tasks necessary for introducing an SCM project (Chap. 15), the selection process of an APS (Chap. 16) and its implementation in industry (Chap. 17).

Recalling the general structure of APS (Chap. 5), Part IV now considers specific APS offered by i2 Technologies, Peoplesoft and SAP. It starts by pointing out differences in architecture (Chap. 18), followed by several case studies. Here we demonstrate how concepts and ideas outlined in the preceding chapters are applied to industrial practice with the help of actual APS. The first three case studies (Chaps. 19 to 21) are intended to provide a general idea of the planning problems facing specific industries and their resolution

by making use of an APS. The following three case studies concern specific modules of an APS. These comprise demand planning (Chap. 22), production planning (Chap. 23) and scheduling (Chap. 24). Here, special emphasis has been given to show how to model supply chain elements.

Part V sums up our experiences and gives an outlook of potential future developments.

Finally, a supplement (Part VI) provides a brief introduction to major algorithms used to solve the models mentioned in Parts II and IV and should enable the reader to better understand how APS work and where their limits are. Especially, forecast methods relate to Demand Planning (Chap. 26). Linear and mixed integer programming models are the solution methods needed if optimal master plans or distribution plans are looked for (Chap. 27). Last but not least, constraint programming and genetic algorithms constitute alternative solution engines within the scheduling module, where suitable sequences of jobs (orders) on multiple resources have to be generated (Chaps. 28 and 29).

References

- Cachon, G.; Fisher, M. (1998) *Campbell Soup's continuous replenishment program: Evaluation and enhanced inventory decision rules*, in: Lee, H. L.; Ng, S. M. (Eds.) Global supply chain and technology management, POMS series in technology and operations management, Vol. 1, Miami, Florida, 130–140
- Forrester, J. W. (1958) *Industrial dynamics: A major breakthrough for decision makers*, Harvard Business Review, Vol. 36, No. 4, 37–66
- Grupp, K. (1998) *Mit Supply Chain Management globale Transparenz in der Distribution*, PPS Management, Vol. 3, No. 2, 50–52
- Lee, H. L.; Billington, C. (1995) *The evolution of supply-chain-integration models in practice at Hewlett-Packard*, Interfaces, Vol. 25, No. 5, 42–63
- Lin, G.; Ettl, M.; Buckley, S.; Yao, D. D.; Naccarato, B. L.; Allan, R.; Kim, K.; Koenig, L. (2000) *Extended-enterprise supply-chain management at IBM personal systems group and other divisions*, Interfaces, Vol. 30, No. 1, 7–25

Part I

Basics of Supply Chain Management

1 Supply Chain Management – An Overview

Hartmut Stadtler

Darmstadt University of Technology, Department of Production & Supply Chain Management, Hochschulstraße 1, 64289 Darmstadt, Germany

What is the essence of *Supply Chain Management* (SCM)? How does it relate to Advanced Planning? In which sense are the underlying planning concepts “advanced”? What are the origins of SCM? These as well as related questions will be answered in this chapter.

1.1 Definitions

During the nineties several authors tried to put the essence of SCM into a single definition. Its constituents are

- the object of the management philosophy,
- the target group,
- the objective(s) and
- the broad means for achieving these objectives.

The object of SCM obviously is the *supply chain* which represents a “... network of organizations that are involved, through upstream and downstream linkages, in the different processes and activities that produce value in the form of products and services in the hands of the ultimate consumer” (Christopher, 1998, p. 15). In a broad sense a supply chain consists of two or more legally separated organizations, being linked by material, information and financial flows. These organizations may be firms producing parts, components and end products, logistic service providers and even the (ultimate) customer himself. So, the above definition of a supply chain also incorporates the target group – the ultimate customer.

As Fig. 1.1 shows, a network usually will not only focus on flows within a (single) chain, but will have to deal with divergent and convergent flows within a complex network resulting from many different customer orders to be handled in parallel. In order to ease complexity, a given organization may concentrate only on a portion of the overall supply chain. As an example, looking in the downstream direction the view of an organization may be limited by the customers of its customers while it ends with the suppliers of its suppliers in the upstream direction.

In a narrow sense the term supply chain is also applied to a large company with several sites often located in different countries. Coordinating material, information and financial flows for such a multinational company in an efficient manner is still a formidable task. Decision-making, however, should

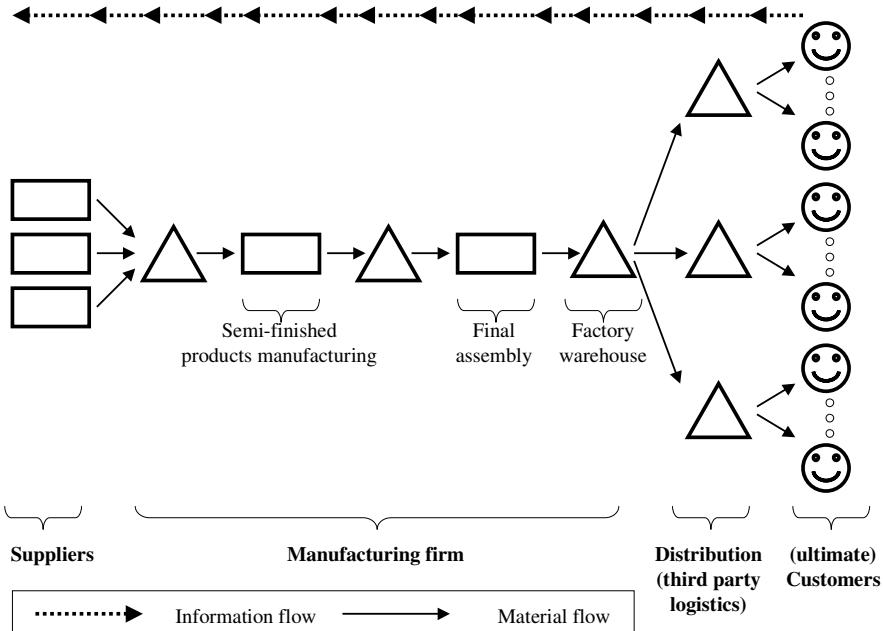


Fig. 1.1. Supply chain (example)

be easier, since these sites are part of one large organization with a single top management level. A supply chain in the broad sense is also called an *inter-organizational* supply chain, while the term *intra-organizational* relates to a supply chain in the narrow sense. Irrespective of this distinction, a close cooperation between the different functional units like marketing, production, procurement, logistics and finance is mandatory – a prerequisite being no matter of course in today's firms.

The objective governing all endeavours within a supply chain is seen as increasing competitiveness. This is because no single organizational unit now is solely responsible for the competitiveness of its products and services in the eyes of the ultimate customer, but the supply chain as a whole. Hence, competition has shifted from single companies to supply chains. Obviously, to convince an individual company to become a part of a supply chain requires a *win-win situation* for each participant in the long run, while this may not be the case for all entities in the short run. One generally accepted impediment for improving competitiveness is to provide superior customer service which will be discussed in greater detail below (Sect. 1.2). Alternatively, a firm may increase its competitiveness by fulfilling a prespecified, generally accepted customer service level at minimum costs.

There are two broad means for improving the competitiveness of a supply chain. One is a closer *integration* of the organizations involved and the other

is a better *coordination* of material, information and financial flows (Lee and Ng, 1998, p. 1). Overcoming organizational barriers, aligning strategies and speeding up flows along the supply chain are common subjects in this respect.

We are now able to define the term *Supply Chain Management* as the task of integrating organizational units along a supply chain and coordinating material, information and financial flows in order to fulfil (ultimate) customer demands with the aim of improving the competitiveness of a supply chain as a whole.

1.2 Building Blocks

The *House of SCM* (see Fig. 1.2) illustrates the many facets of SCM. The roof stands for the ultimate goal of SCM – competitiveness – customer service indicates the means. Competitiveness can be improved in many ways, e.g. by reducing costs, increasing flexibility with respect to changes in customer demands or by providing a superior quality of products and services.

The roof rests on two pillars representing the two main components of SCM, namely the integration of a network of organizations and the coordination of information, material and financial flows. The figure also shows that there are many disciplines that formed the foundations of SCM.

The two main components which incur some degree of novelty, will now be broken down into their building blocks. Firstly, forming a supply chain requires the *choice of suitable partners* for a mid-term partnership. Secondly, becoming an effective and successful *network organization*, consisting of legally separated organizations calls for actually practicing *inter-organizational collaboration*. Thirdly, for an inter-organizational supply chain, new concepts of *leadership* aligning strategies of the partners involved are important.

The coordination of flows along the supply chain can be executed efficiently by utilizing the latest developments in *information and communication technology*. These allow processes formerly executed manually to be automated. Above all, activities at the interface of two entities can be scrutinized, while duplicate activities (like keying in the data of a consignment) can be reduced to a single activity. *Process orientation* thus often incorporates a redesign followed by a standardization of the new process.

For executing customer orders, the availability of materials, personnel, machinery and tools has to be planned. Although production and distribution planning as well as purchasing have been in use for several decades, these mostly have been isolated and limited in scope. Coordinating plans over several sites and several legally separated organizations represents a new challenge that is taken up by *Advanced Planning (Systems)*.

Subsequently, we will describe the house of SCM in greater detail, starting with the roof, followed by its two pillars and ending with some references to its foundations.

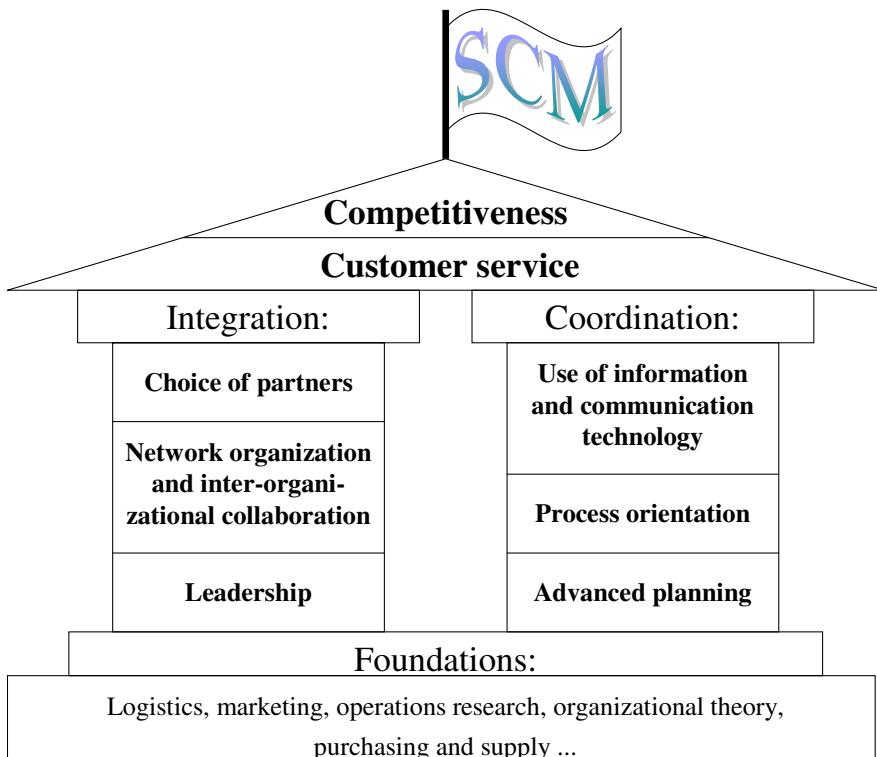


Fig. 1.2. House of SCM

1.2.1 Customer Service

Customer service is a multi-dimensional notion. According to a survey conducted by LaLonde and Zinszer (cited in Christopher (1998, pp. 39)) there are three elements of customer service:

- pre-transaction,
- transaction and
- post-transaction elements.

Some of these elements will be illustrated in the following text.

Pre-transactional elements relate to a company's activities preceding a contract. They concern customer access to information regarding the products and services a firm offers and the existence of an adequate link between organizations involved. Obviously, for standard products ordered routinely (like screws), an impersonal purchase via the Internet may be sufficient. Large projects, however, like a construction of a business building will require several, intense personal links between the organizations involved at different

levels of the hierarchy. Finally, flexibility to meet individual customer requirements may be an important element for qualifying for and winning an order.

Transactional elements are all those which contribute to order fulfilment in the eyes of a customer. The availability of products (from stock) may be one option. If a product or service has to be made on demand, order cycle times play an important role. During delivery times a customer may be provided with information on the current status and location of an order. The delivery of goods can include several additional services, like an introduction into the use of a product, its maintenance, etc.

Post-transactional elements mostly concern the service provided once the order is fulfilled. This includes elements like repairing or exchanging defective parts and maintenance, the way customer complaints are dealt with and product warranties (Christopher, 1998, pp. 41).

For measuring customer service and for setting targets, key performance indicators are used in practice, such as the maximum order lead-time, the portion of orders delivered within x days, the portion of orders without rejects or the fill rate (for details see Sect. 2.3 and Silver et al. (1998, pp. 243)).

If a certain level or standard of customer service has been agreed upon, it must be broken down so that each entity of the supply chain knows how to contribute to its achievement. Consider order lead-times offered to customers as an example (Fig. 1.3).

Assume a delivery time of nine days has to be offered to customers. Now, following each activity upstream in the supply chain with its expected lead-times for information and material flows, it becomes clear, where the *decoupling point* between the two options production-to-stock and production-to-order currently can be located. Since the actual lead-times for assembly totals 11 days, this would require to assemble-to-stock.

Stocks held at the decoupling point incur costs and increase overall throughput times. A decoupling point requires that no customized items or components have to be produced upstream. Ideally, items produced on stock have a large communality so that they can be used within several products. This will reduce the risk of holding the “wrong” stocks, if there is an unexpected shift in products’ demand.

If accumulated lead-times of customer specific parts exceed expected delivery times, the supply chain as a whole – perhaps including key customers – has to look for either reducing lead-times for material or for information flows (e.g. transferring orders by electronic means may save one day while an additional day may be saved by advanced scheduling techniques at the assembly plant, thereby allowing to *assemble-to-order* while suppliers *manufacture-to-stock*).

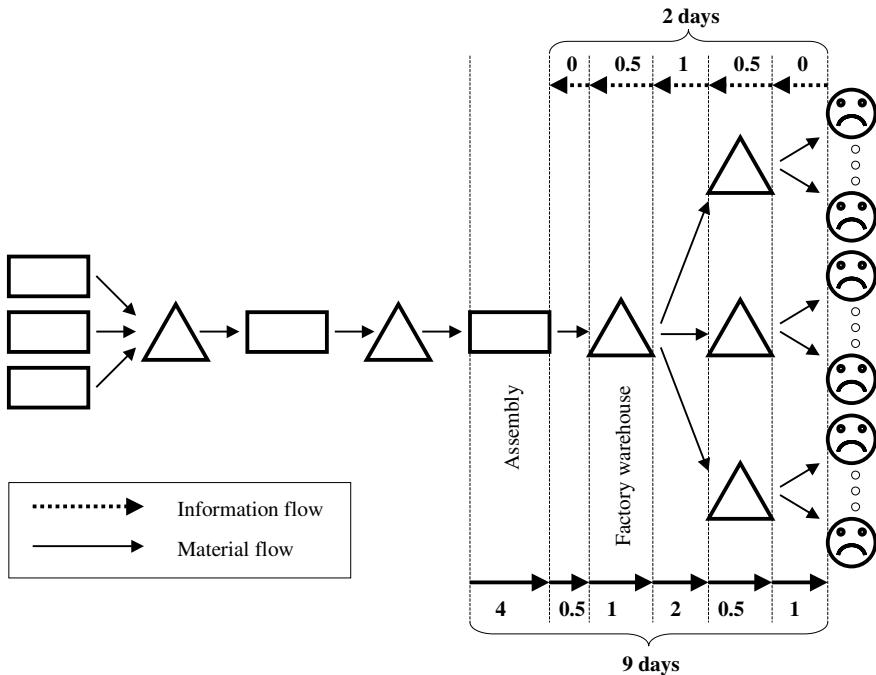


Fig. 1.3. Order lead-time and decoupling point

1.2.2 Integration

As has been stated above, a supply chain in the broad sense consists of several legally separated firms collaborating in the generation of a product or service with the aim of improving the competitiveness of a supply chain as a whole. Integration refers to the special building blocks that cause these firms to collaborate in the long term, namely

- choice of partners,
- network organization and inter-organizational collaboration,
- leadership.

The *choice of partners* starts with analyzing the activities associated with generating a product or service for a certain market segment (see also Chap. 2). Firstly, activities will be assigned to existing members of a supply chain, if these relate to their core competencies. Secondly, activities relating to standard products and services widely available on the market and with no potential of differentiation in the eyes of the ultimate customers, will be bought from outside the supply chain. Thirdly, for all remaining activities, a partner to join the supply chain has to be looked for in the course of a make-or-buy decision procedure (Schneider and Bauer, 1994).

Selection criteria should not be based solely on costs, but on the future potential of a partner to support the competitiveness of the supply chain. A suitable organizational culture and a commitment to contribute to the aims of the supply chain will be of great importance. A possible partner may bring in specialized know-how regarding a production process or know-how of products and their development. In case of a global supply chain, additional criteria have to be considered (like taxes, exchange rates, etc. (see Chap. 6)).

The assignment of activities to those members within the supply chain who can perform them best as well as the ability to adapt the structure of a supply chain quickly according to market needs are seen as a major advantage compared with traditional hierarchies.

From the perspective of organizational theory, supply chains are a special form of a *network organization*. They consist of loosely coupled, independent actors with equal rights. Their organizational structure is adapted dynamically according to the tasks to be performed and the aims of the network organization as a whole (Sydow, 1992; Hilse et al., 1999, p. 30). A supply chain may be regarded as a single (virtual) entity by its customers. The term virtual firm, however, is used for a network of firms collaborating only in the short term, sometimes only for fulfilling a single customer order.

Inter-organizational collaboration is a necessity for an effective supply chain. A supply chain is regarded as a cross between a pure market interaction and a hierarchy. It tries to combine the best features of the two. Ideally, each entity within a supply chain will concentrate on its core competencies and will be relieved from stringent decision procedures and administrative routines attributed to a large hierarchy. Information and know-how is shared openly among members. Competition among members along the supply chain is substituted by the commitment towards improving the competitiveness of the supply chain as a whole. A risk still remains, however, that collaboration is cancelled at some time. These features are assumed to enhance innovativeness and flexibility with respect to taking up new market trends (Burns and Stalker, 1961, pp. 121).

Although legally independent, entities within a supply chain are economically dependent on each other. Obviously, the structure of a supply chain will remain stable, only if there is a win-win situation for each member – at least in the long run. If this is not achieved in the short term by usual price mechanisms, compensation schemes must be looked for. To enforce the coherence of supply chain members several types of bonds may be used. These are

- “technical bonds which are related to the technologies employed by the firms,
- knowledge bonds related to the parties’ knowledge about their business,
- social bonds in the form of personal confidence,
- administrative bonds related to the administrative routines and procedures of the firms, and

- legal bonds in the form of contracts between the firms” (Håkansson and Johanson, 1990, p. 462).

An additional bond may be introduced by exchanging contributions to capital. Bonds must be practiced continuously to build up a certain degree of trust – the basis of a long-term partnership. In the case of a global supply chain special attention has to be paid to intercultural business communications (Ulijn and Strother, 1995).

Leadership, being the third building block of integration, is a delicate theme in light of the ideal of self-organizing, poly-centric actors forming a supply chain. At least some decisions should be made for the supply chain as a whole, like the cancellation of a partnership or the integration of a new partner. Similarly aligning strategies among partners may require some form of leadership (as an example see Rockhold et al. (1998)).

In practice, leadership may be executed either by a focal company or a steering committee. A *focal company* is usually a member having the largest (financial) power, the best know-how of products and processes or has the greatest share of values created during order fulfilment. In some cases, the focal company may also be the founder of a supply chain. For these reasons, decisions made by the focal company will be accepted by all members. On the other hand, a *steering committee* may be introduced, consisting of representatives of all members of a supply chain. The rules of decision-making – like the number of votes per member – are subject to negotiations.

Despite the advantages attributed to a supply chain, one should bear in mind that its structure is vulnerable – the exit of one partner may jeopardize the survival of the supply chain as a whole. Also, a member may run the risk of becoming unattractive and of being substituted by a competitor once his know-how has been dispensed within the supply chain.

Last but not least, the coordination of activities across organizations must not exceed comparable efforts within a hierarchy. In light of the latest developments in information and communication technology as well as software for planning material flows, this requirement has now been fulfilled to a large extent.

1.2.3 Coordination

The coordination of information, material and financial flows – the second main component of SCM – comprises three building blocks:

- utilization of information and communication technology,
- process orientation and
- advanced planning.

Advances in *information technology* (IT) made it possible to process information at different locations in the supply chain and thus enable the application

of advanced planning. Cheap and large storage devices allow for the storage and retrieval of historical mass data, such as past sales. These Data Warehouses may now be used for a better analysis of customer habits as well as for more precise demand forecasts. Graphical user interfaces allow users to access and manipulate data more easily.

Communication via electronic data interchange (EDI) can be established via private and public nets, the most popular being the Internet. Members within a supply chain can thus be informed instantaneously and cheaply. As an example, a sudden breakdown of a production-line can be distributed to all members of a supply chain concerned as a so-called alert.

Rigid standards formerly introduced for communication in special lines of businesses (like ODETTE in the automotive industry) are now being substituted by more flexible meta-languages (like the extensible markup language (XML)).

Communication links can be differentiated according to the parties involved (Corsten and Gössinger, 2001): business (B), consumer (C) or administration (A). Two communication links will be discussed here:

Business-to-business (B2B) communications allow companies to redesign processes, like that of purchasing. Manual tasks, e. g. placing an order for a standard item, can now be taken over by computer. It then controls the entire process, from transmitting the order, order acceptance by the supplier and order execution, until the consignment is received and checked. Finally, the amount payable is transferred to the supplier's account automatically. Automated purchasing allowed the Ford Motor Company to reduce its staff in the purchasing function drastically (Hammer and Champy, 1993, pp. 57). Other advantages stem from increased speed and reduced errors.

Furthermore, firms can make use of Internet based marketplaces, also called e-hubs (Kaplan and Sawhney, 2000). These marketplaces can be distinguished by four characteristics:

- The specificity of goods (either being manufacturing or operating inputs),
- the duration of the relationship (discriminated by systematic or spot sourcing),
- the pricing mechanism (with either fixed prices, e. g. an electronic catalogue, or price negotiations in the form of an auction) and
- the bias of an e-hub, which may favour either the seller, the buyer or take a neutral position.

Due to the global access to the Internet, not only strong competition and reduced purchasing prices may result, but also new sales opportunities. Note that market places play a role especially at the interface between two or more supply chains while the coordination of flows among different companies within a supply chain is supported by collaborative planning (see Chap. 14).

Business-to-consumer (B2C) communications aim at approaching the individual end user via the Internet. Several new challenges have to be addressed here, like a user-friendly access to information regarding products and services, securing safety of payments and finally the transport of goods or services to the customer. B2C opens up a new marketing channel to end users and offers a means for incorporating end users within a supply chain.

The second building block, *process orientation*, aims at coordinating all the activities involved in customer order fulfilment in the most efficient way. It starts with an analysis of the existing supply chain, the current allocation of activities to its members. Key performance indicators can reveal weaknesses, bottlenecks and waste within a supply chain, especially at the interface between its members. A comparison with best practices may support this effort (for more details see Chap. 2). As a result, some activities will be subject to improvement efforts, while some others may be reallocated. The building block “process orientation” has much in common with business process reengineering (Hammer and Champy, 1993); however, it will not necessarily result in a radical redesign. As Hammer (2001, p. 84) puts it, “streamlining cross-company processes is the next great frontier for reducing costs, enhancing quality, and speeding operations.”

Advanced planning – the third building block – incorporates long-term, mid-term and short-term planning levels. Software products – called *Advanced Planning Systems* – are now available to support these planning tasks. Although an Advanced Planning System (APS) is separated into several modules, effective information flows between these modules should make it a coherent software suite. Customizing these modules according to the specific needs of a supply chain requires specific skills, e.g. in systems and data modelling, data processing and solution methods.

APS do not substitute, but supplement existing *Enterprise Resource Planning* (ERP) systems. APS now take over the planning tasks, while an ERP system is still required as a transaction and execution system (for orders). The advantages of the new architecture have to be viewed in light of well-known deficiencies of traditional ERP systems with regard to planning (Drexel et al., 1994). In essence, an ERP system models the different planning tasks inadequately. Furthermore, these planning tasks are executed sequentially, without allowing for revisions to upper-level decisions. Some tasks, like bill-of-materials processing (BOMP), do not consider capacities at all. Furthermore, lead-times are used as a fixed input for the BOMP, even though it is common knowledge that lead-times are the result of planning. It is not surprising that users of ERP systems complain about long lead-times and many orders exceeding dead lines. Also, production planning and distribution planning are more or less separated systems. Last but not least, the focus of ERP systems has been a single firm, while APS have been designed also for inter-organizational supply chains.

Although separated in several modules, APS are intended to remedy the defects of ERP systems through a closer integration of modules, adequate modelling of bottleneck capacities, a hierarchical planning concept and the use of the latest algorithmic developments. Since planning is now executed in a computer's core storage, plans may be updated easily and continuously (e.g. in the case of a breakdown of a production line).

Planning now results in the capability to realize bottlenecks in advance and to make the best use of them. Alternative modes of operations may be evaluated, thus reducing costs and improving profits. Different scenarios of future developments can be planned for in order to identify a robust next step for the upcoming planning interval. Furthermore, it is no longer necessary to provide lead-time estimates as an input for planning. This should enable companies using APS to reduce planned lead-times drastically compared with those resulting from an ERP system.

A most favourable feature of APS is seen in its ability to check whether a (new) customer order with a given due date can be accepted (ATP, see Chap. 9). In case there are insufficient stocks at hand, it is even possible to generate a tentative schedule, inserting the new customer order into a current machine schedule where it fits best. Obviously, these new features allow a supply chain to comply better with accepted due dates, to become more flexible and to operate more economically.

We would like to add that proposals for a better integration of organizational units cannot be separated from the notion of the coordination of flows and vice versa. The choice of partners in a supply chain or the effectiveness of a postponement strategy can best be evaluated by advanced planning. On the other hand, the structure of a network organization sets up the frame for optimizing flows within a supply chain.

1.2.4 Relating SCM to Strategy

According to Porter (1998b, p. 55) a “strategy is the creation of a unique and valuable position, involving a different set of activities.” A company can obtain a unique and valuable position by either performing different activities than its rivals or by performing similar activities in different ways.

This can best be demonstrated by means of an example. The IKEA company has focused on the home furnishing needs of a specific customer group. The target group is price-sensitive and prepared to do its own pickup and delivery as well as the final assembly. IKEA's activities have been created according to these customer needs, which also have influenced the products' design and the structure of the SC. For instance, IKEA's showroom and warehouse are under one roof. A more precise description of the activities relating to IKEA's strategic position is given by the following activity-system map (see Fig. 1.4).

Here, activities, like “self assembly by customers”, are exhibited as well as the major links between dependent activities. For instance, “inhouse product

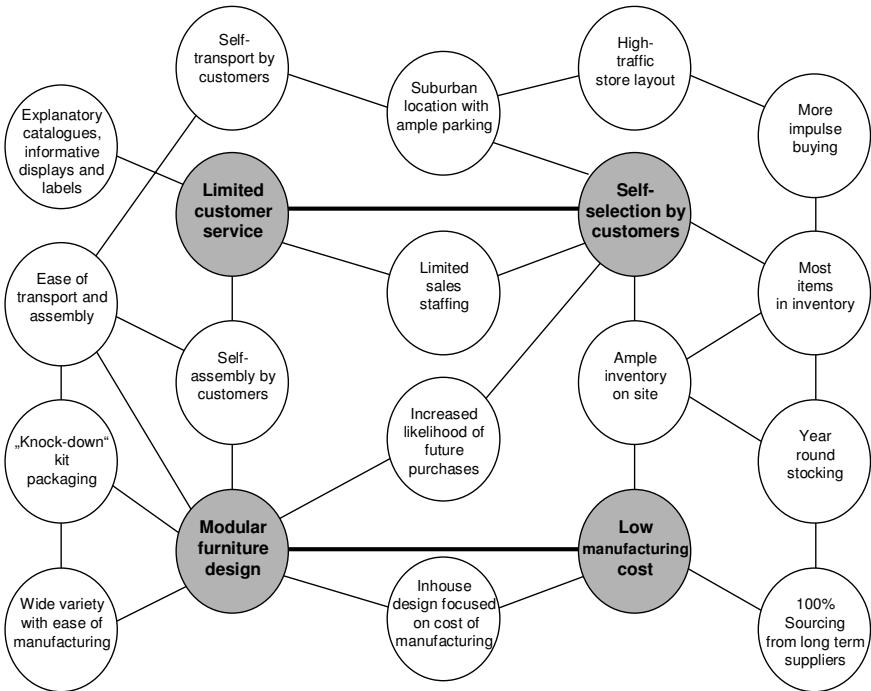


Fig. 1.4. Activity-systems map describing IKEA's strategic position (Porter, 1998b, p. 50)

design focused on cost of manufacturing” together with “100 % sourcing from long term suppliers” directly contribute to “low manufacturing cost”. Shaded activities represent high-order strategic themes. IKEA’s activity-system map also demonstrates that there are usually many interacting activities contributing to an overall strategy.

Another important part of strategy is the creation of fit among a SC’s activities. “The success of a strategy depends on doing many things well – not just a few – and integrating among them” (Porter, 1998b, p. 64). A given strategy will be successful only if all these activities will be aligned, or even better, if they reinforce each other.

The highest level of fit between all these activities – called optimization of effort (Porter, 1998b, p. 62) – is reached when there is coordination and information exchange across activities to eliminate redundancy and minimize wasted effort.

Now recall that SCM has been defined as integrating organizational units along a SC and coordinating activities related to information, material and financial flows. Hence, SCM is *not* a strategy on its own. Instead, SCM can and should be an integral part of a SC’s strategy as well as the individual partners’ business strategies. For example,

- SCM is an approach for generating competitive advantage by integrating organizational units and coordinating flows.
- SCM comprises specific activities, especially those concerning the order fulfilment process, which may be part of a SC's strategy.
- SCM utilizes specific tools best suited to reach the aspired level of fit among all strategic activities of a given SC.

There are a number of excellent textbooks (e. g. Aaker (2001)) on generating a strategy for an intra-organizational SC (company), which we will not review in detail here. In summary two main lines of thought prevail:

- the resource-based view and
- the market-based view.

A resource can be “... all assets, capabilities, organizational processes, firm attributes, information, knowledge, etc., controlled by a firm that enable the firm to conceive of and implement strategies that improve its efficiency and effectiveness” (Barney, 1991, p. 101). The focus here is on developing the resources' potentials.

Considering the market-based view (Porter, 1998a, pp. 3) an industry – usually consisting of several markets – is looked for, where the company can best exist against *competitive forces* given by

- industry competitors,
- potential entrants,
- power of buyers and suppliers or
- new product or service substitutes.

As one might expect the two views are not antagonistic but rather complement each other. For a deeper understanding the reader is referred to two case studies describing the generation of SC strategies in the apparel (Berry et al., 1999) and the lighting industry (Childerhouse et al., 2002).

Note, that creating and implementing a strategy within a single corporation may already be a difficult task, but it will be even more challenging in an inter-organizational SC. Namely, strategies of individual partners have to be aligned with the SC's overall strategy. In an inter-organizational SC further issues have to be addressed. Some of these, like the fit of companies, have already been discussed as part of the pillar “integration” of the House of SCM (see Sect. 1.2.2). Now, when formulating a SC-wide strategy, aspiration levels for the different issues of integration have to be added as well as (rough) paths for their achievement.

Even if contracts are binding SC partners, a SC is vulnerable and only created for a limited period of time. Hence, it seems wise to take into account and prepare “emergency plans” in case of separation. These may require

- good relations to alternative suppliers and customers currently not part of the SC, enabling a company (or SC) to become part of another SC and

- the installation of flexible (production) capacities that may also be used in another SC,
- engaging in several SCs to balance risks.

We would like to add that the discussion of strategies in the literature is dominated by the premise of pure competition. In the area of SCM, strategies for *collaboration* come into play. One of the difficulties is in finding a fair compromise of the sometimes diverging interests among SC partners. As an example consider the setting of fair *transfer prices* for products and services among SC partners. Given a fixed sales price the ultimate consumer is willing to pay for the end-product an increase of the transfer price granted to one SC member will incur a “loss” for the others. Furthermore, SC partners must be concerned that decentral investment decisions are made for the benefit of the SC as a whole, which may require specific subsidies, incentives or guarantees by the other SC partners.

Since generally applicable rules for calculating fair transfer prices or compensations are still missing (proposals for special situations can be found in Cachon and Lariviere (2004); Dudek (2004); Pfeiffer (1999)), negotiations come into play in practice. These become even more delicate if SC partners are reluctant to reveal their (true) cost structure and if the power of SC partners governs the outcome of negotiations.

Collaboration may also exist among competing SCs, e.g. in product distribution to consolidate consignments for the same destination (as in the food industry (Fleischmann, 2000)) or in combining demands for standard parts to increase the purchasing power (as in the automobile industry).

By now it should be clear that a favourable SC strategy always has to be specific in considering an SC’s potentials. Copying recipes drawn from *benchmarking studies* or an analysis of *success factors* (see e.g. Fröhlich and Westbrook (2001); Jayaram et al. (2004)) may be a good starting point but will not result in a unique and valuable position. In any case, a SC’s strategy will guide the specific design of building blocks best serving a SC’s needs (see Fig. 1.5).

For those interested in learning more about the first ideas and publications having influenced our current view of SCM, a section about its origins follows.

1.2.5 Foundations

For operating a supply chain successfully, many more ingredients are needed than those that have been reported in the literature in recent years in subjects like

- logistics and transportation,
- marketing,
- operations research,
- organizational behaviour, industrial organization and transaction cost economics,

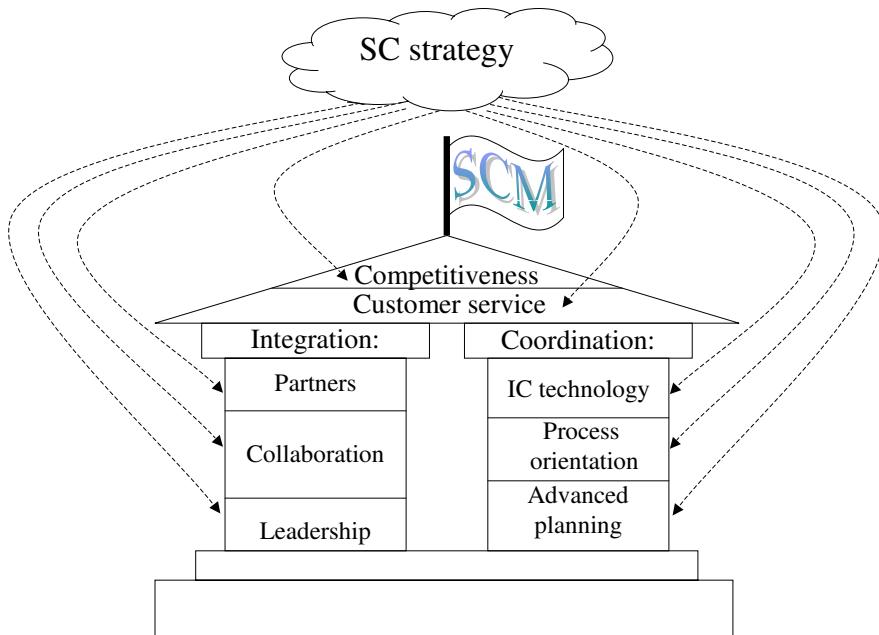


Fig. 1.5. The impact of a SC's strategy on the building blocks of SCM

- purchasing and supply,
- ...

to name only a few (for a complete list see Croom et al. (2000, p. 70)).

Certainly there are strong links between SCM and logistics, as can be observed when looking at the five principles of logistics thinking (Pfohl, 2004, pp. 20):

- Thinking in values and benefits,
- systems thinking,
- total cost thinking,
- service orientation and
- striving for efficiency.

Thinking in terms of values and benefits implies that it is the (ultimate customer) who assigns a value to a product. The value and benefit of a product can be improved with its availability when and where it is actually needed. Systems thinking requires examination of all entities involved in the process of generating a product or service simultaneously. Optimal solutions are aimed at the process as a whole, while being aware that optimal solutions for individual entities may turn out to be suboptimal. All activities are oriented towards a given service level. Service orientation is not limited towards the ultimate customer, but also applies to each entity receiving a product or

service from a supplier. Efficiency comprises several dimensions. The technological dimension requires the choice of processes, which results in a given output without wasting inputs. Furthermore, decision-making will be guided by economical goals, relating to current profits and future potentials. These two dimensions will be supplemented by a social and ecological dimension.

Another subject, operations research, has contributed to the model building and model solving required for coordinating flows along the supply chain. The basics of model building have already been developed in the sixties and seventies. However, only with the rise of powerful computers, large in-core storage devices and the availability of adequate solution methods, like Mathematical Programming and powerful meta-heuristics (e.g. genetic algorithms and tabu search), are these models now solvable with reasonable computational efforts (see Part VI).

Note that the vast body of literature on SCM has concentrated so far on the *integration* of inter-organizational supply chains. However, with regard to the *coordination* of flows, efforts still concentrate on intra-organizational supply chains. While it will not be too difficult to apply APS to an inter-organizational supply chain with a central planning unit, new challenges arise in decentralized planning (like the availability of data required for planning, coordinating plans, compensation schemes, etc.). Recalling that ERP systems only incorporate unconnected, insufficient analytical models (like for single level, uncapacitated lot-sizing), APS – even for intra-organizational supply chains – represent great progress. So, the term *advanced* in APS has to be evaluated in view of the insufficient decision support offered by ERP systems until now.

For those interested in learning more about the first ideas and publications that have influenced our current view of SCM, a section about its origins will follow.

1.3 Origins

The term SCM has been created by two consultants, Oliver and Webber, as early as 1982. The supply chain in their view lifts the mission of logistics to become a top management concern, since “...only top management can assure that conflicting functional objectives along the supply chain are reconciled and balanced ... and finally, that an integrated systems strategy that reduces the level of vulnerability is developed and implemented” (Oliver and Webber, 1992, p. 66). In their view, coordinating material, information and financial flows within a large multi-national firm is a challenging and rewarding task. Obviously, forming a supply chain out of a group of individual companies so that it acts like a single entity is even harder.

Research into the integration and coordination of different functional units started much earlier than the creation of the term SCM in 1982. These efforts can be traced back in such diverse fields as logistics, marketing, orga-

nizational theory, operations management and operations research. Selected focal contributions are briefly reviewed below without claiming completeness (for further information see Ganeshan et al. (1998)). These contributions are

- channel research (Alderson, 1957),
- collaboration and cooperation (Bowersox, 1969),
- location and control of inventories in production-distribution networks (Hanssmann, 1959),
- bullwhip effect in production-distribution systems (Forrester, 1958) and
- hierarchical production planning (Hax and Meal, 1975).

1.3.1 Channel Research

Alderson (1957) put forward *channel research* as a special field of marketing research. He had already argued that the principles of *postponement* require that “... changes in form and identity occur at the latest possible point in the marketing flow; and changes in inventory location occur at the latest possible point in time” (Alderson, 1957, p. 424). Postponement serves to reduce market risk, because the product will stay in an undifferentiated state as long as possible allowing to better cope with unexpected market shifts. Also postponement can reduce transportation costs, since products will be held back in the supply chain as far as possible (e. g. at the factory warehouse) until they are actually needed downstream (e. g. at a distribution centre) thereby reducing the need for the transport of goods between distribution centres in the case of a shortage of goods or an imbalance in the distribution of stocks. Thirdly, when examining the postponability of a (production) step, it might be discovered that it can be eliminated entirely, i. e. “... if a step is not performed prematurely, it may never have to be performed” (Alderson, 1957, p. 426). As an example, Alderson reported on the elimination of bagging wheat in sacks. Instead, a truck with an open box body had been chosen.

The three principles of postponement are still applied today. With regard to elimination, we can see that customers pick their goods directly from pallets thus eliminating the need for the retailer to put the goods on shelves. Another example are the customers of IKEA, who perform the assembly of furniture by themselves.

However, one should bear in mind that postponement in product differentiation requires that a product has already been designed for it, i. e. modifying a product to become customer specific should both be possible technically and economically later on. The capability of assessing the effects of postponement in a supply chain wide context is the achievement of advanced planning today. Thus, the different alternatives of postponement had been analyzed and simulated before Hewlett Packard introduced postponement successfully for its deskjet printer lines (Lee and Billington, 1995).

1.3.2 Collaboration and Coordination

Bowersox (1969) described the state of knowledge in marketing, physical distribution and systems thinking. There had already been an awareness that the individual objectives of the different functional units within a firm may counteract overall efficiency. For example (Bowersox, 1969, p. 64),

- manufacturing traditionally desires long production runs and the lowest procurement costs,
- marketing traditionally prefers finished goods inventory staging and broad assortments in forward markets,
- finance traditionally favours low inventories and
- physical distribution advocates total cost considerations relating to a firm's physical distribution mission.

Long production runs reduce the setup costs per product unit while resulting in higher inventory holding costs. Similarly end product inventories allow short delivery times, but increase inventory holding costs. Furthermore, raw materials and parts used up in the production of end products may no longer be used within other end products, thus limiting the flexibility to cope with shifts in end product demands (see postponement).

Furthermore, Bowersox criticized the fact that physical distribution systems mainly have been studied from the vantage point of vertically integrated organizations. "A more useful viewpoint is that physical distribution activities and related activities seldom terminate when product ownership transfer occurs." (Bowersox, 1969, p. 65). If the interface between two or more physical distribution systems is not properly defined and synchronized, this "...may well lead to excessive cost generation and customer service impairment" (Bowersox, 1969, p. 67).

Although arguing from the viewpoint of physical distribution, Bowersox had already advocated a need for intra-organizational as well as inter-organizational *cooperation and coordination*.

1.3.3 Location and Control of Inventories in Production-Distribution Networks

Hanssmann (1959) was the first to publish an analytical model of interacting inventories in a supply chain with three serial inventory locations. At each location a periodic review, order-up-to-level inventory system is used. There are positive lead-times, which are integer multiples of the review period. Customer demands are assumed to be normally distributed. Decision support is provided for two cases: the location of inventory, if only one single inventory location is allowed in the supply chain and the control of inventories if all three inventory locations may be used. Shortage costs and inventory holding costs are considered as well as revenues from sales which are assumed to be

a function of delivery time. As a solution method, dynamic programming is proposed.

The location and allocation of inventories in serial, convergent and divergent supply chains is still an important topic of research today.

1.3.4 Bullwhip Effect in Production-Distribution Systems

The *bullwhip effect* describes the increasing amplification of orders occurring within a supply chain the more one moves upstream. Surprisingly, this phenomenon also occurs even if end item demand is fairly stable. This phenomenon will be explained more deeply, since it is regarded as a classic of SCM.

The dynamic behaviour of industrial production-distribution systems has already been analyzed by Forrester (1958). The simplest system studied is a supply chain made of a retailer, a distribution centre, a factory warehouse and a production site (Fig. 1.6). Each entity can only make use of locally available information when making its ordering decisions for coping with demands. Another important feature are time delays between decision-making (e.g. ordering) and its realization (e.g. receipt of the corresponding shipment). These delays are indicated in Fig. 1.6 as numbers on top of respective arcs measured in weeks). The assumption is that a customer order comes in. Then the retailer requires one week to deliver it from stock. The lead-time between an incoming customer order until a decision to replenish inventory is made is three weeks (including processing the order), while order transmission to the distribution centre takes another half week. The distribution centre requires one week to process the order, while shipping the order to the retailer takes another week. Thus, five and a half weeks pass from an incoming customer order until the replenishment of the retailer's inventory (see Fig. 1.6: sum of bold numbers). Further lead-times for upstream entities can be derived in the same way from Fig. 1.6.

Forrester has shown the effects of a single, sudden 10% increase in retail sales on orders placed and inventory levels of each entity in the supply chain (see Fig. 1.7). He concludes (Forrester, 1961, p. 25) that “... orders at factory warehouse reach, at the 14th week, a peak of 34% above the previous December” and “... the factory output, delayed by a factory lead-time of six weeks, reaches a peak in the 21st week, an amount 45% above the previous December.” Obviously, these amplified fluctuations in ordering and inventory levels result in avoidable inventory and shortage costs and an unstable system behaviour. Although the time unit of one week seems outdated nowadays, replacing it by a day may reflect current practices better and will not disturb the structure of the model. These so-called information-feedback systems have been studied extensively with the help of a simulation package (DYNAMO).

In order to show the relevance of the work of Forrester on today's topics in SCM, we will add some newer findings here.

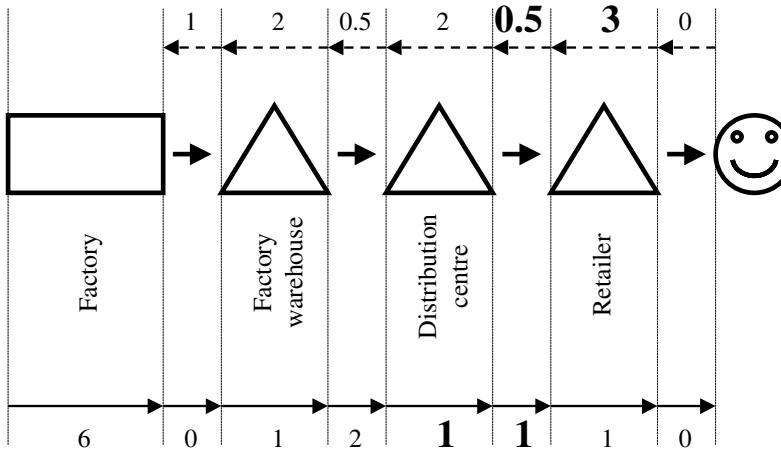


Fig. 1.6. Supply chain modelled by Forrester (1961, p. 22)

The introduction of the so-called *beer distribution game*, by Sterman (1989), has drawn great attention from researchers and practitioners alike to study the bullwhip effect again. Looking at an industrial production-distribution system from the perspective of bounded human rationality, Sterman studied the ordering behaviour of individuals possessing only isolated, local information.

In such an environment, where an individual's knowledge is limited to its current inventory status, the actual amount ordered by its direct successors in the supply chain and knowledge about its past performance, a human being tends to overreact by an amplification of orders placed. Even worse, amplification and phase lags of ordering increase steadily the more one moves upstream the supply chain. This has to be interpreted in light of a given, nearly stable end item demand with just one (large) increase in demand levels at an early period of the game.

This behaviour which is far from optimal for the total supply chain, has been observed in many independent repetitions of the beer distribution game as well as in industrial practice. Actually, the term bullwhip effect has been coined by managers at Procter & Gamble when examining the demand for Pampers disposable diapers (according to Lee et al., 1997).

Obviously, real world production-distribution systems are a lot more complex than those described above. However, examining behavioural patterns

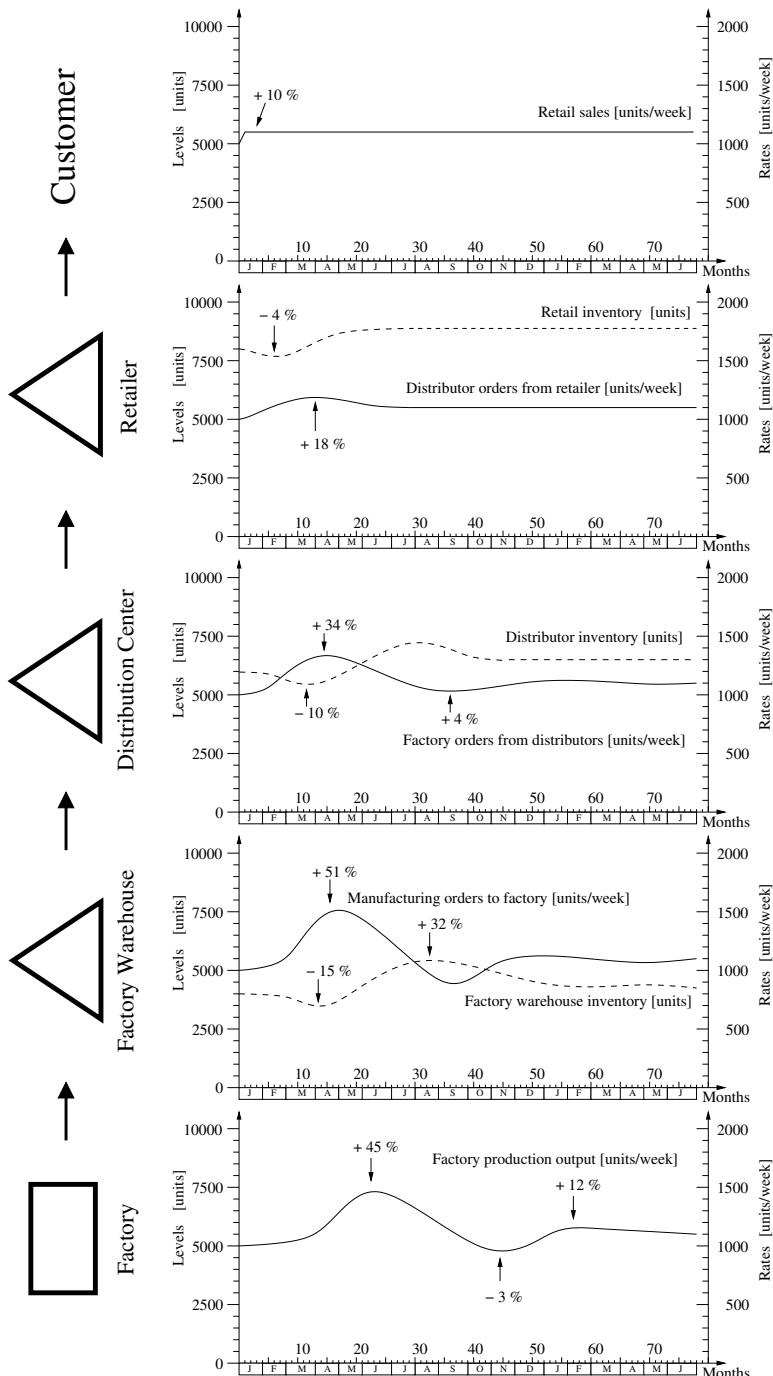


Fig. 1.7. The bullwhip effect (along the lines of Forrester (1961, p. 24))

and policies often adopted by local managers, may amplify fluctuations even further. Studying the causes of the bullwhip effect and its cures have become a very rich area of research in SCM. Recently, Lee et al. (1997) divided recommendations to counteract the bullwhip effect into four categories:

- avoid multiple demand forecast updates,
- break order batches,
- stabilize prices and
- eliminate gaming in shortage situations.

Avoiding *multiple demand forecasts* means that ordering decisions should always be based on ultimate customer demand and not on the ordering behaviour of an immediate downstream partner, since the ordering behaviour of an immediate downstream partner usually will show amplifications due to order batching and possible overreactions. With the advent of EDI and the capability to input sales made with the ultimate customer (point-of-sale (POS) data), accurate and timely data can be made available to each entity in the supply chain, thus also reducing the time-lag in the feedback system drastically. If ultimate customer demands are not available, even simple forecasting techniques (see Chap. 7) will prevent human overreactions and smooth demand forecasts.

In a more radical approach, one could change from decentralized decision-making to generating procurement plans centrally. Even the ultimate customer may be included in these procurement plans, as is the case in *vendor managed inventory* (VMI). Here the supply chain, however, has to bear the responsibility that the ultimate customer will not run out of stock. Finally, the downstream entity(s) could even be bypassed by executing sales directly with the ultimate customer (a well-known example are direct sales of Dell Computers).

Order batching is a common decision for cutting fixed costs incurred in placing an order. Ordering costs can be cut down drastically by using EDI for order transmission as well as a standardization of the (redesigned) ordering procedure. Transportation costs can be reduced if full truck loads are used. This should not, however, be achieved by increasing batch sizes, but rather by asking distributors to order assortments of different products simultaneously. Likewise, the use of third-party logistics companies helps making small batch replenishments economical by consolidating loads from multiple suppliers that are located near each other and thereby achieving economies of scale resulting from full truck loads. Similarly, a third-party logistics company may use assortments to full truckloads when delivering goods. This may give rise to cutting replenishment intervals drastically, resulting in less safety stocks needed without sacrificing service levels or increasing transportation costs.

Since marketing initiatives, which try to influence demands by wholesale price discounting, also contribute to the bullwhip effect, they should be abandoned. This understanding has moved companies to *stabilize prices* by guaranteeing their customers an every day low price.

The fourth category for counteracting the bullwhip effect intends to *eliminate gaming* in shortage situations. Here, gaming means that customers order additional, non-required amounts, since they expect to receive only a portion of outstanding orders due to a shortage situation. This behaviour can be influenced by introducing more stringent cancellation policies, accepting only orders in proportion to past sales records and sharing capacity and inventory information.

Many of the recommendations given above for counteracting the bullwhip effect profit from recent advances in communication technology and large database management systems containing accurate and timely information about the current and past states of each entity in the supply chain. Many time delays existing in production-distribution systems either are reduced drastically or even no longer exist, thus reducing problems encountered in feedback systems. Furthermore, to overcome cognitive limitations, a mathematical model of the supply chain may be generated and used to support the decision-making of individuals (Haehling von Lanzenauer and Pilz-Glombik, 2000). This research also indicates that an APS, with its modelling features and state-of-the-art solution procedures, can be a means to counteract the bullwhip effect.

1.3.5 Hierarchical Production Planning

Although detailed mathematical models have been proposed for production planning much earlier, Hax and Meal (1975) have shown how to build hierarchically coordinated, solvable models that provide effective decision support for the different decision-making levels within a hierarchical organization. Although first presented as a decision support system for a real world tire manufacturing firm, the versatility of the approach soon became clear. In brief, *hierarchical (production) planning* is based on the following five elements:

- decomposition and hierarchical structure,
- aggregation,
- hierarchical coordination,
- model building and
- model solving.

The overall decision problem is decomposed into two or more decision levels. Decisions to be made are assigned to each level so that the top level includes the most important, long-term decisions – i. e. those with the greatest impact on profitability and competitiveness. A separation into distinct decision levels is called *hierarchical* if for each level a single upper level can be identified which is allowed to set the frame within which decisions of the subordinated level have to take place (with the exception of the top level of the hierarchy). Note, there may be several separate decision units (e. g. production sites) within a given decision level coordinated by a single upper level.

Like decomposition, *aggregation* serves to reduce problem complexity. It also can diminish uncertainty (e. g. of demand forecasts). Aggregation is possible in three areas: time, products and resources. As an example, consider an upper level where time may be aggregated into time buckets of one week and only main end products are taken into account – irrespective of their variants, while available capacities at a production site are viewed as a rough maximum (weekly) output rate.

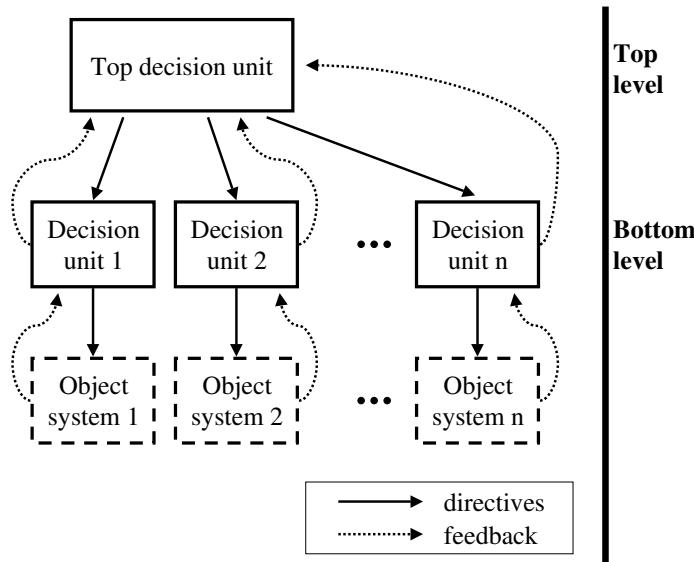


Fig. 1.8. Basic structure of a hierarchical planning system

Hierarchical coordination is achieved by directives and feedback. The most obvious directive is target setting by the upper level (e. g. setting a target inventory level for an end product at the planning horizon of the lower level). Another way is to provide prices for utilizing resources (e. g. a price for using additional personnel). A decision unit, on the other hand, may return a feedback to its upper level regarding the fulfilment of targets. These now allow the upper level to revise plans, to better coordinate lower-level decisions and to enable feasible plans at the lower level. These explanations are illustrated in Fig. 1.8. Here, the object system can be interpreted as the production process to be controlled.

For each decision unit a *model* is generated that adequately represents the decision situation and anticipates lower level reactions on possible directives. It also links targets set by the upper level to detailed decisions to be made at the decision unit considered. Thereby the upper level plan will be disag-

gregated. If a mathematical model is chosen, solvability has to be taken into account, too.

Finally, a suitable *solution procedure* has to be chosen for each model. Here, not only optimum seeking algorithms may be employed, but also manual procedures or group decision-making may be possible.

Hierarchical planning has attracted both researchers and practitioners alike. Thus, a large amount of knowledge has been accumulated so far (for more details see Schneeweiss, 1999). Since hierarchical planning represents an appealing approach in conquering complex decision problems, while incorporating the experience of human decision-makers at different levels of an organization, it is not surprising that today's APS are constructed along the principles of hierarchical planning (see Chap. 4 for more details).

References

- Aaker, D. A. (2001) *Developing business strategies*, 6th ed., New York
- Alderson, W. (1957) *Market behavior and executive action*, Homewood, Illinois
- Barney, J. B.; (1991) *Firm resources and sustained competitive advantage*, Journal of Management Studies, Vol. 17, 99–120
- Berry, W. L.; Hill, T.; Klompaker, J. E. (1999) *Aligning marketing and manufacturing strategies with the market*, International Journal of Production Research, Vol. 37, 3599–3618
- Bowersox, D. J. (1969) *Physical distribution development, current status, and potential*, Journal of Marketing, Vol. 33, January, 63–70
- Burns, T.; Stalker, G. M. (1961) *The management of innovation*, Oxford
- Cachon, G. P.; Lariviere, M. A. (2004) *Supply Chain Coordination with Revenue-Sharing Contracts: Strengths and Limitations*, Management Science (forthcoming)
- Childerhouse, P.; Aitken, J.; Towill, D. R. (2002) *Analysis and design of focused demand chains*, Journal of Operations Management, Vol. 20, 675–689
- Christopher, M. (1998) *Logistics and Supply Chain Management – Strategies for reducing cost and improving service*, 2nd ed., London et al.
- Corsten, H.; Gössinger, R. (2001) *Einführung in das Supply Chain Management*, München, Wien
- Croom, S.; Romano, P.; Giannakis, M. (2000) *Supply chain management: An analytical framework for critical literature review*, European Journal of Purchasing & Supply Management, Vol. 6, 67–83
- Drexl, A.; Fleischmann, B.; Günther, H.-O.; Stadtler, H.; Tempelmeier, H. (1994) *Konzeptionelle Grundlagen kapazitätsorientierter PPS-Systeme*, Zeitschrift für betriebswirtschaftliche Forschung, Vol. 46, 1022–1045
- Dudek, G. (2004) *Collaborative planning in supply chains – a negotiation-based approach*, Lecture notes in economics and mathematical systems, No. 533, Berlin et al.
- Fleischmann, B. (2000) *Kooperation von Herstellern in der Konsumgüterdistribution*, in: Sinz, E.J., et al. (Eds.), *Kooperation im Wettbewerb*, Wiesbaden, 167–186

- Forrester, J. W. (1958) *Industrial dynamics: A major breakthrough for decision makers*, Harvard Business Review, Vol. 36, July–August, 37–66
- Forrester, J. W. (1961) *Industrial dynamics*, New York, London
- Fröhlich, M. T.; Westbrook, R. (2001) *Arcs of integration: an international study of supply chain strategies*, Journal of Operations Management, Vol. 19, 185–200
- Ganeshan, R.; Jack, E.; Magazine, M. J.; Stephens, P. (1998) *A taxonomic review of supply chain management research*, in: Tayur, S.; Ganeshan, R.; Magazine, M. (Eds.) Quantitative models for supply chain management, Dordrecht, The Netherlands, 839–879
- Haehling von Lanzenauer, C.; Pilz-Glombik, K. (2000) *A supply chain optimization model for MIT's beer distribution game*, Zeitschrift für Betriebswirtschaft, 70 Jg., H. 1, 101–116
- Håkansson, H.; Johanson, J. (1990) *Formal and informal cooperation strategies in international networks*, in: Ford, D. (Ed.) Understanding business markets, London, 100–111
- Hammer, M.; Champy, J. (1993) *Reengineering the corporation*, New York
- Hammer, M. (2001) *The superefficient company*, Harvard Business Review, Vol. 79, September, 82–91
- Hanssmann, F. (1959) *Optimal inventory location and control in production and distribution networks*, Operations Research, Vol. 7, No. 4, 483–498
- Hax, A. C.; Meal, H. C. (1975) *Hierarchical integration of production planning and scheduling*, in: Geisler, M. A. (Ed.) Studies in Management Science, Vol. I, Logistics, Amsterdam, 53–69
- Hilse, H.; Götz, K.; Zapf, D. (1999) *Netzwerk contra Hierarchie: Die Abbildung organisationalstruktureller Widersprüche in einem neuartigen Potential für Rollenstress*, in: Götz, K. (Ed.) Führungskultur: Die organisationale Perspektive, München, 29–44
- Jayaram, J.; Kannan, V. R.; Tan, K. C. (2004) *The influence of initiators on supply chain value creation*, International Journal of Production Research, Vol. 42 (in print)
- Kaplan, S. M.; Sawhney, M. (2000) *E-hubs: The new B2B marketplaces*, Harvard Business Review, Vol. 78, May–June, 97–103
- Lee, H. L.; Billington, C. (1995) *The evolution of supply-chain-integration models in practice at Hewlett-Packard*, Interfaces, Vol. 25, No. 5, 42–63
- Lee, H. L.; Ng, S. M. (1998) *Preface to global supply chain and technology management*, in: Lee, H. L.; Ng, S. M. (Eds.) Global supply chain and technology management, POMS series in technology and operations management, Vol. 1, Miami, Florida, 1–3
- Lee, H. L.; Padmanabhan, V.; Whang, S. (1997) *The bullwhip effect in supply chains*, Sloan Management Review, Spring, Vol. 38, 93–102
- Oliver, R. K.; Webber, M. D. (1992) *Supply-chain management: Logistics catches up with strategy (reprint from Outlook (1982))*, in: Christopher, M. (Ed.) Logistics – The strategic issues, London et al., 63–75
- Pfeiffer, T. (1999) *Transfer pricing and decentralized dynamic lot-sizing in multistage, multiproduct production processes*, European Journal of Operational Research, Vol. 116, 319–330
- Pfohl, H.-C. (2004) *Logistiksysteme. Betriebswirtschaftliche Grundlagen*, 7th ed., Berlin et al.
- Porter, M. E. (1998a) *Competitive strategy, techniques for analyzing industries and competitors*, New York (originally published 1980)

- Porter, M. E. (1998b) *On competition*, Boston
- Rockhold, S.; Lee, H.; Hall, R. (1998) *Strategic alignment of a global supply chain for business success*, in: Lee, H. L.; Ng, S. M. (Eds.) Global supply chain and technology management, POMS series in technology and operations management, Vol. 1, Miami, Florida, 16–29
- Schneeweiss, C. (1999) *Hierarchies in distributed decision making*, Berlin et al.
- Schneider, D.; Bauer, C. (1994) *Re-Design der Wertkette durch make or buy*, Wiesbaden
- Silver, E. A.; Pyke, D. F.; Peterson, R. (1998) *Inventory Management and Production Planning and Scheduling*, 3rd ed., New York et al.
- Sterman, J. D. (1989) *Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment*, Management Science, Vol. 35., No. 3, 321–339
- Sydow, J. (1992) *Strategische Netzwerke*, Wiesbaden
- Ulijn, J. M.; Strother, J. B. (1995) *Communicating in business and technology: From psycholinguistic theory to international practice*, New York, Frankfurt

2 Supply Chain Analysis

Christopher Sürie¹ and Michael Wagner²

¹ Darmstadt University of Technology, Department of Production & Supply Chain Management, Hochschulstraße 1, 64289 Darmstadt, Germany

² Paul Hartmann AG, Corporate Materials Management, P.O.Box 1420, 89504 Heidenheim, Germany

Before starting an improvement process one has to have a clear picture of the structure of the existing supply chain and the way it works. Consequently a detailed *analysis* of operations and processes constituting the supply chain is necessary. Tools are needed that support an adequate description, modelling and evaluation of supply chains. In Sect. 2.1 several issues regarding supply chain analysis are discussed. Then, Sect. 2.2 presents modelling concepts and tools with a focus on those designed to analyse (supply chain) processes. The well known SCOR-model is introduced in this section. Building on these concepts (key) performance measures are presented in order to assess supply chain excellence (Sect. 2.3). Inventories are often built up at the interface between partners. As a seamless integration of partners is crucial to overall supply chain performance, a thorough analysis of these interfaces (i.e. inventories) is very important. Consequently, Sect. 2.4 gives an overview on inventories and introduces a standardized analysis methodology.

2.1 Motivation and Goals

An accurate analysis of the supply chain serves several purposes and is more a continuous task than a one time effort. In today's fast changing business environment, although a supply chain partnership is intended for a longer duration, supply chains keep evolving and changing to accommodate best to the customers' needs. In the beginning or when a specific supply chain is analysed for the first time in its entirety the result can be used as a starting point for improvement processes as well as a benchmark for further analyses. While the initial analysis itself often helps to identify potentials and opportunities it may well be used for target-setting, e.g. for APS implementation projects (see Chap. 15) to measure the benefit a successful implementation has provided. On the other hand, the supply chain analysis should evolve in parallel to the changes in the real world. In this way the associated performance measures keep track of the current state of the supply chain and may be used for supply chain controlling.

Many authors, researchers as well as practitioners, thought about concepts and frameworks as well as detailed metrics to assess supply chain performance (see e.g. Dreyer (2000), Lambert and Pohlen (2001) and Bullinger

et al. (2002)). In most concepts two fundamental interwoven tasks play an important role: *process modelling* and *performance measurement*. These two topics will be reviewed in detail in the following two sections, but beforehand some more general remarks are appropriate.

Supply chains differ in many attributes from each other (see also Chap. 3 for a detailed supply chain typology). A distinctive attribute often stressed in literature is the division into innovative product supply chains and functional product supply chains (see e.g. Fisher (1997) and Ramdas and Spekman (2000)). Innovative product supply chains are characterized by short product life cycles, unstable demands, but relatively high profit margins. This leads to a strong market orientation to match supply and demand as well as flexible supply chains to adapt quickly to market swings. On the contrary, functional product supply chains face a rather stable demand with long product life cycles, but rather low profit margins. These supply chains tend to focus on cost reductions of physical material flows and on value creating processes. Naturally, performance measures for both types of supply chains differ. Where time-to-market may be an important metric for innovative product supply chains, this metric does only have a minor impact when assessing performance of a functional product supply chain. Consequently, a supply chain analysis does not only have to capture the correct type of the supply chain, but should also reflect this in the performance measures to be evaluated. Supply chain's visions or strategic goals should also mirror these fundamental values.

Furthermore, a meaningful connection between the process model and the underlying real world as well as between the process model and the performance measures is of utmost importance. Although participating companies are often still organised according to functions, the analysis of supply chains has to be process oriented. Therefore, it is essential to identify those units that contribute to the joint output. These units are then linked to the supply chain processes as well as to the cost accounting systems of the individual companies. Therefore, they can provide the link between the financial performance of the supply chain partners and the non-financial performance metrics which may be used for the whole supply chain.

Finally, a holistic view on the supply chain needs to be kept. This is especially true here, because overall supply chain costs are not necessarily minimized, if each partner operates at his optimum given the constraints imposed by supply chain partners. This is not apparent and will therefore be illustrated by means of an example. Consider a supplier-customer relationship which is enhanced by a vendor managed inventory (VMI) implementation (see Chap. 29). At the customer's side the VMI implementation reduces costs yielding to a price reduction in the consumer market which is followed by a gain in market share for the product. Despite this success in the marketplace the supplier on the other hand may not be able to totally recover the costs he has taken off the shoulders of his customer. Although some cost components decreased (e.g. order processing costs and costs of forecasts),

these did not offset his increased inventory carrying costs. Summing up, although the supply chain as a whole profited from the VMI implementation, one of the partners was worse off. Therefore, when analyzing supply chains one needs to maintain such a holistic view, but simultaneously mechanisms need to be found to compensate those partners that do not profit directly from supply chain successes.

2.2 Process Modelling

2.2.1 Concepts and Tools

Supply chain management's process orientation has been stressed before and since Porter's introduction of the *value chain* a paradigm has been developed in economics that process oriented management leads to superior results compared to the traditional focus on functions. When analysing supply chains, the modelling of processes is an important first cornerstone. In this context several questions arise. First, which processes are important for the supply chain and second, how can these processes be modelled.

To answer the first question, the Global Supply Chain Forum identifies eight core supply chain processes (Croxton et al., 2001):

- Customer relationship management,
- Customer service management,
- Demand management,
- Order fulfillment,
- Manufacturing flow management,
- Supplier relationship management (procurement),
- Product development and commercialization,
- Returns management (returns).

Although the importance of each of these processes as well as the activities/operations performed within these processes may vary between different supply chains, these eight processes make up an integral part of the business to be analysed. Both, a strategic view, especially during implementation, and an operational view have to be taken on each of these processes. Figure 2.1 gives an example for the order fulfillment process and shows the sub-processes for either view as well as potential interferences with the other seven core processes.

Going into more detail, processes can be traced best by the flow of materials and information flows. For example, a flow of goods (material flow) is most often initiated by a purchase order (information flow) and followed by an invoice and payment (information and financial flow) to name only a few process steps. Even though several functions are involved: purchasing as initiator, manufacturing as consumer, logistics as internal service provider and finance as debtor. Furthermore, these functions interact with corresponding

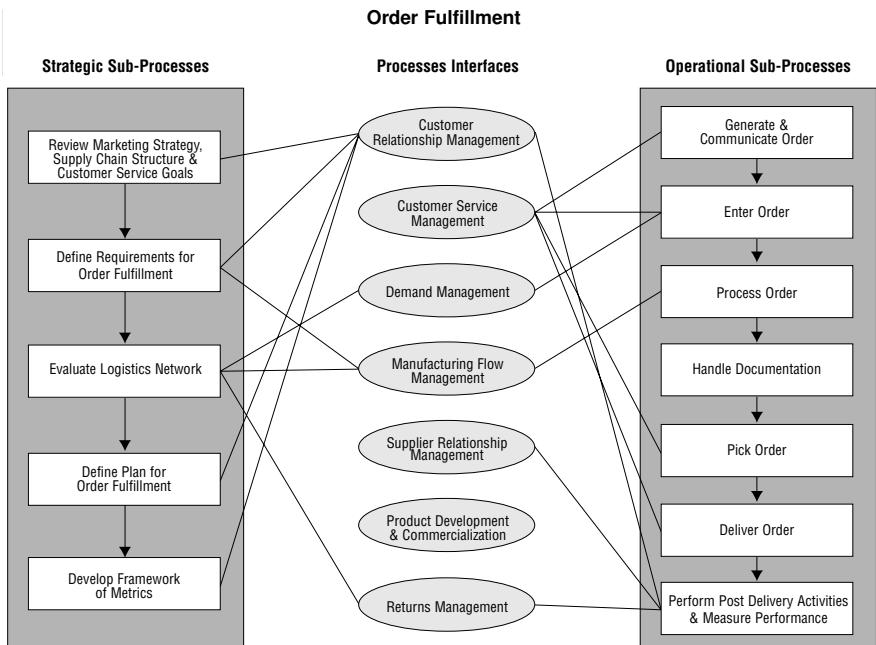


Fig. 2.1. Order fulfillment process (Croxton et al., 2001, p. 21)

functions of the supplier. When analysing supply chains the material flow (and related information flows) need to be mapped from the point of origin to the final customer and probably all the way back, if returns threaten to have a significant impact. Special care needs to be taken at the link between functions, especially when these links bridge two companies, i. e. supply chain members. Nonetheless, a functional view can be helpful when structuring processes.

Various tools and languages have been developed to map processes. One modelling language often cited in the context of supply chain management is the *process chain notation* which has been originally developed by Kuhn (1995) (Brause and Kaczmarek, 2001; Arns et al., 2002).

This notation supports a hierarchical structuring of processes which is a prerequisite to model supply chains because these are often large and complex systems. Furthermore a hierarchical structure allows to model different parts of the supply chain in different levels of detail allowing to focus on the most important sub-processes. Process chains are characterized by sources (e. g. an order of a customer) and sinks (e. g. acceptance of delivery by the customer) which are connected via a chronological sequence of process chain elements. These elements describe the activities/operations to be performed and may be refined into sub-processes.

The concept of event-driven process chains is also very popular to model processes (e. g. ARIS toolset (Scheer, 2002)) and may be used to support supply chain analysis.

The final process models need to be connected to performance measures. Furthermore, the process models can serve a second purpose. They may be used to simulate different scenarios by assigning each process chain element certain attributes (e. g. capacities, process times, availability) and then checking for bottlenecks (Arns et al., 2002). At this point simulation can help to validate newly designed processes and provide the opportunity to make process changes well in time.

The by far most widespread process model especially designed for modelling of supply chains is the SCOR-model which will therefore be introduced in the following subsection in more detail.

2.2.2 The SCOR-Model

The *Supply Chain Operations Reference* (SCOR-)model (current version is 6.0) is a tool for representing, analyzing and configuring supply chains. The SCOR-model has been developed by the *Supply-Chain Council* (SCC) founded in 1996 as a non-profit organization by *AMR Research*, the consulting firm *Pittiglio Rabin Todd & McGrath* (PRTM) and 69 companies (Supply-Chain Council, 2003a). In 2003 SCC had more than 800 members, mainly practitioners (40%), enabling technology providers (25%) and consultants (20%) (Supply-Chain Council, 2003b).

The SCOR-model is a reference model. It does not provide any optimization methods, but aims at providing a standardized terminology for the description of supply chains. This standardization allows benchmarking of processes and the extraction of best practices for certain processes.

Standardized Terminology

Often in different companies different meanings are associated with certain terms. The less one is aware upon the different usage of a term, the more likely misconceptions occur. The use of a standardized terminology that defines and unifies the used terms improves the communication between entities of a supply chain. Thereby, misconceptions are avoided or at least reduced. SCC has established a standard terminology within its SCOR-model.

Levels of the SCOR-Model

The SCOR-model consists of a system of process definitions that are used to standardize processes relevant for SCM. SCC recommends to model a supply chain from the suppliers' suppliers to the customers' customers. Processes such as customer interactions (order entry through paid invoice), physical

material transactions (e.g. equipment, supplies, products, software), market interactions (e.g. demand fulfillment) and (since release 4.0) returns management are supported. Sales and marketing as well as product development and research are not addressed within the SCOR-model (Supply-Chain Council, 2003a, p. 3).

The standard processes are divided into four hierarchical levels: *process types*, *process categories*, *process elements* and *implementation*. The SCOR-model only covers the upper three levels, which will be described in the following paragraphs (following Supply-Chain Council, 2002, p. 10–221), while the lowest (implementation) level is out of the scope of the model, because it is too specific for each company.

Level 1 – Process Types

Level 1 consists of the five elementary process types: *plan*, *source*, *make*, *deliver* and *return* (see Fig. 2.2). These process types comprise operational as well as strategic activities (see Chap. 4). The description of the process types follows Supply-Chain Council (2002) and Supply-Chain Council (2003a).

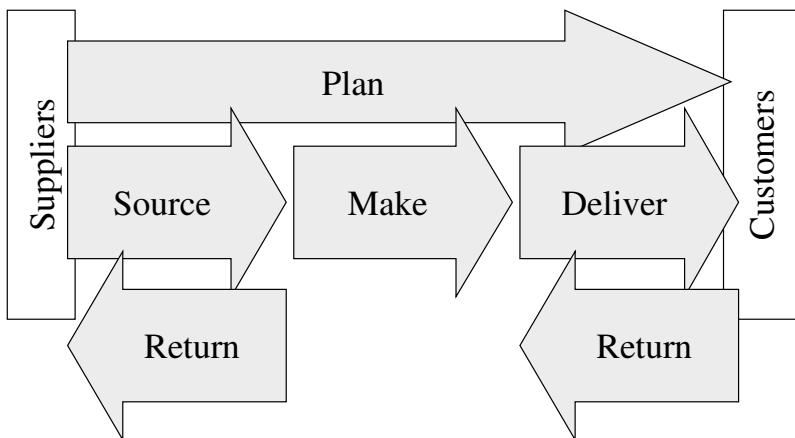


Fig. 2.2. SCOR-model's level 1 (Supply-Chain Council, 2003a, p. 6)

Plan. Plan covers processes to balance resource capacities with demand requirements and the communication of plans across the supply chain. Also in its scope are measurement of the supply chain performance and management of inventories, assets and transportation among others.

Source. Source covers the identification and selection of suppliers, measurement of supplier performance as well as scheduling of their deliveries, receiving of products and processes to authorize payments. It also includes the management of the supplier network and contracts as well as inventories of delivered products.

Make. In the scope of make are processes that transform material, intermediates and products into their next state, meeting planned and current demand. Make covers processes to schedule production activities, produce and test, packaging as well as release of products for delivery. Furthermore, make covers the management of in-process products (WIP), equipment and facilities.

Deliver. Deliver covers processes like order reception, reservation of inventories, generating quotations, consolidation of orders, load building and generation of shipping documents and invoicing. Deliver includes all steps necessary for order management, warehouse management and reception of products at a customer's location together with installation. It manages finished product inventories, service levels and import/export requirements.

Return. In the scope of return are processes for returning defective or excess supply chain products as well as MRO products. The return process extends the scope of the SCOR-model into the area of post-delivery customer service. It covers the authorization of returns, scheduling of returns, receiving and disposition of returned products as well as replacements or credits for returned products. In addition return manages return inventories as well as the compliance to return policies.

Level 2 – Process Categories

The five process types of level 1 are decomposed into 26 *process categories*, including five enable process categories, one for each process type (see Fig. 2.3). At this level typical redundancies of established businesses, such as overlapping planning processes and duplicated purchasing, can be identified. Delayed customer orders indicate a need for integration of suppliers and customers. Each process category is assigned to either *planning*, *execution* or *enable* (see Table 2.1).

Planning. Process categories assigned to planning support the allocation of resources to the expected demand. They incorporate balancing of supply and demand in an adequate planning horizon. Generally, these processes are executed periodically. They directly influence the supply chain's flexibility in respect to changes in demand.

Execution. Execution process categories are those triggered by planned or current demand. In the SCOR-model, they regularly incorporate scheduling and sequencing as well as transforming and/or transporting products. The process types source, make and deliver are further decomposed with respect to the nature of customer orders (e.g. make-to-stock, make-to-order and engineer-to-order). Execution process categories depict the core processes of a supply chain.

Enable. Process categories assigned to enable are support processes for the other process categories. They prepare, preserve and control the flow of information and the relations between the other processes.

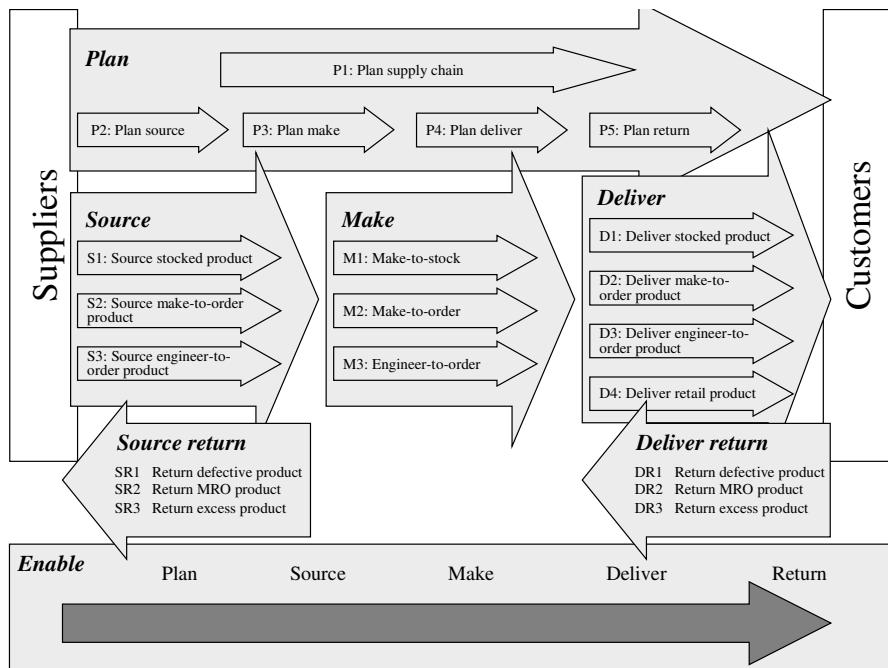


Fig. 2.3. SCOR-model's levels 1 and 2 (Supply-Chain Council, 2003a, p. 9)

Table 2.1. Process types and categories

	Plan	Source	Make	Deliver	Return
Planning	P1	P2	P3	P4	P5
Executing		S1–S3	M1–M3	D1–D4	SR1–SR3, DR1–DR3
Enabling	EP	ES	EM	ED	ER

Level 3 – Process Elements

At this level, the supply chain is tuned. The process categories are further decomposed into *process elements*. Detailed metrics and best practices for these elements are part of the SCOR-model at this level. Furthermore, most process elements can be linked and possess an *input* stream (information and material) and/or an *output* stream (also information and material). Figure 2.4 shows an example for the third level of the “P1: Plan supply chain” process category. Supply-Chain Council (2002, pp. 10–51) gives the following definitions for this process category and its process elements:

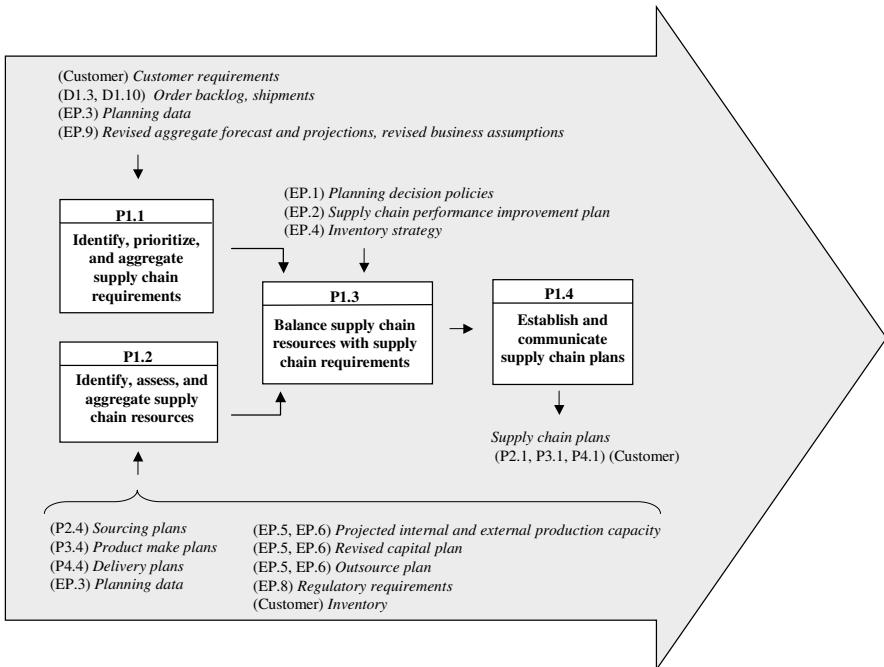


Fig. 2.4. Example of SCOR-model's level 3 (Supply-Chain Council, 2002, p. 11)

P1. The development and establishment of courses of action over specified time periods that represent a projected appropriation of supply chain resources to meet supply chain requirements.

P1.1. The process of identifying, prioritizing and considering, as a whole with constituent parts, all sources of demand in the supply chain of a product or service.

P1.2. The process of identifying, evaluating, and considering, as a whole with constituent parts, all things that add value in the supply chain of a product or service.

P1.3. The process of developing a time-phased course of action that commits supply-chain resources to meet supply-chain requirements.

P1.4. The establishment of course of action over specified time periods that represent a projected appropriation of supply-chain requirements.”

The input and output streams of a process element are not necessarily linked to input and output streams of other process elements. However, the indication in brackets depicts the corresponding supply chain partner, process type, process category or process element from where information or material comes. Thus, the process elements are references, not examples of possible sequences.

The process elements are decomposed on the fourth level. Companies implement their specific management practices at this level. Not being part of the SCOR-model, this step will not be subject of this book.

Metrics and Best Practices

The SCOR-model supports performance measurement on each level. Level 1 metrics provide an overview of the supply chain for the evaluation by management (see Table 2.2). Levels 2 and 3 include more specific and detailed metrics corresponding to process categories and elements. Table 2.3 gives an example of level 3 metrics that are corresponding to the “S1.1: Schedule product deliveries” process element.

Table 2.2. SCOR's level 1 metrics (Supply-Chain Council, 2003b, p. 17–18)

External, customer facing			Internal facing	
Reliability	Responsiveness	Flexibility	Cost	Assets
Delivery performance	Order fulfillment lead time	Supply chain response time	Total supply chain management cost	Cash-to-cash cycle time
Perfect order fulfillment		Production flexibility	Value-added productivity	Inventory days of supply
Fill rate			Warranty / returns processing cost	Asset turns
			Cost of goods sold	

The metrics are systematically divided into the five categories *reliability*, *flexibility*, *responsiveness*, *cost* and *assets*. Reliability as well as flexibility and responsiveness are external (customer driven), whereas cost and assets are metrics from an internal point of view.

In 1991 PRTM initiated the *Supply Chain Performance Benchmarking Study* (now: Supply-Chain Management Benchmarking Series) for SCC members (Stewart, 1995). Within the scope of this study all level 1 metrics and selected metrics of levels 2 and 3 are gathered. The information is evaluated regarding different lines of business. Companies joining the Supply Chain Performance Benchmarking Study are able to compare their metrics with the evaluated ones. Furthermore, associated best practices are identified. Selected

Table 2.3. SCOR's level 3 metrics – example “S1.1: Schedule product deliveries” (Supply-Chain Council, 2003a, p. 11)

Performance attributes	Metric
Reliability	% schedules generated within supplier's lead time % schedules changed within supplier's lead time
Responsiveness	Average release cycle of changes
Flexibility	Average days per schedule change Average days per engineering change
Cost	Product management and planning costs as a % of product acquisition costs
Assets	None identified

best practices, corresponding to process categories and process elements, are depicted in the following paragraph.

An example of an identified best practice for the “P1: Plan supply chain” process category is high integration of the supply/demand process from gathering customer data and order receipt, through production to supplier request. SCC recommends performing this integrated process by using an APS with interfaces to all supply/demand resources. Moreover, the utilization of tools that support balanced decision-making (e.g. trade-off between service level and inventory investment) is identified as best practice (Supply-Chain Council, 2002, p. 12). To perform process element “P1.3: Balance supply chain resources with supply chain requirements” (see Fig. 2.4) effectively, balancing of supply and demand to derive an optimal combination of customer service and resource investment by using an APS is recognized as best practice.

A Procedure for Application of the SCOR-Model

Having described the elements of the SCOR-model, a procedure for its application will be outlined that shows how the SCOR-model can be configured for a distinct supply chain (Supply-Chain Council, 2003a, p. 17–19). This configuration procedure consists of seven steps:

1. Define the business unit to be configured.
2. Geographically place entities that are involved in source, make, deliver and return process types. Not only locations of a single business, but also locations of suppliers (and suppliers' suppliers) and customers (and customers' customers) should be denoted.
3. Enter the major flows of materials as directed arcs between locations of entities.
4. Assign and link the most important source, make, deliver and return processes categories to each location (see Fig. 2.5).

5. Define partial process chains of the (modelled) supply chain (e.g. for distinct product families). A partial process chain is a sequence of processes that are planned for by a single “P1” planning process category.
6. Enter planning process categories (“P2”–“P5”) using dashed lines to illustrate the assignment of execution to planning process categories (see Fig. 2.5 and also Table 2.1).
7. Define a top-level “P1” planning process if possible, i.e. a planning process category that coordinates two or more partial process chains.

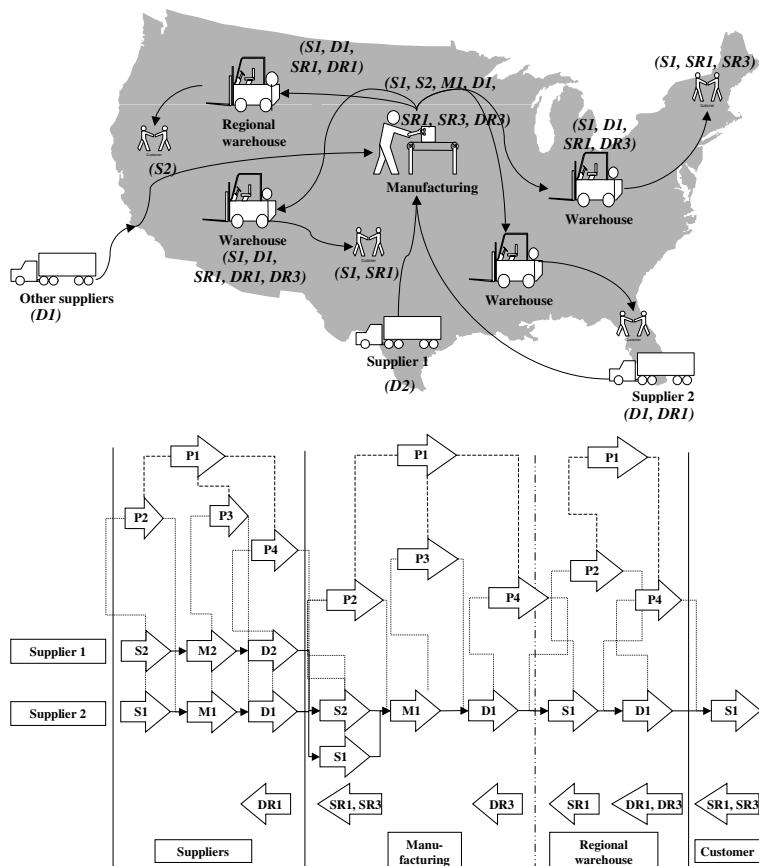


Fig. 2.5. Example results of steps 4 and 6 (cf. Supply-Chain Council, 2003a, pp. 17–19)

After configuring the supply chain, performance levels, practices and systems are aligned. Critical process categories of level 2 can be detailed in level 3.

At this level the most differentiated metrics and best practices are available. Thus, detailed analysis and improvements of process elements are supported.

The *implementation of supply chain processes and systems* is, as already mentioned, not part of the SCOR-model. However, it is recommended to continue to use the metrics of the SCOR-model. They provide data for internal and external benchmarking studies to measure and document consequences of change processes within a supply chain.

2.3 Performance Measurement

Having mapped the supply chain processes it is important to assign measures to these processes to evaluate changes and to assess the performance of the complete supply chain as well as of the individual processes. Thereby it is crucial not to measure “something”, but to find the most relevant metrics. These not only need to be aligned with the supply chain strategy (see Sect. 1.2.4), but also need to reflect important goals in the scope and within the influence of the part of the organisation responsible for the individual process under consideration. In the two subsequent subsections, first some general issues of performance measurement within a supply chain setting will be discussed, whereas second some key performance indicators for supply chains will be introduced.

2.3.1 Issues Regarding Performance Measurement

Indicators are defined as numbers that inform about relevant criteria in a clearly defined way (see e.g. (Horváth, 2003) for a comprehensive introduction to indicators and systems of indicators). Performance indicators (measures, metrics) are utilized in a wide range of operations. Their primary application is in *operational controlling*. Hardly a controlling system is imaginable that does not make use of performance measures regularly. In fact, the utilization of a wide variety of measures (as necessary) to model all business processes of a company enables the company to run its business according to management-by-exception.

Three functions can be attributed to indicators:

Informing. Their main purpose is to inform management. In this function, indicators are applied to support decision-making and to identify problem areas. Indicators can therefore be compared with standard or target values.

Steering. Indicators are the basis for target setting. These targets guide those responsible for the process considered to accomplish the desired outcome.

Controlling. Indicators are also well suited for the supervision of operations and processes.

The main disadvantage inherent to indicators is that they are only suited to describe *quantitative facts*. “Soft” facts are difficult to measure and likely to be neglected when indicators are introduced (e.g. motivation of personnel). Still, non-quantitative targets which are not included in the set of indicators should be kept in mind.

When using indicators, one key concern is their *correct interpretation*. It is essential to keep in mind that variations observed by indicators have to be linked to a *causal model* of the underlying process or operation. A short example will illustrate this. To measure the productivity of an operation the ratio of revenue divided by labour is assumed here as an appropriate indicator:

$$\text{productivity} = \frac{\text{revenue}[\$]}{\text{labour}[h]} \quad (2.1)$$

Revenue is measured in currency units (\$), whereas labour is measured in hours worked (per plant, machine or personnel), where the relevance of the different measures for labour depends on the specific product(s) considered. Supposed productivity is 500 \$/h in one period and 600 \$/h in the next period, there is definitely a huge difference. In fact, when calculating productivity a *causal link* between revenue and labour is assumed implicitly. On the other hand, there are many more rationales that could have caused this increase in productivity. These have to be examined too before a final conclusion can be derived. In this example price hikes, changes in product mix, higher utilization of resources or decreased inventories can account for substantial portions of the observed increase in productivity. Therefore, it is essential to find appropriate measures with clear links connecting the indicator and the causal model of the underlying process (root causes).

Furthermore, indicators have to be evaluated how they translate to the strategic goals of the supply chain. If indicators and strategy are not aligned, it may well happen that one supply chain entity pursues a conflicting goal. For example, one partner increases its inventory turn rate by reducing safety stock, which negatively affects the downstream delivery performance of its partners.

When choosing supply chain performance metrics it is essential to keep in mind the cross-functional process-oriented nature of the supply chain. Functional measures may be too narrow-minded and should be substituted by cross-functional measures, therefore helping that not individual entities optimize only their functional goals (e.g. maximizing capacity utilization), but shared goals (e.g. a superior order fill rate compared to a rival supply chain).

Historically, indicators and systems of indicators have been based on *financial data*, as financial data have been widely available for long. Improvements in terms of superior financial performance that are caused by the successful application of SCM can be measured by these indicators. Nevertheless some additional, more appropriate measures of supply chain performance

should be derived, since the focal points of SCM are customer orientation, the integration of organizational units and their coordination.

The transition to incorporate non-financial measures in the evaluation of business performance is widely accepted, though. Kaplan and Norton (1992) introduced the concept of a balanced scorecard (BSC) that received broad attention not only in scientific literature but also in practical applications. In addition to financial measures, the BSC comprises a customer perspective, an innovation and learning perspective as well as an internal business perspective. These perspectives integrate a set of measures into one management report that provides a deeper insight into a company's performance. The measures chosen depend on the individual situation faced by the company. Figure 2.6 gives an example of a BSC used by a global engineering and construction company.

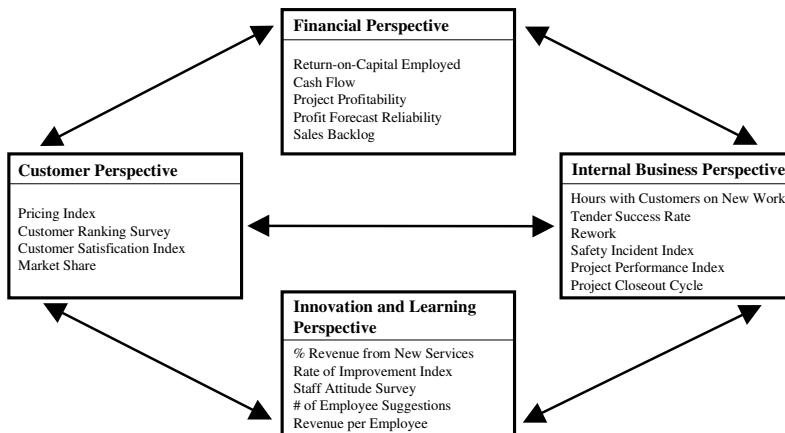


Fig. 2.6. Example of indicators used by a balanced scorecard (Kaplan and Norton, 1993, p. 136)

An increasing number of contributions in the literature is dealing with the adaption of BSCs to fit the needs of SCM (see e.g. Brewer and Speh (2000) and Bullinger et al. (2002)). Adaptions are proposed within the original framework consisting of the four perspectives introduced above, but also structural changes are proposed. For example, Weber et al. (2002) propose a BSC for supply chains consisting of a financial perspective, a process perspective and two new perspectives relating to cooperation quality and cooperation intensity. In addition to the supply chain BSC they propose individual company BSCs on a second hierarchical level. In contrast to the supply chain BSC these still might comprise of a customer perspective (for the most downstream supply chain partner) and a learning perspective.

Non-financial measures have the advantage that they are often easier to quantify as there is no allocation of costs necessary for their calculation. Moreover, they turn attention to physical processes more directly. As will be discussed later delivery performance is a critical performance measure for supply chains. How superior delivery performance can relate to the financial performance (here: the economic value added (EVA)) is depicted in Fig. 2.7. It is shown which financial data are affected by delivery performance and how these financial data affect EVA as the ultimate performance criteria in this example. On the other hand, it is important to identify by which measures in terms of process improvements better delivery performance can be achieved. A similar instrument providing connections of root causes and financial performance measures via non-financial/logistical key performance indicators are the Enabler-KPI-Value networks presented in Chap. 15.

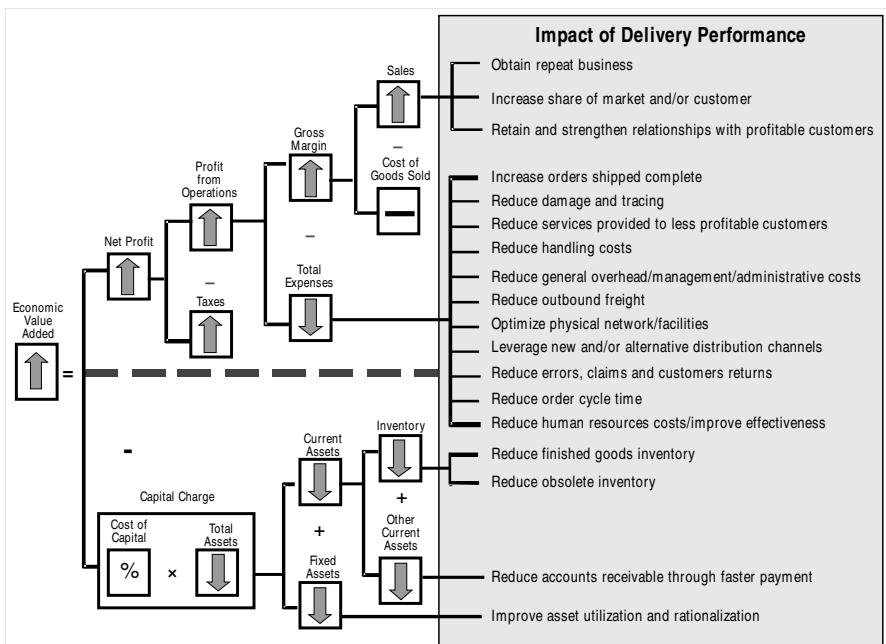


Fig. 2.7. Relation of delivery performance and EVA (Lambert and Pohlen, 2001, p. 13)

Specifically when assessing supply chain performance it is important to bear in mind the following issues:

Definition of indicators. As supply chains usually span over several companies or at least several entities within one company a *common definition* of all indicators is obligatory. Otherwise the comparison of indicators and their uniform application can be counterproductive.

Perspective on indicators. The view on indicators might be different considering the roles of the two supply chain partners, the supplier and the customer. A supplier might want to calculate the order fill rate based on the order receipt date and the order ship date, as these are the dates he is able to control. From the customer's point of view the basis would be the request date and the receipt date at customer's warehouse. If supplier's and customer's dates do not match, this will lead to different results with respect to an agreed order fill rate. This is why both have to agree on *one perspective*.

Capturing of data. Data needed to calculate the indicators should be captured in a consistent way throughout the supply chain. *Consistency* with respect to units of measurement and the availability of *current* data for the supply chain partners are essential. Furthermore, *completeness* of the used data is obligatory, i. e. all necessary data should be available in adequate systems and accessible by supply chain partners.

Confidentiality. Confidentiality is another major issue if more than one company form the supply chain. As all partners are separate legal entities, they might not want to give complete information about their internal processes to their partners. Furthermore, there might be some targets which are not shared among partners.

Nevertheless, it is widely accepted that supply chain integration benefits from the utilization of key performance indicators. They support communication between supply chain partners and are a valuable tool for the coordination of their individual, but shared plans.

2.3.2 Key Performance Indicators for Supply Chains

A vast amount of literature has been published suggesting performance indicators for supply chains (e. g. Lapide (2000), Gunasekaran et al. (2001), Bullinger et al. (2002) and Hausman (2003)). Although each supply chain is unique and might need special treatment, there are some performance measures that are applicable in most settings. In the following paragraphs these will be presented as key performance indicators. As they tackle different aspects of the supply chain they are grouped into four categories corresponding to the following attributes: delivery performance, supply chain responsiveness, assets and inventories, and costs.

Delivery Performance

As customer orientation is a key component of SCM, delivery performance is an essential measure for total supply chain performance. As promised delivery dates may be too late in the eye of the customer, his expectation or even request fixes the target. Therefore delivery performance has to be measured in terms of the actual delivery date compared to the delivery date mutually

agreed upon. Only perfect order fulfillment which is reached by delivering the right product to the right place at the right time ensures customer satisfaction. An on time shipment containing only 95% of items requested will often not ensure 95% satisfaction with the customer. Increasing delivery performance may improve the competitive position of the supply chain and generate additional sales. Regarding different aspects of delivery performance, various indicators called *service levels* are distinguished in inventory management literature (see e.g. Tempelmeier (2003, pp. 397-398) or Silver et al. (1998, p. 245)). The first one, called α -service level (P_1 , cycle service level), is an event-oriented measure. It is defined as the probability that an incoming order can be fulfilled completely from stock. Usually, it is determined with respect to a predefined period length (e.g. day, week or order cycle). Another performance indicator is the quantity-oriented β -service level (P_2), which is defined as the proportion of incoming order quantities that can be fulfilled from inventory on-hand. In contrast to the α -service level, the β -service level takes into account the extent to which orders cannot be fulfilled. The γ -service level is a time- and quantity-oriented measure. It comprises two aspects: the quantity that cannot be met from stock and the time it takes to meet the demand. Therefore it contains the time information not considered by the β -service level. An exact definition is:

$$\gamma\text{-service level} = 1 - \frac{\text{mean backlog at end of period}}{\text{mean demand per period}} \quad (2.2)$$

The *order fill rate* as it is defined in the SCOR-model is closely related to the α -service level and can be described as the percentage of ship-from-stock orders shipped within 24 hours.

Furthermore, *on time delivery* is an important indicator. It is defined as the proportion of orders delivered on or before the date requested by the customer. A low percentage of on time deliveries indicates that the order promising process is not synchronized with the execution process. This might be due to order promising based on an infeasible (production) plan or because of production or transportation operations not executed as planned.

Measuring *forecast accuracy* is also worthwhile. Forecast accuracy relates forecasted sales quantities to actual quantities and measures the ability to forecast future demands. Better forecasts of customer behaviour usually lead to smaller changes in already established production and distribution plans. An overview of methods to measure forecast accuracy is given in Chap. 7.

A further important indicator in the context of delivery performance is the *order lead-time*. Order lead-times measure, from the customer's point of view, the average time interval from the date the order is placed to the date the customer receives the shipment. As customers are increasingly demanding, short order lead-times become important in competitive situations. Nevertheless, not only short lead-times but also reliable lead-times will sat-

isfy customers and lead to a strong customer relationship, even though the two types of lead-times (shortest vs. reliable) have different cost aspects.

Supply Chain Responsiveness

Responsiveness describes the ability of the complete supply chain to react according to changes in the marketplace. Supply chains have to react to significant changes within an appropriate time frame to ensure their competitiveness. To quantify responsiveness separate flexibility measures have to be introduced to capture the ability, extent and speed of adaptations. These indicators shall measure the ability to change plans (flexibility within the system) and even the entire supply chain structure (flexibility of the system). An example in this field is the upside production flexibility determined by the number of days needed to adapt to an unexpected 20% growth in the demand level.

A different indicator in this area is the *planning cycle time* which is simply defined as the time between the beginning of two subsequent planning cycles. Long planning cycle times prevent the plan from taking into account the short-term changes in the real world. Especially planned actions at the end of a planning cycle may no longer fit to the actual situation, since they are based on old data available at the beginning of the planning cycle. The appropriate planning cycle time has to be determined with respect to the aggregation level of the planning process, the planning horizon and the planning effort.

Assets and Inventories

Measures regarding the assets of a supply chain should not be neglected. One common indicator in this area is called *asset turns*, which is defined by the division of revenue by total assets. Therefore, asset turns measure the efficiency of a company in operating its assets by specifying sales per asset. This indicator should be watched with caution as it varies sharply among different industries.

Another indicator worthy of observation is *inventory turns*, defined as the ratio of total material consumption per time period over the average inventory level of the same time period. A common approach to increase inventory turns is to reduce inventories. Still, inventory turns is a good example to illustrate that optimizing the proposed measures may not be pursued as isolated goals. Consider a supply chain consisting of several tiers each holding the same quantity of goods in inventory. As the value of goods increases as they move downstream the supply chain, an increase in inventory turns is more valuable if achieved at a more downstream entity. Furthermore, decreasing downstream inventories reduces the risk of repositioning of inventories due to bad distribution. However, reduced inventory holding costs may be offset by increases in other cost components (e. g. production setup costs) or unsatisfied customers (due to poor delivery performance). Therefore, when using

this measure it needs to be done with caution, keeping a holistic view on the supply chain in mind.

Lastly, the *inventory age* is defined by the average time goods are residing in stock. Inventory age is a reliable indicator for high inventory levels, but has to be used with respect to the items considered. Replacement parts for phased out products will usually have a much higher age than stocks of the newest released products. Nevertheless, the distribution of inventory ages over products is suited perfectly for identifying unnecessary “pockets” of inventory and for helping to increase inventory turns.

Determining the right inventory level is not an easy task, as it is product- and process-dependent. Furthermore, inventories not only cause costs, but there are also benefits to holding inventory. Therefore, in addition to the aggregated indicators defined above, a proper analysis not only regarding the importance of items (e.g. an ABC-analysis), but also a detailed investigation of inventory components (as proposed in Sect. 2.4) might be appropriate.

Costs

Last but not least some financial measures should be mentioned since the ultimate goal will generally be profit. Here, the focus is on cost based measures. Costs of goods sold should always be monitored with emphasis on substantial processes of the supply chain. Hence, an integrated information system operating on a joint database and a mutual cost accounting system may prove to be a vital part of the supply chain.

Further, productivity measures usually aim at the detection of cost drivers in the production process. In this context *value-added employee productivity* is an indicator which is calculated by dividing the difference between revenue and material cost by total employment (measured in (full time) equivalents of employees). Therefore, it analyses the value each employee adds to all products sold.

Finally, *warranty costs* should be observed, being an indicator for product quality. Although warranty costs depend highly on how warranty processing is carried out, it may help to identify problem areas. This is particularly important because superior product quality is not a typical supply chain feature, but a driving business principle in general.

2.4 Inventory Analysis

Often claimed citations like “inventories hide faults” suggest to avoid any inventory in a supply chain. This way of thinking is attributed to the Just-In-Time-philosophy, which aligns the processes in the supply chain such that almost no inventories are necessary. This is only possible in some specific industries or certain sections of a supply chain and for selected items.

In all other cases inventories are necessary and therefore need to be managed in an efficient way. Inventories in supply chains are always the result of inflow and outflow processes (transport, production etc.). This means that the isolated minimization of inventories is not a reasonable objective of SCM, instead they have to be managed together with the corresponding supply chain processes.

Inventories cause costs (holding costs), but also provide *benefits*, in particular reduction of costs of the inflow and/or outflow processes. Thus, the problem is to find the right trade-off between the costs for holding inventories and the benefits.

Inventory decomposes into different *components* according to the *motives* for holding inventory. The most important components are shown in Table 2.4 and will be described in detail in the following paragraphs. The distinction

Table 2.4. Stock components, determinants, and benefits

stock component	determinants	benefits
production lot-sizing stock	setup frequency	reduced setup time and costs
transportation lot-sizing stock	shipment quantity	reduced transportation costs
inventory in transit	transportation time	reduced transportation costs
seasonal stock	demand peaks, tight capacity	reduced costs for overtime and for investments
work-in-process	lead time, production planning and control	increased utilization, reduced investments in additional capacity
safety stock	demand and lead time uncertainty, process uncertainties	increased service level, reduced costs for emergency shipments and lost sales

of stock components is necessary for

- the identification of benefits,
- the identification of determinants of the inventory level, and
- setting target inventory levels (e.g. in APS).

The *inventory analysis* enables us to decompose the *average* inventory level in a supply chain. It shows the different causes for inventories held in the past and indicates the relative importance of specific components. The *current* inventory of certain stock keeping units (SKUs) on the other hand might be higher or lower depending on the point in time chosen. Thus, the current inventory is not suitable for a proper inventory analysis.

In an ex-post analysis it is possible to observe whether the trade-off between the benefits and the stock costs has been managed efficiently for each

component and SKU (*inventory management*). In the following paragraphs we will show the motives, the benefits, and determinants of some important components (see also Chopra and Meindl (2001), pp. 52).

Production Lot-sizing or Cycle Stock

The cycle stock (we use 'production lot-sizing stock', 'lot-sizing stock' and 'cycle stock' synonymously) is used to cover the demand between two consecutive production runs of the same product. For example, consider a color manufacturing plant, which produces blue and yellow colours, alternating between each bi-weekly. Then, the production lot has to cover the demand in the current and the following week. Thus, the production quantity (lot) equals the two-week demand and the coverage is two weeks. The role of cycle stock is to reduce the costs for setting up and cleaning the production facility (setup or changeover costs). Finding the right trade-off between fixed setup costs and inventory costs is usually a critical task, as this decision may also depend on the lot-size of other products. An overview on the problems arising here is given in Chap. 10.

For the inventory analysis of final items in a make-to-stock environment it is mostly sufficient to consider a cyclic production pattern with average lot-sizes q^p over a time interval that covers several production cycles. Then, the inventory level follows the so-called "saw-tooth"-pattern, which is shown in Fig. 2.8. The average cycle stock CS is half the average lot-size: $CS = q^p/2$.

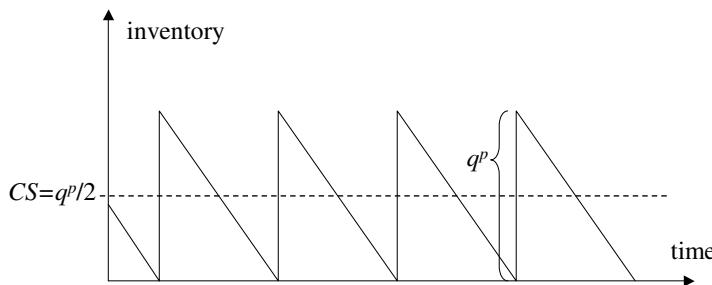


Fig. 2.8. Inventory pattern for cycle stock calculation

The average lot-size can be calculated from the total number of production setups su and the total demand d^p during the analysis interval: $q^p = d^p/su$. Thus, all you need to analyse cycle stock is the number of production setups and the total demand.

Transportation Lot-sizing Stock

The same principle of reducing the amount of fixed costs per lot applies to transportation links. Each truck causes some amount of fixed costs which arise

for a transport from warehouse A to warehouse B. If this truck is only loaded partially, then the cost per unit shipped is higher than for a full truckload. Therefore, it is economical to batch transportation quantities up to a full load and to ship them together. Then, one shipment has to cover the demand until the next shipment arrives at the destination. The decision on the right transportation lot-size usually has to take into account the dependencies with other products' shipments on the same link and the capacity of the transport unit (e.g. truck, ship etc.) used (see Chap. 12).

For the inventory analysis we can calculate the average transportation quantity q^t from the number of shipments s during the analysis interval and the total demand d^t for the product at the destination warehouse by $q^t = d^t/s$. In contrast to the production lot-sizing stock, the average transportation lot-sizing stock equals not half, but the whole transportation quantity q^t , if we consider both the “source warehouse”, where the inventory has to be built up until the next shipment starts and the “destination warehouse” where the inventory is depleted until the next shipment arrives. Therefore, the average stock level at each warehouse is one half of the transportation lot-size and, the transportation lot-sizing stock sums up to $TLS = q^t$.

This calculation builds on the assumption of a continuous inflow of goods to the source warehouse, which is valid if the warehouse is supplied by continuous production or by production lots which are not coordinated with the shipments. This is the case for most production-distribution chains.

Inventory in Transit

While the transportation lot-sizing stock is held at the start and end stock points of a transportation link, there exists also inventory that is currently transported in-between. This stock component only depends on the transportation time and the demand because on average the inventory “held on the truck” equals the demand which occurs during the transportation time. The inventory in transit is independent of the transportation frequency and therefore also independent of the transportation lot-size. The inventory in transit can be reduced at the expense of increasing transportation costs, if the transportation time is reduced by a faster transportation mode (e.g. plane instead of truck transport).

The average inventory in transit TI is calculated by multiplying the average transportation time by the average demand. For instance, if the transportation time is two days and the average amount to be transported is 50 pieces per day, then $TI = 100$ pieces.

Seasonal Stock or Pre-built Stock

In seasonal industries (e.g. consumer packaged goods) inventories are held to buffer future demand peaks which exceed the production capacities. In

this sense, there is a trade-off between the level of regular capacity, additional overtime capacity and seasonal stock. The seasonal stock can help to reduce lost sales, costs for working overtime or opportunity costs for unused machines and technical equipment. In contrast to the previous stock components which are defined by SKU, the seasonal stock is common for a group of items sharing the same tight capacity. Figure 2.9 shows how the total amount of seasonal inventory can be calculated from the capacity profile of a complete seasonal cycle. In this case, the seasonal stock is built up in periods 3 and 4

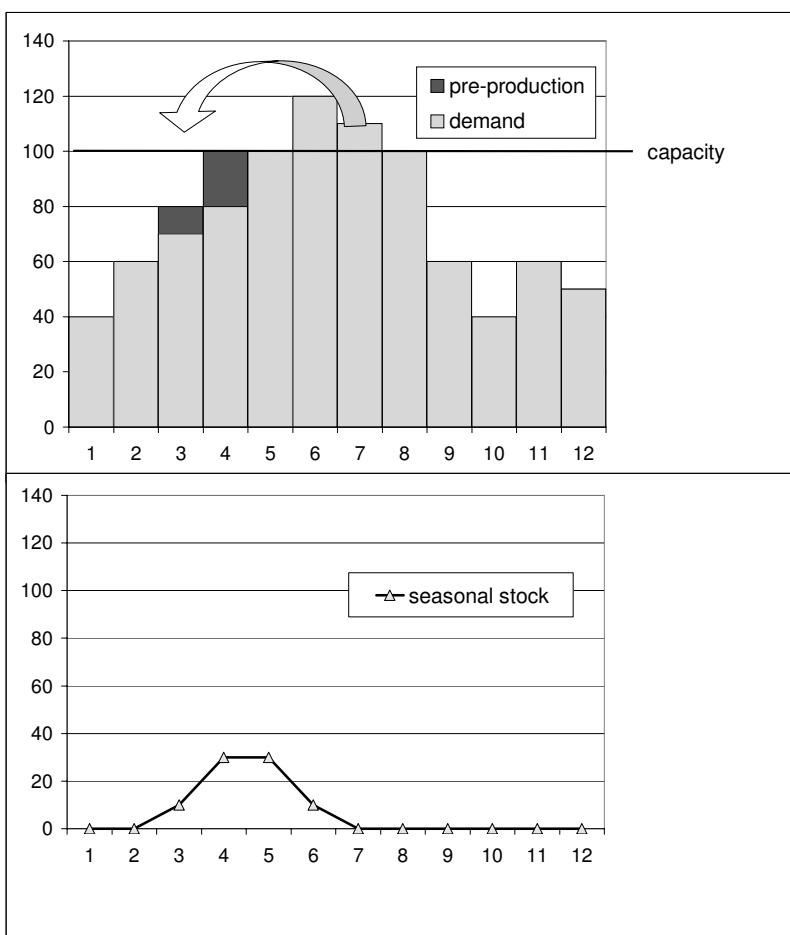


Fig. 2.9. Example for the determination of seasonal stock

and used for demand fulfilment in periods 6 and 7. The total seasonal stock shown in the figure is calculated using the assumption that all products are pre-produced in the same quantity as they are demanded in the bottleneck

periods. In practice one would preferably pre-build those products, which create only small holding costs and which can be forecasted with high certainty. In Chap. 8 we will introduce planning models, which help to decide on the right amount of seasonal stock.

Work-in-process Inventory (WIP)

The WIP inventory can be found in every supply chain, because the production process takes some time during which the raw materials and components are transformed to finished products. In a multi-stage production process the production lead time consists of the actual processing times on the machines and additional waiting times of the products between the operations, e.g. because required resources are occupied. The *benefits* of the WIP are that it prevents bottleneck machines from starving for material and maintains a high utilization of resources. Thus, WIP may avoid investments in additional capacities. The waiting time part of production lead time is also influenced by the production planning and control system (see also Chap. 10), which should schedule the orders so as to ensure short lead times. Therefore, it is possible to reduce the WIP by making effective use of an APS. In this sense, the opinion “inventories hide faults” indeed applies to the WIP in the modified form: Too high WIP hides faults of production planning and control.

According to Little’s law (see e. g. Silver et al. (1998), pp. 697) the average production lead time LT is proportional to the WIP level. If d^w is the average demand per unit of time, then $WIP = LT \cdot d^w$.

Safety Stock

Safety stock has to protect against uncertainty which may arise from internal processes like production lead time, from unknown customer demand and from uncertain supplier lead times. This implies that the main drivers for the safety stock level are production and transport disruptions, forecasting errors, and lead time variations. The benefit of safety stock is that it allows quick customer service and avoids lost sales, emergency shipments, and the loss of goodwill. Furthermore, safety stock for raw materials enables smoother flow of goods in the production process and avoids disruptions due to stock-outs at the raw material level. Besides the uncertainty mentioned above the main driver for safety stock is the length of the lead time (production or procurement), which is necessary to replenish the stock.

In the inventory analysis, the *observed* safety stock is the residual level, which is left after subtracting all of the components introduced above from the average observed inventory level. This observed safety stock can then be compared with the level of safety stock that is necessary from an economical standpoint. A short introduction on how *necessary* safety stocks can be calculated is given in Chap. 7.

A further component which may occur in a distribution centre is the order picking inventory. It comprises the partly filled pallets from which the small quantities per customer order are picked.

The main steps of the inventory analysis are summarized in the following:

1. Calculate the average inventory level (AVI) from past observations over a sufficiently long period (e.g. half a year) of observations (e.g. inventory levels measured daily or weekly).
2. Identify possible stock components (e.g. cycle stock, safety stock) and their corresponding drivers (e.g. lot-size, lead time).
3. Decompose the AVI into the components including the *observed* safety stock.
4. Calculate the *necessary* safety stock and compare it to the *observed* safety stock.
5. The remaining difference (+/-) shows avoidable buffer stock (+) or products which didn't have enough stock (-).
6. For the most important components of the observed inventory calculate the optimal target level w. r. t. inventory costs and benefits.

For the optimization of inventory, the main principle of inventory management has to be considered: The objective is to balance the costs arising from holding inventories and the benefits of it. Furthermore, this trade-off has to be handled for each separate component. In Part II we will show how APS can support this critical task of inventory management.

References

- Arns, M.; Fischer, M.; Kemper, P.; Tepper, C. (2002) *Supply chain modelling and its analytical evaluation*, Journal of the Operational Research Society, Vol. 53, 885–894
- Brause, F.; Kaczmarek, M. (2001) *Modellierung und Analyse von Supply Chains*, Wirtschaftsinformatik, Vol. 43, No. 6, 569–578
- Brewer, P. C.; Speh, T. W. (2000) *Adapting the balanced scorecard to supply chain management*, Supply Chain Management Review, Vol. 5, No. 2, 48–56
- Bullinger, H.-J.; Kühner, M.; van Hoof, A. (2002) *Analysing supply chain performance using a balanced measurement method*, International Journal of Production Research, Vol. 40, No. 15, 3533–3543
- Chopra, S.; Meindl, P. (2001) *Supply Chain Management*, Upper Saddle River, New Jersey
- Croxton, K. L.; Garcia-Dastugue, S. J.; Lambert, D. M.; Rogers, D. S. (2001) *The supply chain management processes*, The International Journal of Logistics Management, Vol. 12, No. 2, 13–36
- Dreyer, D. E. (2000) *Performance measurement: a practitioner's perspective*, Supply Chain Management Review, Vol. 5, No. 5, 62–68
- Fisher, M. L. (1997) *What is the right supply chain for your product?*, Harvard Business Review, Vol. 75, No. 2, 105–116

- Gunasekaran, A.; Patel, C.; Tirtiroglu, E. (2001) *Performance measures and metrics in a supply chain environment*, International Journal of Operations & Production Management, Vol. 21, No. 1/2, 71–87
- Hausman, W. H. (2003) *Supply chain performance metrics*, in: Harrison, T. P.; Lee, H. L.; Neale, J. J. (Eds.) *The practice of supply chain management: Where theory and application converge*, Boston
- Horváth, P. (2003) *Controlling*, 9th ed., München
- Kaplan, R. S.; Norton, D. P. (1992) *The balanced scorecard – Measures that drive performance*, Harvard Business Review, Vol. 70, No. 1, 71–79
- Kaplan, R. S.; Norton, D. P. (1993) *Putting the balanced scorecard to work*, Harvard Business Review, Vol. 71, No. 5, 134–142
- Kuhn, A. (1995) *Prozeßketten in der Logistik. Entwicklungstrends und Umsetzungsstrategien*, Dortmund
- Lambert, D. M.; Pohlen, T. L. (2001) *Supply chain metrics*, The International Journal of Logistics Management, Vol. 12, No. 1, 1–19
- Lapide, L. (2000) *What about measuring supply chain performance*, URL: <http://lapide.ascet.com>, Date: January 30, 2004
- Ramdas, K.; Spekman, R. E. (2000) *Chain or shackles: understanding what drives supply-chain performance*, Interfaces, Vol. 30, No. 4, 3–21
- Scheer, A.-W. (2002) *ARIS - Vom Geschäftsprozeß zum Anwendungssystem*, 4th ed., Berlin et al.
- Silver, E. A.; Pyke, D. F.; Petersen, R. (1998) *Inventory management and production planning and scheduling*, 3rd ed., New York et al.
- Stewart, G. (1995) *Supply chain performance benchmarking study reveals keys to supply chain excellence*, Logistics Information Management, Vol. 8, No. 2, 38–44
- Supply-Chain Council (2002) *Supply chain operations reference-model – version 5.0*, Tech. Paper, Pittsburgh
- Supply-Chain Council (2003a) *Supply chain operations reference model – overview of SCOR version 6.0*, Pittsburgh, URL: <http://www.supplychainassociates.com/sitebuildercontent/sitebuilderfiles/scor6.0overview.pdf>, Date: January 30, 2004
- Supply-Chain Council (2003b) *Supply chain operations reference-model (SCOR) – overview presentation*, URL: http://www.supply-chain.org/News/SCOR_6.0_Overview_7-23-03.zip, Date: January 30, 2004
- Tempelmeier, H. (2003) *Material-Logistik: Modelle und Algorithmen für die Produktionsplanung und -steuerung und das Supply Chain Management*, 5th ed., Berlin et al.
- Weber, J.; Bacher, A.; Groll, M. (2002) *Konzeption einer Balanced Scorecard für das Controlling von unternehmensübergreifenden Supply Chains*, Kostenrechnungspraxis, Vol. 46, No. 3, 133–141

3 Types of Supply Chains

Herbert Meyr¹ and Hartmut Stadtler²

¹ Vienna University of Economics and Business Administration, Institute of Transport Economics and Logistics, Nordbergstraße 15, A–1090 Vienna, Austria

² Darmstadt University of Technology, Department of Production & Supply Chain Management, Hochschulstraße 1, 64289 Darmstadt, Germany

The SCOR-model presented in Sect. 2.2.2 is an excellent tool to analyze, visualize, and discuss the structure of the supply chain, and to reveal redundancies and weaknesses. It enables the formulation of structural changes and strategies to improve the performance of the supply chain as a whole.

However, when it comes to planning, the SCOR-model needs to be supplemented. To be able to identify the type of decision problems facing the supply chain and guide the selection of standard or specialized modules, models and algorithms for decision making, this chapter defines a “supply chain typology”, supporting the SCOR-model at level 2. Two examples illustrate the use of the typology and will be resumed in Chap. 4 in order to design planning concepts fitting the particular requirements of these two types of supply chains.

3.1 Motivation and Basics

In the early days of production planning and control a single concept and software system was applied in industry – material requirements planning (MRP) – irrespective of the many different requirements existing in diverse areas such as the production of foods or automobiles. On the other hand, if a production manager was asked whether the production system he manages is unique and requires special purpose decision-making tools, most probably the answer would be “yes”. As regards the type of decisions to be made, the truth lies somewhere in the middle of these two extremes. Abstracting from minor specialties usually reveals that there are common features in today’s production and distribution systems which require similar decision support and thus can be supported by the same software modules.

APS are much more versatile than MRP and ERP systems due to their modelling capabilities and different solution procedures (even for one module). Modules offered by a software vendor may still better fit one type of supply chain than another. So, it is our aim to outline a *supply chain typology* which allows to describe a given supply chain by a set of attributes which we feel might be important for decision-making and the selection of an APS. Attributes may have nominal properties (e.g. a product is storable or not), ordinal properties (e.g. an entity’s power or impact on decision-making is re-

garded higher or lower than average) or cardinal properties (i. e. the attribute can be counted, like the number of legally separated entities within a supply chain).

Attributes with a similar focus will be grouped into a peculiar category to better reveal the structure of our typology (see Tables 3.1 and 3.2). We will discriminate “functional” attributes to be applied to each organization, entity, member, or location of a supply chain as well as “structural” attributes describing the relations among its entities.

3.2 Functional Attributes

Functional attributes (see Tab. 3.1) of an entity are grouped into the four categories

- procurement type,
- production type,
- distribution type and
- sales type.

The **procurement type** relates to the *number* (few ... many) and *type of products* to be procured, the latter one ranging from standard products to highly specific products requiring special product know-how or production process know-how (or equipment). The following attribute depicts the *sourcing type*, better known by its properties: single sourcing, double sourcing and multiple sourcing. Single sourcing exists if there is a unique supplier for a certain product to be procured. In double sourcing there are two suppliers, each fulfilling a portion of demand for the product to be procured (e. g. 60% of the demand is fulfilled by the main supplier, 40 % by the second supplier). Sourcing contracts with suppliers are usually valid in the medium-term (e. g. a product's life cycle). Otherwise, products can be sourced from multiple suppliers. Next, the *flexibility of suppliers* with respect to the amounts to be supplied may be important. Amounts may either be fixed, have a lower or upper bound due to given contracts with suppliers or may be freely available. *Lead time* and *reliability of suppliers* are closely related. The lead time of a supplier defines the average time interval between ordering a specific material and its arrival. Usually, the shorter lead times are, the more reliable the promised arrival dates are. The *life cycles of components or materials* have direct impact on the risk of obsolescence of inventories. The shorter the life cycles are, the more often one has to care about substituting old materials with newer ones.

The **production type** is formed by many attributes. The two most prominent attributes are the *organization of the production process* and the *repetition of operations*. Process organization and flow lines represent well-known properties of the production process. Process organization requires

Table 3.1. Functional attributes of a supply chain typology

Functional attributes	
Categories	Attributes
Procurement type	number and type of products procured sourcing type flexibility of suppliers supplier lead time and reliability materials' life cycle
Production type	organization of the production process repetition of operations changeover characteristics bottlenecks in production working time flexibility etc.
Distribution type	distribution structure pattern of delivery deployment of transportation means loading restrictions
Sales type	relation to customers availability of future demands demand curve products' life cycle number of product types degree of customization bill-of-materials (BOM) portion of service operations

that all resources capable of performing a special task (like drilling) are located in the same area (a shop). Usually a product has to pass through several shops until it is finished. A flow shop exists if all products pass the shops in the same order, otherwise it is a job shop. A flow line exists in case resources are arranged next to each other corresponding to the sequence of operations required by the products to be manufactured on it. Usually capacities within a flow line are synchronized and intermediate inventories are not possible. Hence, for planning purposes a flow line can be regarded as a single entity.

The attribute *repetition of operations* has three broad properties, mass production, batch production and making one-of-a-kind products. In mass production the same product is generated constantly over a long period of time. In batch production several units of a given operation are grouped together to form a batch (or lot) and are executed one after the other. Several batches are loaded on a resource sequentially. At the start of a batch a setup is required, incurring some setup costs or setup time. When making one-of-a-

kind products which are specific to a (customer) order, special care is needed to schedule the many operations usually belonging to a (customer) order.

The influence of setup costs and setup times may be higher or lower. Therefore, their degree can further be specified by an optional attribute *changeover characteristics*. If setup costs (or times) even vary with respect to sequence of the batches or lots, “sequence dependent” changeover costs are given. If production capacity is a serious problem, the attribute *bottlenecks in production* tries to characterize why. In a multi-stage production system, the bottleneck machines may be stationary and known, or shifting (frequently) depending on the mix of demand. One way to increase capacity is to provide more working time (e.g. by means of overtime or additional shifts). The capability and lead times to adapt working time to changing demand pattern are described by the attribute *working time flexibility*. For further specifications of the production type see Schneeweiss (2002, pp. 10) and Silver et al. (1998, pp. 36).

The **distribution type** consists of the distribution structure, the pattern of delivery, the deployment of transportation means, and possible loading restrictions. The *distribution structure* describes the network of links between the factory (warehouse) and the customer(s). A one-stage distribution structure exists if there are only direct links between a factory (warehouse) and its customers. In case the distribution network has one intermediate layer (e. g. either central warehouses (CW) or regional warehouses (RW)) a two stage distribution structure is given. A three stage distribution structure incorporates an additional layer (e. g. CW and RW).

The *pattern of delivery* is either cyclic or dynamic. In a cyclic pattern, goods are transported at fixed intervals of time (e. g. round-the-world ship departures). A dynamic pattern is given if delivery is made depending on demand (for transportation). As regards the *deployment of transportation means* one can distinguish the deployment of vehicles on routes (either standard routes or variable routes depending on demand) and simply a given transportation capacity on individual links in the distribution network. It may even be possible to assume unlimited transportation capacities and to consider only a given cost function (e. g. based on a contract with a large third-party service provider). *Loading restrictions* (like the requirement of a full truck load) may form a further requirement.

The **sales type** of an entity in the supply chain largely depends on the *relation to its customers*. One extreme may be a downstream entity in the supply chain (with some kind of “agreement” regarding expected demands and an open information flow) while the other extreme may be a pure market relation with many competitors (e. g. auctions via Internet conducted by the purchasing departments of a large company). This attribute is closely related to the *availability of future demands*. These may be known (by contract) or have to be forecasted. The existence of (reliable) demand forecasts is best described by the length of the forecast horizon. Besides the general availability

of demand information, the shape of the *demand curve* is of interest. Demand for a specific product may, for example, be quite static, sporadic, or seasonal.

The typical length and the current stage of a *product's life cycle* significantly influences appropriate marketing, production planning and financial strategies. As regards the products to be sold one should discriminate the *number of product types* offered and the *degree of customization*. The latter one may range from standard products to highly specific products (in accordance to the products procured). In the light of mass customization some way in the middle becomes more and more important: constituting customer-specific products from a variety of product options and alternatives being offered. The attribute *bill-of-materials (BOM)* shows the way that raw materials and components are composed or decomposed in order to generate the final products. If raw materials are just changed in their sizes and shapes, a serial structure is given. In a convergent structure, several input products are assembled (or mixed) to form a single output product. Whereas in a divergent structure, a single input product is disassembled (or split) and several output products are the result. Of course, a structure of a mixture type – combining both convergent and divergent properties – is also possible.

Apart from selling tangible goods the *portion of service operations* is constantly growing (e.g. the training of a customer's personnel).

3.3 Structural Attributes

Structural attributes (see Table 3.2) of a supply chain are grouped into the two categories

- topography of a supply chain and
- integration and coordination.

Table 3.2. Structural attributes of a supply chain typology

Structural attributes	
Categories	Attributes
Topography of a supply chain	network structure degree of globalization location of decoupling point(s) major constraints
Integration and coordination	legal position balance of power direction of coordination type of information exchanged

As regards the **topography of a supply chain** the attribute *network structure* describes the material flows from upstream to downstream entities which are either serial, convergent, divergent, or a mixture of the three. Note that the network structure often coincides with the BOM. The *degree of globalization* ranges from supply chains operating in a single country to those with entities in several continents. Global supply chains not only have to take into account tariffs and impediments to trade as well as exchange rates varying over time, but also can profit from it. Also the *location of the decoupling point(s)* within the supply chain has to be mentioned. It is the first stage (or location) in the flow of materials where a further processing step or a change in the location of a product will only be executed with respect to a customer order (see also Sect. 1.2). Note, the decoupling point may differ between product groups. Starting with the most upstream location of a decoupling point we have engineer-to-order (with no make-to-stock at all), followed by manufacture-to-order of parts, then assemble-to-order and deliver-to-order. In a vendor managed inventory system a supplier even has to deliver-to-stock since there are no orders from the buyer to replenish inventories. The attribute *major constraints* gives an impression what the main bottlenecks of the supply chain (as a whole) are. These may, for example, be limited production capabilities of some member(s) or the limited availability of some critical materials.

Integration and coordination concerns the attributes legal position, balance of power, direction of coordination and type of information exchanged. The *legal position* of entities has already been mentioned. In case entities are legally separated, an inter-organizational supply chain exists, otherwise it is called intra-organizational. For intra-organizational supply chains it will be much easier to coordinate flows centrally than for inter-organizational supply chains. Also the *balance of power* within an inter-organizational supply chain plays a vital role for decision-making. A dominant member in the supply chain can act as a focal firm. On the other hand, we have a supply chain of equals, named a polycentric supply chain.

As regards information flows, several attributes may be considered. As an example consider the *direction of coordination*. It may be purely vertical or purely horizontal or a mixture of both. Vertical information flows comply with hierarchical planning. On the other hand, horizontal flows may exist between two adjacent entities within the supply chain which can easily and quickly make use of local information (e. g. to overcome the effects of a breakdown of a machine). Also the *type of information exchanged* between members influences planning (e. g. some entities may hesitate to reveal their manufacturing costs but are willing to provide information about available capacities).

While attributes describing a production type are generally accepted and validated today, a typology of the service sector is still in its infancy (for a state-of-the-art survey see Cook et al. (1999)). Also, the aforementioned attributes only provide a basis for a rough grouping of decision problems

which may be refined further according to the needs of a given SCM project. For this, special purpose typologies can be of help (e.g. for production processes concerning cutting and packing (Dyckhoff and Finke, 1992)). In some cases, this will also indicate that special purpose solution procedures may be needed, currently not provided by APS.

In order to reduce the burden associated with an (extensive) typology, one should bear in mind its aim. Since decision-making and decision support is of interest here, one might concentrate on activities to be performed on those products and services regarded most important (e.g. “A” products in an ABC-classification based on the annual turnover, see Silver et al. (1998, pp. 32)). Furthermore, attention can be focused on those activities which either have to be performed on potential bottlenecks along the supply chain or which affect critical performance criteria considerably (e.g. order lead-time).

Once a list of functional attributes has been established for each entity of a supply chain, it will show the degree of diversity existing in the supply chain. For partners having similar properties the choice of an appropriate decision-making tool (or module of an APS) can be made jointly, saving costs and time. In order to demonstrate the applicability of the above typology, it will be used in the following two sections for the different supply chain types *consumer goods industries* and *computer assembly*. We will come back to these two examples in Section 4.3 and in our case studies (Part IV).

3.4 Example for the Consumer Goods Industry

First, the typology will be applied for supply chains where consumer goods are produced and sold. Functional attributes are presented for the consumer goods manufacturing entity only. Structural attributes consider the supply chain as a whole comprising both manufacturers and retailers. Some attributes of our typology are not used within the example because they play only a minor role in supply chains of the consumer goods type. This kind of supply chain is considered again in the case study “*Food and Beverages*” (Chap. 20). Therefore, our description is rather detailed and affects additional proprietary attributes not mentioned explicitly in the above (universal) typology.

Table 3.3 summarizes the characteristics of the consumer goods supply chain. Since the products to be sold are the determining factor of our example, we start illustrating the *sales type* category.

Sales Type In the remainder we concentrate on the subset of consumer goods that comprises standard products with a low volume, weight and value per item (e.g. food, beverages, office supplies, or low tech electronics). Since quite often these standard products are just packaged in different sizes or under several brand names, some sort of “divergent” BOM is given. Thus, a

Table 3.3. Supply chain typology for the consumer goods industry

Functional attributes	
Attributes (see Table 3.1)	Contents
number and type of products procured	few, standard (raw materials)
sourcing type	multiple
supplier lead time and reliability	short, reliable
materials' life cycle	long
organization of the production process	flow line
repetition of operations	batch production
changeover characteristics	high, sequ. dep. setup times & costs
bottlenecks in production	known, stationary
working time flexibility	low
distribution structure	three stages
pattern of delivery	dynamic
deployment of transportation means	unlimited, routes (3^{rd} stage)
availability of future demands	forecasted
demand curve	seasonal
products' life cycle	several years
number of product types	hundreds
degree of customization	standard products
bill-of-materials (BOM)	divergent
portion of service operations	tangible goods

Structural attributes	
Attributes (see Table 3.2)	Contents
network structure	mixture
degree of globalization	several countries
location of decoupling point(s)	deliver-to-order
major constraints	capacity of flow lines
legal position	intra-organizational
balance of power	customers
direction of coordination	mixture
type of information exchanged	nearly unlimited

typical consumer goods manufacturer offers several hundreds of final items that are technologically related.

The final customer expects to find his preferred brand in the shelf of a grocery or electronics store. If the desired product is not available, he probably changes his mind and buys a comparable product of another manufacturer. This behaviour is due to the low degree of product differentiation predom-

inant in the consumer goods industry. Therefore, consumer goods manufacturers are forced to produce to stock by means of demand estimates.

Since the product life cycle of standard products typically extends over several years, a solid data basis for forecasting is available. However, demand for some products may be subject to seasonal influences (e.g. for ice cream or light bulbs) or price promotions.

If consumer goods are standardized, the emphasis of marketing has to be set on service level and price. Altogether, a strictly competitive market is given.

Distribution Type Consumer goods are distributed via wholesalers and/or retailers to the final customers. The distribution network of a consumer goods manufacturer quite often comprises three distribution stages (see Fleischmann (1998) and Fig. 3.1).

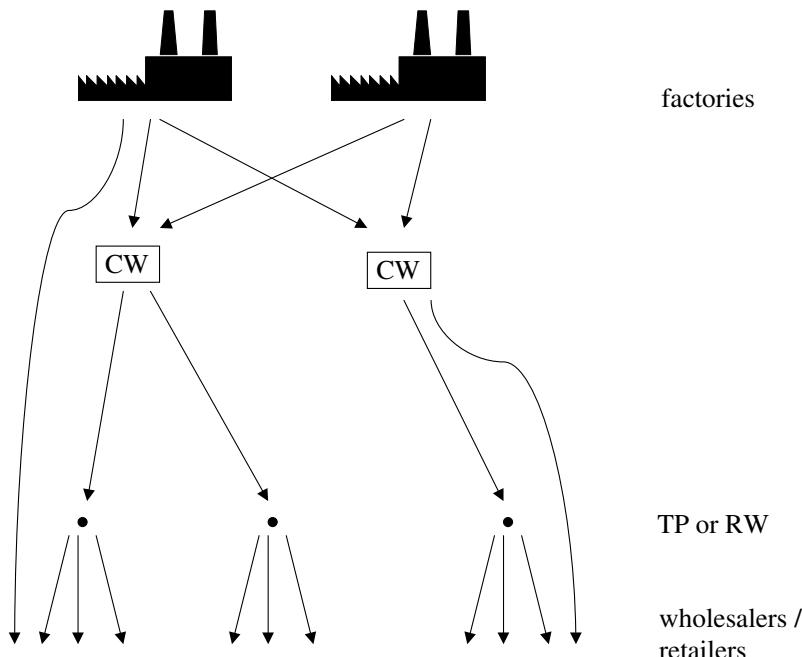


Fig. 3.1. Three-stage distribution system

The product programme of the manufacturer is supplied by one or a few factories. Thereby, some product types may be produced in more than one site. The finished goods can temporarily be stored in a few CWs, each of them offering the whole range of products. Large orders of the manufacturer's customers (i.e. wholesalers, retailers or department stores) can be delivered directly from the factory or CW to the respective unloading point.

Since most orders are of rather small size and have to be transported over long distances, a further distribution stage consisting of RWs or stock-less transshipment points (TP) is often used. The customers in the vicinity (at most 100 km radius) of such a RW/TP are supplied in one-day tours starting from this RW/TP. Over the (typically) long distance between the CW and the RW/TP all orders of the respective region are bundled (usually by third-party service providers) so that a high transport utilization is achieved.

As opposite to RWs, no stock is held in TPs, thus causing lower inventory holding, but higher transportation costs due to the higher delivery frequency. A similar distribution structure may be used by major sales chains which replenish their (large number of) department stores from their own retail CWs.

Production Type Production of consumer goods often comprises only one or two production stages, e. g. manufacturing and packaging. On each production stage one or a few parallel (continuous) production lines (flow lines) are organized in a flow shop. A line executes various operations. But since these operations are strictly coordinated, each line may be planned as a single unit. The lines show a high degree of automation and are very capital intensive. Because of this automation, however, short and reliable throughput times can be achieved.

The capacity of the production lines is limited and they are usually highly utilized. Therefore, they represent potential bottlenecks. For the handling of the lines, few but well-trained operators are necessary. A short-term expansion of working time is normally not possible. The working time of the whole team supervising a line has to be determined on a mid-term time range. However, in many companies the lines are already operating seven days a week, 24 hours a day.

As mentioned above, there are a lot of final items. But these are often technologically related and can be assigned to a few *setup families*. Changeovers between items of the same family are negligible. But changeovers between items of different families cause high setup costs and setup times. Therefore, batch production is inevitable. The degree of these costs and times may vary notably with respect to the family produced last on the same line (sequence dependent setup times and costs).

Procurement Type Consumer goods frequently have a rather simple BOM. In these cases only few suppliers have to be coordinated. As long as not sophisticated components, but mainly standard products (e. g. raw materials) are needed, procurement is not really a problem. The lead time of raw materials is short and reliable. The life cycles of these materials are rather long. Therefore, mid- and long-term contracts and cooperations ensure the desired flow of raw materials from the suppliers to the manufacturer. Nevertheless, if there should be any unexpected problems in sourcing material, because of

the high degree of standardization it is quite easy to fall back on alternative suppliers on the short-term (multiple sourcing).

Topography of the Supply Chain The production network (maybe several sites producing the same product), the distribution network of the manufacturer and possibly the distribution network of large wholesalers/retailers contain both divergent and convergent elements thus forming a network structure of the mixture type. Production and distribution networks usually extend over several countries, sometimes even over multiple continents. Since products are made to stock, the decoupling point of the manufacturer is settled in CWs or RWs, from which goods are delivered to order. While procurement is quite unproblematic, the limited capacity of the flow lines is the major constraint of the whole supply chain.

Integration and Coordination Because of the low differentiation the balance of power is shifted towards the customers, i.e. the retailers. As regards the consumer goods manufacturing entity, there is a strong need for intra-organizational coordination. Several organizational units of the same company (e.g. order management, sales, manufacturing, procurement) have to exchange information horizontally. Furthermore, the central planning unit has to coordinate the bulk of decentral units by sending directives and gathering feedback, thus inducing heavy vertical information traffic. Since all of these units belong to the same company, information should be freely available.

In addition, new logistical concepts of SCM result in special emphasis on inter-organizational relations within the supply chain, particularly on the interface between consumer goods manufacturers and large retailers. Current trends can be outlined as follows:

- The flow of information between the manufacturers and retailers is improved by EDI or WWW connections.
- Short delivery cycles (with rather small quantities) are established in order to closely connect the material flow with the demand of final customers (*Continuous Replenishment / Efficient Consumer Response*).
- Traditional responsibilities are changed. Large retailers abstain more and more from sending orders to their suppliers, i.e. the consumer goods manufacturers. Instead they install consignation stores whose contents are owned by their suppliers until the goods are withdrawn by the retailer. A supplier is responsible for filling up his inventory to an extent which is convenient for both the supplier and the retailer. As already mentioned, such an agreement is called *vendor managed inventory* (VMI).

3.5 Example for Computer Assembly

Now a second application of the above general typology will be presented. In order to offer a quite contrary example, a *computer assembly* supply chain has been chosen. A particular instance of this type of supply chain will be described in the case study in Chap. 21. Table 3.4 summarizes the properties of that type so that a direct comparison with the consumer goods type (Table 3.3) is possible. Again, functional attributes are only shown for the computer manufacturing entity, whereas structural attributes characterize the interrelations between different entities of the supply chain.

Sales Type Computers have a strictly convergent BOM. The system unit is assembled from several components like the housing, the system board, the Central Processing Unit (CPU), hard disk(s), a sound card etc. The degree of customization varies between the two extremes

- *standard products* with *fixed configurations*, i.e. only some predefined types are offered. Customers merely can choose between these types, but no changes or extensions (at least at the system unit) are possible.
- *customized products* which are completely *configurable*. In this case the customer specifies which components he wants to get from what supplier or at least the options of the components he wants to get (like a “slow” CPU, but a “high-end” graphics card). The manufacturer tests whether the requested configuration is technically feasible and calculates the price. Because of the ability to combine many different components – again obtainable from several alternative suppliers – an incredibly large number of possible final items is given.

Of course, the usual practice is somewhere in between. For example, some standard computers are defined with a few options like additional RAM or a CDRW instead of a DVD. Or only a limited number of hard disks, CPUs, housings etc. is offered and the customer can only choose between these alternatives. The corresponding final items then have already been tested for technical feasibility and prices have been assigned. In the following, just the two extreme cases are considered.

The computer itself consists of the system unit and some accessories like cables, software, a manual, or a keyboard. A typical order of a customer comprises several order lines for different product families (e.g. desktop computers, servers, notebooks) and external units (peripherals) like speakers, monitors, printers and so on. If customers call for delivery of “complete orders”, all order lines of an order have to be delivered simultaneously to the customer (e.g. because printers without computers are useless for the customer). Thus, the BOM comprises several stages like the order itself (consisting of several order lines for computers of different product families and peripherals), computers (system unit and accessories) and system units (housing, main board,

Table 3.4. Supply chain typology for computer assembly

Functional attributes	
Attributes (see Table 3.1)	Contents (fixed / configurable)
number and type of products procured	many, standard & specific
sourcing type	multiple
supplier lead time and reliability	short & long, unreliable
materials' life cycle	short
organization of the production process	flow shop & cellular
repetition of operations	larger / smaller batches
changeover characteristics	irrelevant
bottlenecks in production	low importance
working time flexibility	high
distribution structure	two stages
pattern of delivery	dynamic
deployment of transportation means	individual links
availability of future demands	forecasts & orders
demand curve	weakly seasonal
products' life cycle	few months
number of product types	few / many
degree of customization	standard / customized
bill-of-materials (BOM)	convergent
portion of service operations	tangible goods

Structural attributes	
Attributes (see Table 3.2)	Contents
network structure	mixture
degree of globalization	several countries
location of decoupling point(s)	assemble-/configure-to-order
major constraints	material
legal position	inter- & intra-organizational
balance of power	suppliers & customers
direction of coordination	mixture
type of information exchanged	forecasts & orders

etc.). Some computer manufacturers are also responsible for the assembly of the system board from several components like the Printed Circuit Board, chips, etc.

There is a low product differentiation. Price, speed and reliability of the promised due dates are the key performance indicators. The planned order lead times vary – dependent on the product family – between a few days and a few weeks. Because of technological improvements a fast changing environment has to be mastered. Due to the short product life cycles of only

a few months, there is a high risk of obsolescence. Total customer demand is known for the next few days only. For the further future, the probability of having fully specified customer orders on hand decreases drastically. Then, (not yet known) customer orders have to be anticipated by forecasts. Demand is weakly influenced by seasonal effects like the Christmas business or year's end business of authorities.

Distribution Type Typical customers are system integrators offering overall solutions for big corporate customers, medium and small business customers, and consumer market stores which sell standard computers (“consumer PCs”) to private customers. In this case, often a two-stage distribution system is used where computers and peripherals are merged by logistics service providers in distribution centres to constitute a complete order. Sometimes manufacturers sell directly to private customers via the Internet. Then, a parcel service is responsible for the delivery to the final customer. It is interesting to note that in the “complete order” case the last stage of the BOM is settled in a distribution centre.

Production Type The main production processes are the “assembly of the system board”, the “assembly of the system unit”, the “loading of the software”, a final “testing” and the “packing” (assembly of the computer). The “assembly of the system board” may be done in-house or in an additional upstream factory, also owned by the computer manufacturing company. But system boards may also be bought from external suppliers. Anyway, system boards are assembled on highly automated flow lines with very short throughput times.

The key process “assembly of the system unit” is also done in flow line organization, but manually. Sometimes a cellular organization is given. Despite of the manual work and the possibly high degree of customization, processing times are stable. Only low skilled personnel is necessary. Therefore, additional staff can be hired on the short term and working time flexibility is high. Fixed configurations can be assembled in large batches. Open configurations, however, have to be produced in small batches because of the individuality of customer demand. Nevertheless, due to the nature of the setup processes (e.g. providing components of the next batch in parallel to the assembly of the current batch), there are no significant setup costs or times. Altogether, serious bottlenecks in production are missing and production capacity does not play a critical role.

Procurement Type Because of the rather simple production processes, the key competences of a computer manufacturer actually are the synchronization of suppliers and sales and order management, respectively. Thousands of components, accessories and external units have to be purchased and must be

right in place before the assembly or delivery. The products procured are very inhomogeneous. Standard components as well as highly specific components have to be ordered. Supplier lead times range from a few days to several months and are most of the time very unreliable.

Just as it is the case for computers, life cycles of components are often very short due to technological progress. So there is also a high risk of obsolescence at the supply side. Because of mid- to long-term contracts with critical suppliers, there may exist both upper and lower bounds on supply quantities. Such contracts are particularly important when supply shortages can occur and multiple sourcing is not possible, i.e. when the balance of power is shifted towards the supplier.

For some components like hard disks multiple sourcing is common practice. These components are bought from several suppliers. Thus, at least for standard products the computer manufacturer is free to substitute components and to increase orders for alternative suppliers if the one originally planned runs into trouble. Also “downgrading” of components is a practicable (but expensive) way to deal with shortage situations: in this case, a lower value component – being requested but out of stock – is replaced by an alternative component with higher value. For example, a 3.2 GHz CPU is assembled instead of a 3.0 GHz CPU because the requested lower value component is not in stock any more. Since the price has been fixed earlier and cannot be re-adjusted, the customer does not need to be informed.

Topography of the Supply Chain The network structure is of a mixture type: lots of suppliers (of components, accessories and peripherals) are linked with a few assembly sites (for system boards and several product families), a few distribution centres, and with a large number of customers (of different types as described above). The whole network may extend over several countries.

Nowadays, most computer manufacturers have successfully shifted their deliver-to-order decoupling point upstream in order to reduce the risky and expensive finished product inventory. In case of fixed configurations, an assemble-to-order decoupling point is now common practice, i.e. computers are only assembled if a respective customer order for a standard configuration has arrived. For open configurations an engineer-to-order decoupling point is given, i.e. an incoming customer request has also to be checked for technological feasibility and an individual price has to be set. Shifting the decoupling point upstream reduces finished product inventory and hedges against demand uncertainty, but also increases order lead times (as long as throughput times are not simultaneously decreased). The performance of the supply chain is primarily limited by constraints on material supply and not by scarce assembly capacity.

Integration and Coordination Both inter- and intra-organizational members participate at computer assembly supply chains. So there is a need for collaboration between legally independent companies (e.g. by exchanging demand information like forecasts and orders horizontally) as well as a need for vertical coordination of different organizational units of the computer manufacturing company itself. Thus, the direction of coordination is of a mixture type.

Both suppliers and customers may have a high power within such supply chains. The power is extremely high for suppliers that reside in some sort of monopoly or oligopoly like vendors of operating systems or CPUs. As shown above, long-term contracts may ensure the desired flow of critical components from these suppliers.

We will next time come back to the *consumer goods manufacturing* and *computer assembly* types of supply chains in Section 4.3. There, the particular planning requirements of these two supply chain types and planning concepts fitting them are derived from the attributes shown above.

References

- Cook, D. P.; Goh, C.-H.; Chung, C. H. (1999) *Service typologies: A state of the art survey*, Production and Operations Management, Vol. 8, No. 3, 318–338
- Dyckhoff, H.; Finke, U. (1992) *Cutting and packing in production and distribution: A typology and bibliography*, Heidelberg
- Fleischmann, B. (1998) *Design of freight traffic networks* in: Fleischmann, B.; Nunen, J. A. E. E. van; Speranza, M. G.; Stähly, P. (Eds.) *Advances in Distribution Logistics*, Lecture Notes in Economics and Mathematical Systems, Vol. 460, Berlin et al., 55–81
- Schneeweiss, C. (2002) *Einführung in die Produktionswirtschaft*, 8th ed., Berlin et al.
- Silver, E. A.; Pyke, D. F.; Petersen, R. (1998) *Inventory management and production planning and scheduling*, 3rd ed., New York et al.

4 Advanced Planning

Bernhard Fleischmann¹, Herbert Meyr², Michael Wagner³

¹ University of Augsburg, Department of Production and Logistics,
Universitätsstraße 16, 86135 Augsburg, Germany

² Vienna University of Economics and Business Administration, Institute of
Transport Economics and Logistics, Nordbergstraße 15, A-1090 Vienna, Austria

³ Paul Hartmann AG, Corporate Materials Management, P.O.Box 1420,
89504 Heidenheim, Germany

4.1 What is Planning?

Why planning? Along a supply chain hundreds and thousands of individual decisions have to be made and coordinated every minute. These decisions are of different importance. They comprise the rather simple question “*Which job has to be scheduled next on a respective machine?*” as well as the very serious task whether to open or close a factory. The more important a decision is, the better it has to be prepared.

This preparation is the job of *planning*. Planning supports decision-making by identifying alternatives of future activities and selecting some good ones or even the best one. Planning can be subdivided into the phases (see Domschke and Scholl (2003), pp. 26)

- *recognition* and *analysis* of a decision problem,
- definition of *objectives*,
- *forecasting* of future developments,
- *identification* and *evaluation* of feasible activities (*solutions*), and finally
- *selection* of good solutions.

Supply chains are very complex. Not every detail that has to be dealt with in reality can and should be respected in a plan and during the planning process. Therefore, it is always necessary to abstract from reality and to use a simplified copy of reality, a so-called *model*, as a basis for establishing a plan. The “art of model building” is to represent reality as simple as possible but as detailed as necessary, i. e. without ignoring any serious real world constraints.

Forecasting and *simulation models* try to predict future developments and to explain relationships between input and output of complex systems. However, they do not support the selection of one or a few solutions that are good in terms of predefined criteria from a large set of feasible activities. This is the purpose of *optimization models* which differ from the former ones by an additional *objective function* that is to be minimized or maximized.

Plans are not made for eternity. The validity of a plan is restricted to a predefined *planning horizon*. When reaching the planning horizon, at the

latest, a new plan has to be made that reflects the current status of the supply chain. According to the length of the planning horizon and the importance of the decisions to be made, planning tasks are usually classified into three different planning levels (see Anthony (1965)):

Long-term planning: Decisions of this level are called *strategic decisions* and should create the prerequisites for the development of an enterprise/supply chain in the future. They typically concern the design and structure of a supply chain and have long-term effects, noticeable over several years.

Mid-term planning: Within the scope of the strategic decisions, mid-term planning determines an outline of the regular operations, in particular rough quantities and times for the flows and resources in the given supply chain. The planning horizon ranges from 6 to 24 months, enabling the consideration of seasonal developments, e. g. of demand.

Short-term planning: The lowest planning level has to specify all activities as detailed instructions for immediate execution and control. Therefore, short-term planning models require the highest degree of detail and accuracy. The planning horizon is between a few days and three months. Short-term planning is restricted by the decisions on structure and quantitative scope from the upper levels. Nevertheless, it is an important factor for the actual performance of the supply chain, e. g. concerning lead-times, delays, customer service and other strategic issues.

The last two planning levels are called *operational*. Some authors call the second level *tactical* (e. g. Silver et al. (1998), Chap. 13.2), but as this notion has several contradictory meanings in the literature, it is not used in this book.

A naive way of planning is to look at the alternatives, to compare them with respect to the given criteria, and to select the best one. Unfortunately, this simple procedure encounters, in most cases, three major difficulties:

First, there are often several criteria which imply conflicting objectives and ambiguous preferences between alternatives. For example, customer service ought to be as high as possible while – at the same time – inventories are to be minimized. In this case no “optimal” solution (accomplishing both objectives to the highest possible degree) exists. A common way to deal with this *multi-objective decision problem* is to set a minimum or maximum satisfaction level for each objective except for one that will be optimized. In the above example one may try to minimize inventories while guaranteeing a minimum customer service level. Another useful way to handle multiple objectives consists in pricing all objectives monetarily by revenues or costs and maximizing the resulting *marginal profit*. However, not every objective can be expressed in monetary values, e. g. the customer service. A more general way is to define scale values or scores for every objective and to aggregate them into a weighted sum. A danger of this procedure is that it yields pretended “optimal” solutions which strongly depend on the arbitrary weights. An APS

supports each of these procedures in principle. The case studies in Part IV give examples of some relevant modelling features of the *i2*, *PeopleSoft* and *SAP* systems.

The second difficulty is caused by the huge number of alternatives that are predominant in supply chain planning. In case of continuous decision variables, e.g. order sizes or starting times of a job, the set of alternatives is actually infinite. But also for discrete decisions, e.g. the sequence of several jobs on a machine, the number of alternatives may be combinatorially large (see Chap. 10). In these cases it is impossible to find an optimal solution by enumeration of all alternatives, and even a feasible solution may be difficult to find. In this situation, mathematical methods of *operations research (OR)* should support the planning process. Some methods are able to determine an exact optimal solution, e.g. Linear Programming (LP) or network flow algorithms, but for most combinatorial problems only near-optimal solutions can be computed by *heuristics*, e.g. local search. The success of these methods also depends on the way a problem is modeled. As examples, for some important types of optimization models the capabilities of OR methods are shown in the Supplement (Part VI).

The third and probably hardest difficulty is dealing with uncertainty. Planning anticipates future activities and is based on data about future developments. The data may be estimated by forecast models, but there will be a more or less important forecast error. This error reduces the availability of products and therefore reduces the customer service a company offers. For improvement of the service safety stocks can be utilized which buffer against demands exceeding the forecast. However, that is not the only way to tackle uncertainty.

Nearly always, reality will deviate from the plan. The deviation has to be controlled and the plan has to be revised if the discrepancy is too large. Planning on a *rolling horizon basis* is an implementation of this plan-control-revision interaction. The planning horizon (e.g. one year) is divided into periods (e.g. months). At the beginning of January a plan is made that covers January to December. But only the first period, the so-called *frozen period*, is actually put into practice. At the beginning of the second period (February) a new plan is made considering the actual developments during the first period and updated forecasts for the future periods. The new planning horizon overlaps with the previous one, but reaches one period further (until the end of January of the next year; see Fig. 4.1) and so on.

This procedure is a common way of coping with uncertainty in operational planning both in classical planning systems and in APS. A more efficient way of updating the plans is *event-driven planning*: A new plan is not drawn up in regular intervals but in case of an important event, e.g. unexpected sales, major changes in customer orders, breakdown of a machine etc. This procedure requires that all data which are necessary for planning, e.g. stocks, progress of work etc., are updated continuously so that they are available at

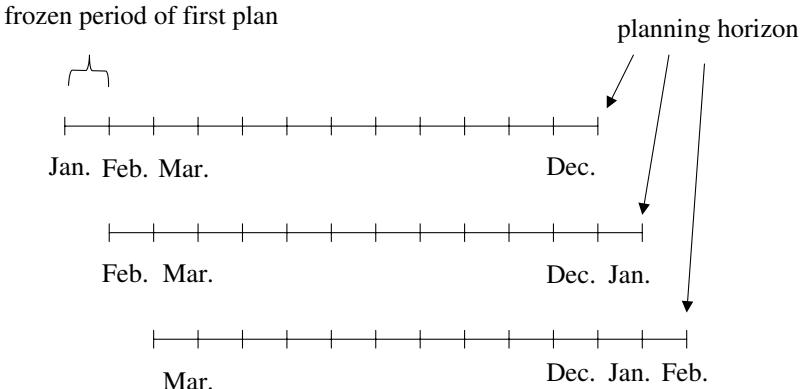


Fig. 4.1. Planning on a rolling horizon basis

any arbitrary event time. This is the case for an APS which is based on data from an Enterprise Resource Planning (ERP) system.

There are three main characteristics of APS:

- *Integral planning* of the entire supply chain, at least from the suppliers up to the customers of a single enterprise, or even of a more comprehensive network of enterprises;
- *true optimization* by properly defining alternatives, objectives, and constraints for the various planning problems and by using optimizing planning methods, either exact ones or heuristics (see Fleischmann and Meyr (2003, Chap. 9.4));
- *a hierarchical planning system* (see Schneeweiss (2003) and Chap. 1) which is the only framework permitting the combination of the two preceding properties: Optimal planning of an entire supply chain is neither possible in form of a monolithic system that performs all planning tasks *simultaneously* – this would be completely impracticable – nor by performing the various planning tasks successively – this would miss optimality. Hierarchical planning is a compromise between practicability and the consideration of the interdependencies between the planning tasks.

Note that the traditional material requirements planning (MRP) concept (see Orlicky (1975)) which is implemented in nearly all ERP systems does not have any of the above properties: It is restricted to the production and procurement area, does not optimize and in most cases even not consider an objective function, and it is a successive planning system.

The main idea of hierarchical planning is to decompose the total planning task into *planning modules*, i. e. partial plans, assigned to different levels where every level covers the complete supply chain but the tasks differ from level to level (see e. g. Miller (2001)): On the upmost level, there is only one module, the development of an enterprise-wide, long-term but very rough

plan. The lower the levels are, the more restricted are the supply chain sections covered by one plan, the shorter is the horizon and the more detailed is the plan. Plans for different supply chain sections on one level are coordinated by a more comprehensive plan on the next upper level in a hierarchical structure (see Fig 4.2).

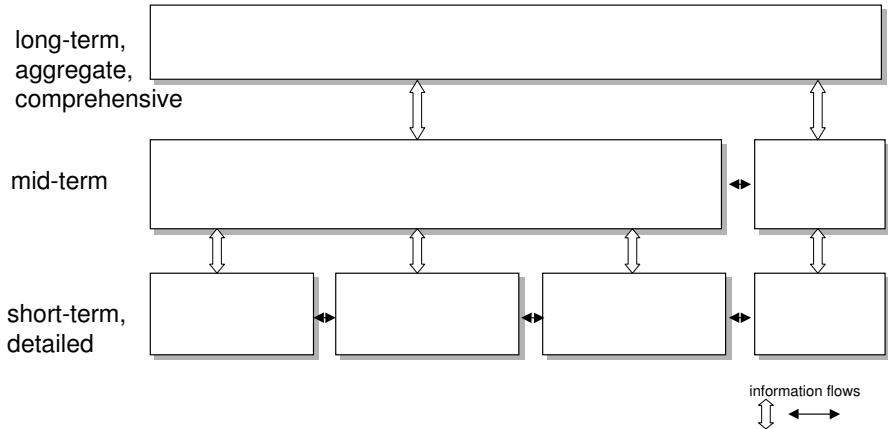


Fig. 4.2. Hierarchy of planning tasks

The increasing (resp. decreasing) degree of detail is achieved by disaggregating (resp. aggregating) data and results when going down (resp. up) in the hierarchy. *Aggregation* concerns

- products, aggregated into groups,
- resources, aggregated into capacity groups, and
- time: periods, aggregated into longer ones.

The modules are linked by vertical and horizontal information flows. In particular, the result of a higher planning module sets restrictions for the subordinate plans, and the results of the latter yield feedback information on performance (e.g. costs, lead-times, utilization) to the higher level. The design of a *hierarchical planning system* (HPS) requires a careful definition of the modular structure, the assignment of planning tasks to the modules, and the specification of the information flows between them. Usually, an HPS works with a rolling horizon, where sophisticated coordination of the planning intervals and horizons on the different levels has been suggested in literature (e.g. Hax and Meal (1975), Stadtler (1986)).

Planning takes into account future developments, identifies alternatives for future activities and provides directives for their implementation. However, the decisions themselves usually are put into practice outside of the planning system. Because of this separation and because of the above mentioned planning intervals, a time gap between planning and the final imple-

mentation has to be bridged which leaves room for unforeseen events. For this reason and in order to keep planning systems manageable, usually not all decisions are prepared in the planning system itself, but there is still some degree of freedom left open (to more precisely specify or revise a plan) until the final *execution* takes place. For the remainder of the book “execution” is defined as the starting and subsequent controlling of activities that have to be carried out immediately. Thus, in contrast to instructions prepared by a planning system, decisions for execution cannot be revised.

An “*execution system*” receives the decisions of a higher-ranked planning system, checks whether the assumptions underlying the plan are still valid, puts in further details when necessary (like assigning transport activities to production orders) and – in case no unexpected events have occurred – brings the overall decisions to final execution. However, if unforeseen events like machine breakdowns etc. have happened, it is up to the execution system to recognize this status and to react immediately. Minor problems may be solved by the execution system directly. If serious problems occur, an “alert” has to be sent back to the planning system, thus initiating an extraordinary re-planning. This event-driven planning simplifies the use of an HPS and makes it more flexible. A prerequisite is a communication system that guides alerts (see Chap. 13) on “events” to the relevant planning levels and tasks. Moreover, the result of one planning task can also generate alerts for other plans.

APS try to “computerize” planning. This might incur some problems for many human planners because they are afraid of being substituted by machines. This fear is based upon three major advantages of APS: they visualize information, reduce planning time, and allow an easy application of optimization methods. However, modelling is always a relaxation of reality. Therefore, human knowledge, experience, and skill is yet required to bridge the gap between model and reality. Planning systems, no matter how *advanced* they might be, remain decision support systems, i. e. they support human decision-makers. Also, in event-driven planning it is usually the human planner (at the interface between the execution and planning system) who decides whether a plan is to be revised. Finally, each planning module requires a human “owner” who is responsible for its function, data, and results.

4.2 Planning Tasks Along the Supply Chain

The whole Supply Chain Network can be split into internal supply chains for every partner in the network, each consisting of four main supply chain processes with substantially different planning tasks. *Procurement* includes all subprocesses which provide resources (e. g. materials, personnel etc.) necessary for production. The limited capacity of resources is the input to the *production* process which may consist of various subprocesses. The *distribu-*

bution bridges the distance between the production site and the customers, either retailers or other enterprises processing the products further. All of the above logistical processes are driven by demand forecasts and/or order figures determined by the *sales* process.

Supply Chain Planning Matrix

The *Supply Chain Planning Matrix* (SCP-Matrix, see Rohde et al. (2000)) classifies the planning tasks in the two dimensions “planning horizon” and “supply chain process”. Fig. 4.3 shows typical tasks which occur in most supply chain types, but with various contents in the particular businesses. In Fig. 4.3 the long-term tasks are shown in a single box to illustrate the comprehensive character of strategic planning. The other boxes represent the matrix entries, but do not correspond exactly to the planning modules of an HPS. The latter may contain only parts of a box – e.g. on the short-term level the planning tasks can be decomposed according to further dimensions like factory sites or product groups – or combine tasks of several boxes. This is a question of the design of the HPS as mentioned in Sect. 4.1. The SCP-Matrix can also be used to position the *software modules* of most APS vendors (see Chap. 5). The construction of an HPS from the software modules of an APS is discussed in Part IV.

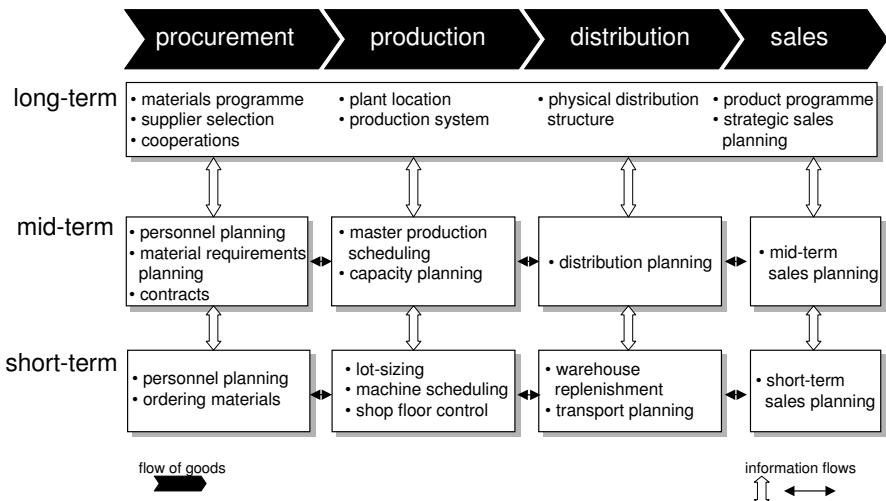


Fig. 4.3. The Supply Chain Planning Matrix

Long-term Planning Tasks

Product Programme and Strategic Sales Planning The decision about the product programme a firm wants to offer should be based on a long-range forecast which shows the possible sales of the whole product range. Such a forecast includes dependencies between existing product lines and future product developments and also the potential of new sales regions. It is often necessary to create different scenarios depending on the product programme decision. Long-range forecasts consider information on product-life-cycles and economical, political, and competitive factors. As it is not possible to estimate long-range sales figures for each item, the products need to be aggregated into groups of items sharing common sales and production characteristics. Marginal profits of potential sales and fixed costs for assets have to be considered in the objective function of the product programme optimization problem.

When a manufacturing member of a supply chain thinks about introducing a new product (group), it has to determine the location of the decoupling points with respect to the specific customers or markets considered. The location of the decoupling point is predefined by the (strategic) decision on the order lead-times (time between order entry and planned delivery) that probably will be accepted by the customers and therefore should be assigned to a respective product / market combination (see Hoekstra and Romme (1991, Chap. 1.5)). The shorter the order lead-time is, the better customers will be satisfied, but – on the other hand – the more downstream the decoupling point has to be settled. As we have seen in the previous chapter (p. 79), this entails some increased demand uncertainty for higher-value products.

Physical Distribution Structure As more and more companies concentrate their production capacities because of high investments in machining, the distance between the production facility and customers and the respective distribution costs increase. Such trends and a changing environment require a reorganization of the distribution system. The physical structure comprises the number and sizes of warehouses and cross docking points including the necessary transportation links.

Typical inputs for the decision are the product programme and the sales forecast, the planned production capacity in each plant, and the underlying cost structure. The objective is to minimize the long-term costs for transportation, inventory, handling, and investments in assets (e.g. warehouses, handling facilities etc.). The question, whether the transports are performed by one's own fleet of vehicles or a third-party carrier, is very closely related to the decision on the physical distribution system. For this reason, the two decision types should be integrated into one model.

Plant Location and Production System Long-term changes in product programmes or sales figures require to review the existing production capac-

ties and locations. Furthermore, the continuous improvement of production technologies leads to new prerequisites. Therefore, the production and decision systems need to be verified. Usually, decisions on plant locations and the distribution structure are made together. They are based on long-term forecasts and production capacities available (without consideration of single machines). Planning the production system means organizing a single production plant, i. e. designing the layout of the plant and the resulting material flows between the machines.

Materials Programme and Supplier Selection The materials programme is often directly connected to the product programme because the final products consist of some predefined components and raw materials. Sometimes different materials could be used alternatively for the same purpose. In order to select one of them for the materials programme, one should consider price (including possible quantity discounts), quality, and availability.

As A-class materials (see e. g. Silver et al. (1998) for an introduction to the ABC-analysis) cause the biggest part of procurement costs, it is reasonable to source those parts through special supply channels. Therefore, the suppliers should be rated according to quality, service, and procurement costs.

Cooperations Further reduction of procurement costs is often achieved by strategic cooperations with suppliers of A-class items. Planning and evaluation of collaboration concepts gain importance because no longer companies but whole supply chains compete against each other. These concepts include simultaneous reduction of inventories and backorders using ideas like *VMI* (*vendor managed inventory*), *EDLP* (*every-day-low-price strategies*), and *JIT* (*just-in-time*) supply. While the above cooperation concepts concern day-to-day operations, *simultaneous engineering* and *consolidation centres* set strategic frames for the daily procurement processes.

Mid-term Planning Tasks

Mid-term Sales Planning The main task in mid-term sales planning is forecasting the potential sales for product groups in specific regions. As the forecasts are input to master production scheduling, the products are grouped according to their production characteristics (e. g. preferred resources, change-over times etc.). The forecast is usually calculated on a weekly or monthly basis for one year or less. It includes the effects of mid-term marketing events and promotions on sales. The necessary safety stocks for finished products are mainly determined by the quality of the forecast. Therefore, it is reasonable to set them on the basis of the forecast error which has to be calculated in the forecasting procedure.

Distribution Planning Mid-term distribution planning comprises the planning of transports between the warehouses and determination of the necessary stock levels. A feasible plan fulfils the estimated demand (forecasts) and considers the available transportation and storage capacities while minimizing the relevant costs. Inventory holding and transportation costs are elements of the objective function. The planning horizon consists of weekly or monthly buckets. Therefore, the underlying model only considers aggregated capacities (e.g. available truck capacity and not single trucks). The distribution plan could also state the usage of the own fleet and the necessary capacity which must be bought from a third-party carrier.

Master Production Scheduling and Capacity Planning The result of this planning task shows how to use the available production capacity of one or more facilities in a cost efficient manner. Master production scheduling (MPS) has to deal with seasonal fluctuations of demand and to calculate a frame for necessary amounts of overtime. As the plan is based on families of products and weekly or monthly time buckets, it does not consider single production processes. The objective is to balance the cost of capacity against the cost of (seasonal) inventories. If more than one production facility is considered, the transportation costs between the locations have to be included in the objective function.

Personnel Planning Capacity planning provides a rough cut overview of the necessary working time for finished products. Personnel planning has to calculate the personnel capacity for components and other production stages which have to be passed before the final assembly of the products. This planning step considers the specific know how of personnel groups and their availability according to labour contracts. If not enough employees are available to fulfil the work load, personnel planning shows the necessary amount of additional part time employees.

Material Requirements Planning As MPS plans only finished products and critical materials (concentration on bottlenecks), material requirements planning (MRP) has to calculate the production and order quantities for all remaining items. This could be done by the traditional MRP-concept (see Orlicky (1975)) which is available in most ERP-systems or by stochastic inventory control systems. Whereas the MRP-concept is suitable for rather important (but non-bottleneck) materials and A-class components, stochastic inventory systems are adequate for C-class items. The calculation of material requirements should support lot-sizing decisions for every item in the bill-of-materials (BOM) and consider the dependencies between the lots on different levels of the BOM. Mid-term planning sets frames for weekly or monthly order quantities and safety stock levels which ensure the desired service level for production.

Contracts On basis of the weekly or monthly requirements obtained from MRP, basic agreements with A-class suppliers can be made. Such contracts set the price, the total amount, and other conditions for the materials to be delivered during the next planning horizon.

Short-term Planning Tasks

Short-term Sales Planning In make-to-stock environments the short-term sales planning comprises the fulfilment of customer orders from stocks. Therefore, the stock on hand can be partitioned in committed stocks and the *available-to-promise* (ATP) quantity. If a customer requests a product, the sales person checks online whether the quantity could be fulfilled from ATP and turns the requested amount in committed stock. For customer inquiries on the availability of products in future periods the ATP quantity is calculated by adding stock on hand and planned production quantities. The *capable-to-promise* (CTP) functionality is an extension of the traditional ATP task which has the additional option of creating new production orders.

Warehouse Replenishment, Transport Planning While the mid-term distribution planning suggests weekly or monthly transportation quantities for product families, the short-term warehouse replenishment particularizes this plan in daily quantities for single products. This time-phased deployment schedule considers detailed transportation capacities (e.g. available trucks) and actual customer orders or short-term forecasts. Planned or actual production quantities set the frame for the transportation plan and also restrict the possible degree of customer service. Every day the planned truck loads have to be deployed to customer locations according to a cost-minimizing routing.

Transports occur not only in the distribution process, but also as part of the procurement and may be controlled by either the supplier or the receiver. In the latter case, transport planning is necessary on the procurement side as well, and the transport processes have to be considered also in the mid-term and long-term levels of procurement planning.

Lot-sizing and Machine Scheduling, Shop Floor Control Short-term production planning comprises the determination of lot-sizes and the sequences of the lots on the machines. Lot-sizing has to balance the costs of changeovers and stock holding with respect to dependencies between different products. These lots are scheduled according to their due dates and the available capacity with minutely accuracy. Both tasks can independently be executed if the changeovers are not dependent on the sequence of the products. As interruptions or delays are common in complex production environments, the shop floor has to be controlled actively and orders have to be rescheduled appropriately.

Short-term Personnel Planning, Ordering Materials The short-term production schedule determines the appropriate personnel of the shop floor with respect to the knowledge and capability. Short-term personnel planning determines the detailed schedule of the staff with consideration of employment agreements and labour costs. As some amount of material might already have been committed by mid-term contracting, the short-term task of filling the commitments in a cost efficient manner still remains.

Coordination and Integration

As already mentioned the planning modules in an HPS need to be connected by information flows. Typical contents of these flows are discussed in the following.

Horizontal Information Flows The main horizontal flows go upstream, consisting of customer orders, sales forecasts, internal orders for warehouse replenishment and for production in the various departments, as well as of purchasing orders to the suppliers. This way, the whole supply chain is driven by the customers. However, the exchange of additional information in both directions and not only between neighbored modules, can improve the supply chain performance significantly (see bullwhip effect, Chap. 1). This concerns in particular actual stocks, available capacity lead-times, and point-of-sales data.

Vertical Information Flows Downwards flows coordinate subordinate plans by means of the results of a higher level plan. Typical informations are aggregate quantities, allocated to production sites, departments, or processes. The timing of quantities is better expressed in form of projected final stocks at the end of the lower level planning horizon because this includes the information about the longer planning horizon on the upper level and provides more flexibility on the lower level. Coordination is also achieved by allocation of capacities and by setting due dates.

Upwards flows provide the upper level with more detailed data on the performance of the supply chain, e. g. actual costs, production rates, utilization of the equipment, lead-times etc. This information can be used in the upper level planning for anticipating the consequences for the more detailed processes on the lower level.

4.3 Examples of Type-Specific Planning Tasks and Planning Concepts

Up to now quite general planning tasks – to some extent appearing in every supply chain – have been described. However, the importance of a specific

planning task may vary with respect to the type of supply chain considered. While some tasks, e.g. *lot-sizing* or *ordering materials*, may be extremely difficult (and thus relevant) in one type of SC, they may be quite simple (and therefore negligible in terms of planning) in another type of SC. In order to illustrate this, the two exemplary “SC-types” of the last chapter, *consumer goods manufacturing* and *computer assembly*, will be picked up, again.

Their most important planning tasks are derived from the characteristics of the respective SC-type. To admit a better differentiation, type-specific names will be introduced for some particularly characteristic tasks. Tables 4.1 (p. 94) and 4.2 (p. 100) try to emphasize the causal linkage between the typology of Sect. 3 (Tables 3.3 and 3.4) and the impact on planning that the respective attributes of an SC-type have. Additionally, hierarchical planning concepts – especially designed to link these respective tasks – will be shown as an example. For sake of brevity, we concentrate on mid- and short-term operational planning tasks, only.

4.3.1 Consumer Goods Industry

Master Production Scheduling, Capacity Planning and Mid-term Distribution Planning As consumer goods manufacturers often face seasonal or strongly fluctuating demand and because the supply chain is capacity-constrained, it is necessary to smooth those effects by pre-production in periods with less customer demand. Here, master production scheduling has to trade off the costs for seasonal stocks due to pre-production and the costs for capacity, especially the additional expenditure for working overtime in periods with peak demand. Up to now, most consumer goods manufacturers had a quite low working time flexibility and therefore changes in the working time pattern already had to be announced on the mid-term. Because of this and because of the scarce capacity, mid-term planning of working time is a crucial task in consumer goods industry. But in the meantime, more and more labour agreements are going to provide flexible working times. Thus, further sophisticated planning methods could lead to lower costs by effectively taking advantage of the additional freedom.

Furthermore, quite a lot of consumer goods companies use more than one site for producing the same product. Thus, the above planning task is getting more complex as capacity problems could be balanced by shifting production quantities from one site to another. Therefore, the costs for transports to the demand point are relevant and have to be considered, too, during the decision process. This extension of master production scheduling leads to a planning model (in general: *capacity-constrained master planning*) which includes both the tasks of mid-term production planning and mid-term distribution planning. If alternative sites producing the same products are sourcing their material from multiple suppliers with substantially different purchasing prices, the master planning model has to integrate the procurement side, too.

Table 4.1. Specific planning tasks of the SC-type “consumer goods industry”

Attributes & contents	Impact on planning
multiple sourcing of material	short- & mid-term supplier allocation
flow line organization	simultaneous ...
batch production	... lot-sizing and ...
sequence dependent changeovers	... scheduling necessary
known, stationary bottlenecks	focus on bottlenecks possible
low working time flexibility	mid-term planning of working time
3-stage distribution system	choice of distribution channels, allocation of safety stocks
seasonal demand	building up seasonal stock
long life cycle	forecasts based on historical data
hundreds of product types	aggregation ...
standard products	... of final items ...
divergent BOM	... necessary & possible
alternative sites	integrated mid-term production & distribution planning
deliver-to-order	forecasts & safety stocks of final items, deployment, shortage planning
capacity constrained	high utilization aspired, master planning w. r. t. capacity
intra-organizational	central coordination by means ...
coordination of mixture type	... of mid-term “master ...
unlimited information	... planning” possible
customer oriented	high service levels aspired

Usually, the main result of master planning in the consumer goods area is *not* the production quantity because the demand or forecast might change in the short run. Therefore, short-term scheduling needs to plan with updated demand data. So, the necessary capacity (especially working time, shift pattern, and overtime), the quantity which has to be pre-built (seasonal stock), and the transport capacity on each link are the decisions aided by master planning.

Mid-term and Short-term Sales Planning Since an deliver-to-order decoupling point is given, all production and most of the planning processes are driven by forecasts, more precisely, by forecasts for final items. Forecasting is often the crucial point in consumer goods industries because inventory of finished products is quite expensive and lost sales or backlogs reduce the customer’s trust in the company. These effects are sometimes amplified by depreciations which arise because of the low shelf-lives of the products. Therefore,

it is necessary to include the seasonal influences and the additional demand which is caused by promotions and marketing activities.

The high number of product types forbids the forecasting of individual final items for a mid-term planning horizon. However, since standard products are considered and since a divergent BOM is given, aggregation of final items to product groups quite often is straightforward. Thus, in mid-term forecasting usually aggregated product groups are considered and the time buckets comprise one week or more. As a general rule, the total planning horizon should at least include a complete seasonal cycle. Usually, the planning task consists of two steps. The first involves statistical forecasting under consideration of trends and seasonal effects. For that purpose, the time series of past demand are analyzed and extrapolated into the future. This can easily be done because the long life cycles of products give access to a long history of sales data. In a second step, the additional demand which is caused by planned marketing activities is added to the base forecast.

The short-term forecasting procedure then considers all products and a more detailed time grid (usually daily buckets). As the sales personnel has exact information on promotions for each time bucket (day), the short-term forecast figures should be composed from the statistical base forecast, supplementary demand resulting from promotions, and the change in demand caused by seasonal fluctuations. The information on seasonal effects (calculated in mid-term sales planning) has to be considered as add-on to the base forecast because the short horizon comprises not a complete cycle which is necessary for a seasonal planning model.

Lot-sizing and Machine Scheduling Production planning in consumer goods industries seems simple as the production process only consists of one or two stages. But in practice one of the hardest planning problems occurs because of high sequence dependent setup costs and times. This dependence enforces the simultaneous determination of lot-sizes and sequences: changes in the sequence of lots cause alterations in setup costs and setup times (i. e. in the net capacity actually remaining for production) which influence the lot-sizing decision. But the sequencing decision in turn is based on known lot-sizes. This problem is the more crucial, the tighter capacities are. However, since often bottlenecks are stationary and known, it is possible to concentrate on a single bottleneck stage comprising several parallel flow lines.

Transport Planning, Warehouse Replenishment A further crucial task in consumer goods industries is to balance the inventories in the multi-stage distribution network. Two major types of stocks are affected on the short-term, namely the lot-size and the safety stock.

In a deliver-to-order (= make-to-stock) environment final items have to be produced on forecast, i. e. without knowing customer orders. These production quantities, the so-called *lot-size stock*, have to be distributed among the

various stocking points of the 3-stage distribution system at which customer orders arrive. The task of *deployment* is to plan the short-term transportation activities such that customer orders can best possibly be fulfilled.

The deliver-to-order decoupling point also enforces *safety stocks* of final items to be placed at the most downstream stage (i. e. before customer delivery) in order to avoid stock-outs. In a 3-stage distribution system it seems – for risk pooling purposes – often reasonable to hold a part of the safety stocks at upstream warehouses (e. g. central warehouses etc.). Thus, not only the determination of the total amount of safety stock, but also the allocation of safety stocks within the distribution system are important planning tasks, seriously influencing customer service.

Because of the intense competition in consumer goods supply chains and because of the high power of customers (wholesalers, retailers) very high service levels are aspired. However, usually not all incoming customer orders can immediately be served from stock. The crucial task of selecting the minor important orders that can best be postponed (but nevertheless may get lost because customers become annoyed) is called “*shortage planning*”.

Coordination and Integration Since an intra-organizational supply chain is given, information could centrally be made available and central coordination should basically be possible. This coordination task should be settled on the mid-term master planning level because – as we have seen above – here an integration of procurement, production, and distribution is necessary, anyway.

After deriving these specific planning tasks of the consumer goods SC-type the question is how to link them together to get an integrated planning concept covering the whole (intra-company) supply chain best possibly. As we have seen in Sect. 4.1, *hierarchical planning* is a proper way to allow such a coordination. Of course, only a rough and very general draft of such a planning concept can be shown here. Details concerning aggregation of products or resources, time buckets of planning modules, and planning frequencies have to be skipped over. Thus, Fig. 4.4 only presents a “skeleton” of planning modules and the basic information flows between them. A planning concept for a real world supply chain has to be adjusted appropriately. A more complex consumer goods supply chain may comprise further planning tasks and require additional modules with the respective information flows in between. However, we hope to give some idea how the specific planning requirements of a consumer goods SC-type have to be reflected in a “fitting” planning concept.

Because of the higher degree of uncertainty only such decisions that cannot be postponed to later, shorter-term planning should be predetermined at the (capacity-constrained) *Master Planning* level. Just this information should be passed on the short-term level by means of instructions. As we have already seen, in consumer goods supply chains such decisions usually

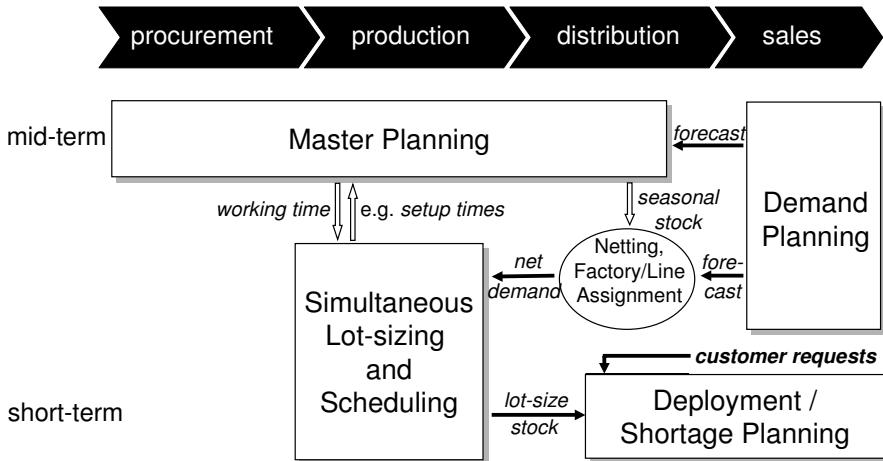


Fig. 4.4. Exemplary operational planning concept for the consumer goods manufacturing SC-type

comprise the determination of working time like shift patterns (because of its low flexibility) and the build up of seasonal stock (because of the long planning horizon being necessary). In order to take sound decisions, all influencing factors should be considered. For mid-term master planning in consumer goods supply chains this means that constraints like

- dynamic forecasts of customer demand (in order to reflect seasonality),
- limited capacity of resources and capabilities of extension,
- minimum stocking levels (safety stock and anticipated lot-size stock),

and further decisions like

- transport flows from factories to central warehouses (CWs) and customers (because stocks can be balanced between CWs) and
- production quantities of factories
(in order to evaluate the amount of overtime being necessary)

altogether have to be integrated in a single, holistic view of the supply chain.

This can (for reasons of complexity) and should (because of uncertainty) only be done in an aggregate manner, e.g. by means of product types, aggregate resources and monthly time buckets. Demand information has to be available at the same aggregation level. Such mid-term forecasts often are made in a further *Demand Planning* task by a central Sales department by consolidating the (more accurate) decentral forecasts of their regional dependencies and upgrading this aggregate forecast with additional, centrally available information like planned TV advertisements etc.

Because of seasonality the planning horizon usually should include at least one seasonal cycle – quite often a year. To make mid-term planning more

realistic, decisions of the short-term level, to be taken at later moments, have to be anticipated. In consumer goods supply chains average setup times (also reducing mid-term capacity, but not being considered in detail in mid-term planning) or the average level of lot-size stock are of relevance. These essentially are a result of the shorter-term lot-sizing and scheduling module.

Short-term planning has to respect the instructions of the mid-term planning level. However, short-term planning has a more detailed view of the supply chain. For example, since *Simultaneous Lot-sizing and Scheduling* (SLS) has to decide about changeovers, now “setup families” have to be considered which have the property that setup costs and setup times only occur for changeovers between items of different families (see Sect. 3.4, p. 74). Usually, a product type consists of several setup families. Thus, there is a higher level of detail than it was at mid-term master planning.

Also a shorter planning horizon suffices (e.g. two months) and capacities of production lines instead of aggregate resources are the limiting factor. Consequently, the aggregate instructions of the mid-term planning level have to be disaggregated into more detailed instructions for the short-term level. That means that working time commitments have to be refined at the decentral factories (maybe within an additional master production scheduling task) and that seasonal stocks of product types have to be assigned to setup families.

On the short-term, usually more accurate forecasts of customer demand are available. These short-term forecasts, the disaggregated seasonal stock, and the planned safety stocks are balanced with the initial stocks that are currently available at the central warehouses to compute the net demand that has to be satisfied by SLS. This net demand furthermore has to be *assigned* to (the production lines of) the factories. Note, if initial inventories have fallen below the safety stock levels, a part of the net demand is used to “refill” safety stocks. Also note that this *Netting* procedure has the character of a disaggregation step and that due to the better demand information the mid-term (virtual) transportation flows between factories, central warehouses, and customers normally have to be revised on the short-term.

At each factory, the decentral SLS is responsible for production line planning, i.e. determining the sizes and sequences of production lots of setup families. The lot-size stock of final items, resulting from a further disaggregation of setup families into final items (within SLS), has to be *deployed* to the CWs at which customer requests arrive. As the deliver-to-order decoupling point indicates, the final *Shortage Planning* at the CWs matches the incoming customer orders against the forecast-based stocks and determines whether and when a certain order will be delivered.

Finally note that – for sake of clarity – only two dimensions are printed in Fig. 4.4, but actually three dimensions would be necessary. This is due to the fact that there may be several factories and CWs where planning tasks like SLS or demand planning have to be tackled decentrally. Furthermore,

additional planning levels and modules may be required, e. g. in order to plan the movement of machines or tools between factories (see e. g. Sect. 20.1.2). This has to be done if total customer demand is stable but regional customer behaviour changes over time. Then, it may be advantageous to serve customer demand always from the nearest factory in order to save transportation costs of finished products, but this also depends on the costs for the movement of machines. Such a planning task would have a lower planning frequency than the ordinary master planning described above.

This example already shows that our typology is by far not (and cannot be) comprehensive. Even a small change in the assumptions being made may have significant impact on planning tasks and planning concepts. As a second example, in our consumer goods supply chain we (implicitly) restricted ourselves to products with a rather long shelf life. If this is not the case (e. g. for fresh food), holding stocks is only possible for a very short time. Then excess capacity instead of inventory has to balance seasonal demand and the lot-size stock has to be restricted, too. So the planning concept of Fig. 4.4 is not appropriate any more and has to be adjusted accordingly. However, we think that quite a lot of supply chains fit the consumer goods SC-type introduced above. Nevertheless, the fresh food example shows that it is very important to document *how* a planning concept has been derived from the specific characteristics of an SC-type. Because only then it is possible to check whether the own supply chain fits the type and where adjustments in the planning concept have to be made.

As a second and quite contrary example of type-specific planning tasks and corresponding planning concepts we now come back to the computer assembly type introduced in Sect. 3.5.

4.3.2 Computer Assembly

As pointed out below and summarized in Table 4.2, the specific characteristics of the computer assembly SC-type necessitate special emphasis on quite different planning tasks.

Master Production Scheduling, Capacity Planning and Mid-term Distribution Planning As opposite to the consumer goods type, less a capacity-constrained, but rather a material-constrained supply chain can be found. Because of the high working time flexibility, capacity of production is only a minor focus of mid-term planning. The limited availability of some important components, however, is a serious problem. If critical suppliers have a high power within the supply chain, mid- to long-term contracts (comprising both maximum supply and minimal purchasing quantities) ought to ensure the desired flow of components. These commitments limit the material supply (upper and lower bounds) that can be utilized. Due to their long lead-times quite a lot of components have to be ordered in good time on basis of demand forecasts.

Table 4.2. Specific planning tasks of the SC-type “computer assembly”

Attributes & contents	Impact on planning
large number of products procured long supplier lead-times unreliable supplier lead-times short materials’ life cycle	mid-term master plan coordinates purchasing & order promising safety stocks of components high risk of obsolescence, mark down, phase-in, phase-out
no bottlenecks in production	only rough capacity planning necessary
2-stage distribution system	merge-on-the-fly
forecasts & orders available short life cycles customized BOM convergent BOM	forecast netting no sales history available configuration check demand-supply matching, component substitution
assemble-to-order	forecasts & safety stocks of components, order promising, allocation planning
material-constrained	master planning synchronizes materials
supplier oriented customer oriented	long- & mid-term contracts short delivery times, high delivery reliability aspired

Both material constraints and long lead-times enforce a mid-term balancing of demand against possible component supply. In so doing backlogs may arise. As will be shown below, *order promising* needs to know component availability in order to set reliable delivery dates as soon as customer requests arrive. The information about availability (the so-called ATP quantities) is a result of this *material-constrained master planning*. Thus master planning has to synchronize the purchasing of a vast number of different components (planned component inflow) and to provide this information about planned component availability for *order promising* in form of ATP.

Mid-term distribution planning is only a relevant topic if an order can be satisfied from alternative sources such that one needs to choose between different distribution channels. Only in this (rather seldom) case, the distribution system has to be incorporated in master planning.

Mid-term Sales Planning In configure-to-order and assemble-to-order environments all assembly processes are kicked off by a specific customer order. Processes upstream from the decoupling point – and especially the *purchasing* – have to be based on forecasts, either directly on forecasts for components or indirectly on forecasts for final items.

In the first case, component demand could be estimated *directly* on basis of the sales histories and the assembly histories, respectively. In case of short life cycles, there is only a very poor history available. Sometimes, knowledge about life cycles of related components with similar functionality (e.g. of the discontinued predecessor) can be utilized as a surrogate. However, such a direct approach is mostly useful for C-components and -materials with minor value and rather long life cycles.

For high tech A-components with rather short life cycles the risk of obsolescence is very high and not only understocking, but also overstocking should be avoided. Then, one may try to *indirectly* derive a (hopefully) more accurate component demand from the production programme. Thus final item demand has to be estimated on basis of aggregate product types. Component demand (= planned component inflow) has to be derived from the planned production quantities in a sort of BOM explosion (as integral part of the master planning process). This task can quite easily be implemented in assemble-to-order environments where standard variants are predominant. In case of a configure-to-order decoupling point, however, also the structure of the BOM, i.e. the share of components within product types (e.g. the share of 120 GB and 160 GB hard disks within consumer PCs) has to be estimated which is an extremely difficult problem. Note that the component demand considered here corresponds to the planned component inflow stated above as a result of master planning. But the master planning process has to simultaneously respect supplier lead-times and material constraints. Thus master planning is more than a simple forecasting procedure.

Short-term Sales Planning On the short-term more accurate demand information is available, i.e. the already known customer orders' share of actual demand is higher. So one has to wonder how to integrate this information into the forecasting process and how to match "old" forecasts with incoming customer orders ("forecast netting"). The latter problem actually comprises the tasks of controlling forecast accuracy and reacting to forecast errors. Since forecast errors should be hedged against by safety stocks, here refilling of safety stocks (in case of too pessimistic forecasts) or reduction of the currently available stock (in case of too optimistic forecasts) are addressed. In consumer goods supply chains this netting procedure is still a relatively simple task because just stocks of final items have to be considered. In computer assembly supply chains, however, stocks of components have to be netted. This implies that forecast accuracy can also be measured on the component level.

Besides the danger of understocking, there is a high risk of overstocking of components because of their short life cycles. Thus, at the end of the life cycle one possibly has to take care about promotions or discounts in order to get rid of obsolete component stocks. In any case, older components have frequently to be replaced with their more modern successors (phase-in, phase-

out). Thus, quite often forecasts for both predecessor and successor have to be aligned (see Chap. 7).

An upstream decoupling point entails rather long order lead-times. Thus – as compared to consumer goods manufacturing – there is a noticeable time span between a customer request and the delivery of the complete order to the customer in computer assembly supply chains. If a customer has to wait anyway, he at least wants to get a reliable promise at which point in time his order will be delivered (a so-called “due date” or “promised date”). So the *order promising* and all subsequent further *demand fulfilment* processes are very important tasks within such a type of supply chain. Whereas short delivery times and early due dates are aspired by order promising, the compliance with that due date has highest priority throughout the demand fulfilment afterwards.

Quite often *order promising* is an online task. A customer wants his due date to be assigned very soon after his request (e.g. within a few minutes). Then order promising has to be executed on a first-come-first-served basis. Thus, there is a high chance that a less lucrative order books components that later on could be assigned to a more lucrative order. In order to realize higher profits, it may be useful to allocate quota of components to specific customer classes (as it is well known from yield management and flight ticketing). Such a “refinement” of ATP is sometimes called *allocation planning*. Note that allocation planning is only required in shortage situations.

Lot-sizing and Machine Scheduling As we have seen, in computer assembly supply chains setup costs and times are negligible. There are no serious bottlenecks in production and working time is quite flexible, even on the short-term. Thus lot-sizing is irrelevant and scheduling the *released* customer orders (“*production orders*”, “*jobs*”) with the objective of meeting the promised due dates also isn’t a very critical task.

However, in order to select the orders to be released next, the currently available, anonymously purchased stocks of components (“*supply*”) have to be assigned to the already promised customer orders (“*demand*”). This *demand-supply matching* is only important in shortage situations. If supply of components is not sufficient to satisfy all customer orders in time, i.e. with respect to the promised due dates, one has to decide which demand should be backlogged and which supply should be accelerated. In the first case, the Order Management department has to contact some carefully selected customers and to inform them about delaying their orders. Of course, simultaneously new second or even third promised dates have to be set (“*repromising*”). In the second case, the Procurement department has to negotiate with some critical suppliers in order to (hopefully) speed up the delivery of their components. Since hundreds of components and thousands of customer orders might be concerned and thus should be considered, this obviously is a very difficult task. Note that there can be further degrees of freedom, e.g. due

to component substitution, because customers might be satisfied by similar components of alternative suppliers not originally agreed on.

Transport Planning, Warehouse Replenishment Like it was the case for mid-term distribution planning, shorter-term transport planning is not a critical task. Sometimes, there may be a choice between alternative transportation modes, e. g. between “normal” delivery by a carrier and “express” delivery by a parcel service.

It is interesting to note that – because of the convergent BOM – an assignment of currently available stock to customer orders, similarly to the demand-supply matching, may be required at several stages downstream from the decoupling point. The latest possible stage in a 2-stage distribution system are the distribution centres where different order lines (e. g. monitors and computers) have to be “matched” to a complete order. Such matching tasks are necessary whenever a customer order initiates the release of material (or the execution of some processes), but the material released (or the output of the process) will not durably remain assigned to this specific order. For example, customer order 1 may initiate the assembly of a system unit, but order 2, having a higher priority, will finally catch this unit. Such a procedure increases flexibility, yet also decreases the stability of a system. The earliest possible “marriage” between an order and its components – as the other extremal – would be the *durable* assignment of ATP on hand at the order promising stage. Then, very reliable due dates can be promised (because the necessary components are already on stock and cannot be caught by other orders) and a complete tracking and tracing of this order is possible. Obviously, such a procedure necessitates a high stock level due to high WIP.

However, the major focus of short-term planning is on the supply side. As introduced above, safety stocks have to be held on component level. This is the more important, the longer and the less reliable supplier lead-times are. As compared to the consumer goods supply chain, determination of correct safety stock levels is more complicated since service levels are usually defined and measured for finished products, whereas safety stocks have to be set for components. Because of the short material life cycles, there is a high risk of obsolescence, too. So at the end of the life cycles, short-term safety stock planning has the character of a newsboy problem (see Nahmias (2001, Chap. 5.3)).

Coordination and Integration Due to the high power of some suppliers and customers, intensive collaboration should be established, e. g. in order to exchange capacity (material availability) or demand information. For the intra-company part of planning, also central coordination by means of a (material-constrained) master plan is useful which synchronizes the activities of the Sales, Production, Procurement, and Order Management departments. The outcome of master planning should be the planned inflow of components.

As can be seen in Fig. 4.5, this information is used to synchronize the purchasing (by means of the aggregate inflow) and order promising (by means of ATP). The input of master planning may be mid-term forecasts for final item demand (aggregated to product types) and attach rates, i.e. forecasts for the share of components within these product types. Both are results of a *Demand Planning* task which usually is in the responsibility of the Sales department. As for consumer goods supply chains, also decentral forecasts of several sales regions have to be consolidated and upgraded to an aggregate forecast for the company.

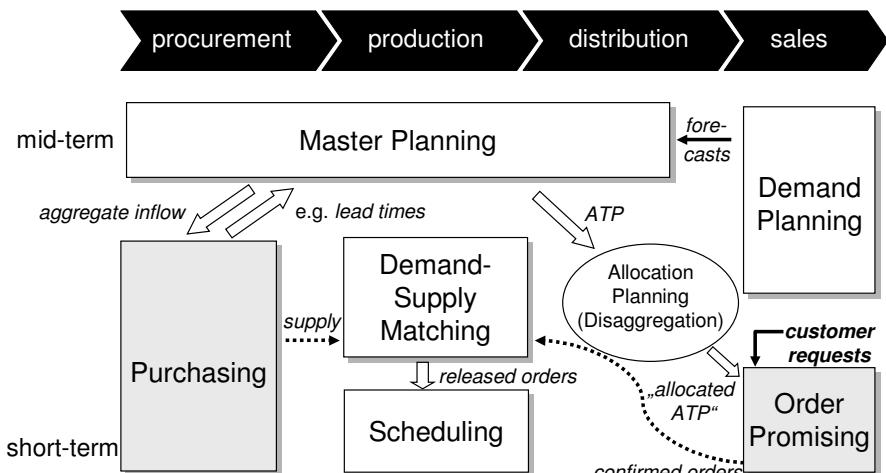


Fig. 4.5. Exemplary operational planning concept for the computer assembly SC-type

Thus, the task of *Master Planning* is to link the planned component inflow with final item demand. This task would be straightforward if there weren't any constraints. While production capacity is a rather loose limitation, the problem is to respect upper and lower bounds for the procurement of some critical components and to respect the varying, partly long lead-times. The objective should be to balance inventory holding costs for components against profit that might be obtained by different product types in several regional markets. Note, however, that purchasing and order promising not necessarily have to be synchronized by taking monetary objectives into account because just a *unique* master plan – no matter whether cheap or expensive – is required.

Purchasing needs to know about the aggregate component inflow master planning calculates with, e.g. about the weekly or monthly inflow of hard disks of a specific size or class of sizes. Concrete purchasing orders to each supplier (which entail a higher level of detail) have to meet this aggregate com-

ponent inflow best possibly. Thereby, multiple sourcing, supplier contracts, economic lot-sizes, and safety stock targets (including forecast netting) have to be taken into consideration. The master plan can only take care of the most critical A-components. Thus, the remaining B- and C-components have to be forecasted and ordered, directly. The result of purchasing is the component inflow (component supply) that arrives at the inbound warehouses and becomes available for assembly. In order to feed master planning with up-to-date data, purchasing has to provide realistic information about lead-times and minimum or maximum purchasing quantities of critical components.

On the other hand, order promising requires information about ATP quantities, i.e. the part of the component stock on hand and the expected component inflow (already in transit or planned by master planning) that has not yet been allocated to specific orders and thus can be promised to customers in the future.

Since final item demand has driven the master plan, there already has been some rough assignment of component stock – and thus ATP – to different markets. However, if detailed quotas for smaller sales regions are required to permit an online order promising, the output of the master plan has to be refined into “allocated ATP” in a further *Allocation Planning* step. Similar to the netting procedure in consumer goods supply chains, this task primarily is a disaggregation step because the major (material-constrained) decisions about assignment of component stock to markets have to be taken on the master planning level. *Order Promising* then suggests a due date for an incoming customer order by searching within allocated ATP for all requested components of the order. In case of customer compliance with the date, the confirmed order finally books the corresponding components within allocated ATP (but usually not within physical stock) so that they cannot be promised a second time.

The coupling to short-term production planning is rather loose. *Demand-Supply Matching* has to balance the available stock of components – which is the actual supply resulting from short-term purchasing activities – with the confirmed orders. Note that actual and planned supply may deviate considerably because of unreliable lead-times. But this discrepancy should be buffered by safety stock (within master planning and purchasing as well). Besides supply acceleration activities and repromising of orders, the confirmed orders, to be released to the shop floor next, are the results of Demand-Supply Matching. These assembly jobs afterwards have to be *scheduled* on the shop floor. As mentioned above, if there is only a temporary assignment of components to customer orders, planning tasks similar to this demand-supply matching may also occur at further downstream stages, the last of them being settled at a distribution centre.

Of course, there may exist other useful ways to hierarchically link the planning tasks and planning modules of a computer assembly supply chain.

However, a planning concept for computer assembly has to take into account the specific requirements of such a type of supply chain.

References

- Anthony, R. N. (1965) *Planning and control systems: A framework for analysis*, Cambridge/Mass.
- Domschke, W.; Scholl, A. (2003) *Grundlagen der Betriebswirtschaftslehre*, 2nd ed., Berlin et al.
- Fleischmann, B. (1998) *Design of freight traffic networks* in: Fleischmann, B.; Nunen, J. A. E. E. van; Speranza, M. G.; Stähly, P. (Eds.) *Advances in Distribution Logistics*, Lecture Notes in Economics and Mathematical Systems, Vol. 460, Berlin et al., 55–81
- Fleischmann, B.; Meyr, H. (2003) *Planning Hierarchy, Modeling and Advanced Planning Systems*, in: de Kok, A.G.; Graves, S.C. (Eds.) *Supply Chain Management: Design, Coordination, Operation*, Handbooks in Operations Research and Management Science, Vol. 11, Amsterdam et al., 457–523
- Hax, A. C.; Meal, H. C. (1975) *Hierarchical integration of production planning and scheduling*, in: Geisler, M. A. (Ed.) *Logistics: TIMS Studies in Management Sciences* Vol. 1, Amsterdam, 53–69
- Hoekstra, S. (1999); Romme, J. (Eds., 1991) *Integral logistic structures*, New York
- Miller, T.C. (2001): *Hierarchical Operations and Supply Chain Planning*, Berlin et al.
- Nahmias, S. (2001) *Production and operations analysis*, 4th ed., Boston et al.
- Orlicky, J. A. (1975) *Material requirements planning*, New York et al.
- Rohde, J.; Meyr, H.; Wagner, M. (2000) *Die Supply Chain Planning Matrix*, in: PPS-Management, Vol. 5, No. 1, 10–15
- Schneeweiss, C. (2003) *Distributed Decision Making*, 2nd ed., Berlin et al.
- Silver, E. A.; Pyke, D. F.; Peterson, R. (1998) *Inventory management and production planning and scheduling*, 3rd ed., New York et al.
- Stadtler, H. (1986) *Hierarchical production planning: Tuning aggregate planning with sequencing and scheduling*, in: Axsäter, S.; Schneeweiss, C.; Silver, E. (Eds.) *Multistage production planning and inventory control*, Lecture Notes in Economics and Mathematical Systems, Vol. 266, Berlin et al., 197–226

Part II

Concepts of Advanced Planning Systems

5 Structure of Advanced Planning Systems

Herbert Meyr¹, Michael Wagner² and Jens Rohde³

¹ Vienna University of Economics and Business Administration, Institute of Transport Economics and Logistics, Nordbergstraße 15, A-1090 Vienna, Austria

² Paul Hartmann AG, Corporate Materials Management, P.O.Box 1420, 89504 Heidenheim, Germany

³ Darmstadt University of Technology, Department of Production & Supply Chain Management, Hochschulstraße 1, 64289 Darmstadt, Germany

APS have been launched independently by different software companies at different points in time. Nevertheless, a common structure underlying most of the APS can be identified. APS typically consist of several *software modules* (eventually again comprising several software components), each of them covering a certain range of planning tasks (see Rohde et al. (2000)).

In Sect. 4.2 the most important tasks of supply chain planning have been introduced and classified in the two dimensions *planning horizon* and *supply chain process* by use of the SCP-Matrix (Fig. 4.3). As Fig. 5.1 shows, certain planning sections of the SCP-Matrix, e. g. mid-term procurement, production and distribution, are typically covered by a respective software module. The names of the modules vary from APS provider to APS provider, but the planning tasks that are supported are basically the same. In Fig. 5.1 supplier-independent names have been chosen that try to characterize the underlying planning tasks of the respective software modules.

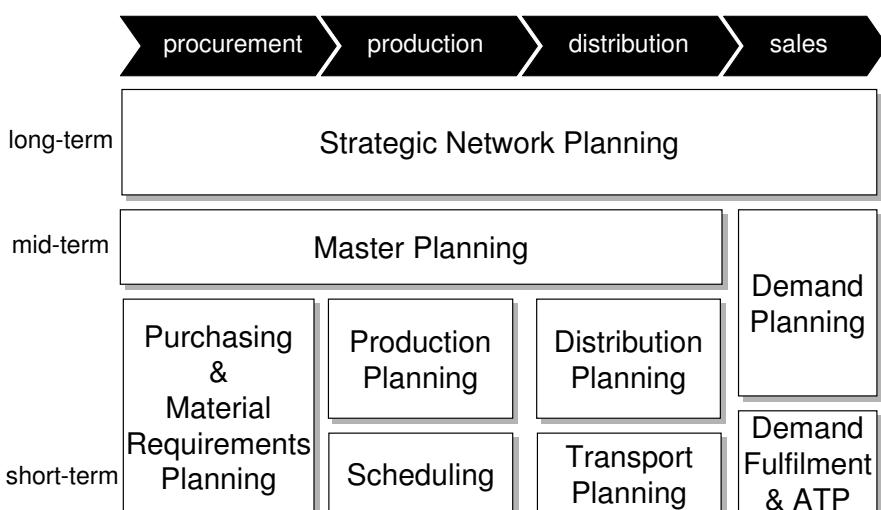


Fig. 5.1. Software modules covering the SCP-Matrix

APS typically do not support all of the planning tasks that have been identified in Sect. 4.2. In the remainder of the book it will be shown which tasks are actually considered (Part II), how to select and implement APS (Part III), how to build models using software modules (Part IV) and which solution methods are commonly used (Part VI). In the meantime, the following provides an overview of the structure of the software modules and the planning tasks concerned:

Strategic Network Planning covers all four long-term planning sections, especially the tasks *plant location* and the design of the *physical distribution structure*. Some questions that arise in *strategic sales planning* (e.g. which products to place in certain markets) can be considered, too. Basically, the design of the supply chain and the elementary material flows between suppliers and customers are determined.

Demand Planning. Further tasks of *strategic sales planning* (e.g. long-term demand estimates) and the *mid-term sales planning* are usually supported by a module for Demand Planning.

Demand Fulfilment & ATP. Most APS providers offer Demand Fulfilment & ATP components that comprise the *short-term sales planning*.

Master Planning coordinates procurement, production, and distribution on the mid-term planning level. The tasks *distribution*, *capacity* and *mid-term personnel planning* are often considered simultaneously. Furthermore, *master production scheduling* is supported.

Production Planning and Scheduling. If there are two separate software modules for Production Planning and Scheduling, the first one is responsible for *lot-sizing* whereas the second one is used for *machine scheduling* and *shop floor control*. Quite often, however, a single software module ought to support all three tasks.

Planning on such a detailed, short-term planning level is particularly dependent on the organization of the production system. Therefore, all bottlenecks have to be considered explicitly. If multi-stage production processes and product structures exist, they have to be coordinated in an integrative manner. In order to meet the specific requirements of particular industries, some software vendors offer alternative Production Planning and Scheduling modules.

Transport Planning and Distribution Planning. The short-term *transport planning* is covered by a corresponding software module. Sometimes an additional Distribution Planning module deals with material flows in a more detailed manner than can usually be done by Master Planning.

Purchasing & Material Requirements Planning. The planning tasks *BOM explosion* and *ordering of materials* are often left to the ERP system(s), which traditionally intend to supply these functionalities and are needed as transaction systems, anyway. As far as non-bottleneck materials are concerned, the BOM explosion indeed can be executed within an ERP system. However, an “advanced” purchasing planning for materials

and components, with respect to alternative suppliers, quantity discounts, and lower (mid-term supply contracts) or upper (material constraints) bounds on supply quantities, is not supported by ERP systems. Not all APS providers launch a special software module Purchasing & Material Requirements Planning that supports (mid- to) short-term procurement decisions directly. Sometimes, at least a further Collaboration module helps to speed up the traditional interactive (collaborative) procurement processes between a manufacturer and its suppliers.

The software modules of APS are dedicated to deterministic planning. However, there are uncertainties on both the inbound (unreliable suppliers, machine breakdowns) and the outbound (unknown customer demand) side. In order to hedge against uncertainty, buffers have to be installed – either in the form of safety stocks or safety times. Buffering against uncertainty is a task that covers all supply chain processes and actually cannot be assigned to a single software module because it depends on the particular industry and the locations of the decoupling points (see Tempelmeier (2001)). However, in accordance with some software providers, we describe the safety stock planning functionality of APS in Chap. 7, when discussing the details of the Demand Planning module.

The planning tasks may vary substantially dependent on the particular industries and supply chains, especially. This is especially true for the short-term planning tasks (see e.g. Drexel et al. (1994)). APS providers are increasingly becoming aware of this situation. Therefore, they offer several software components and/or software modules covering the same planning tasks, yet respect the peculiarities of the particular type of supply chain considered. So actually, a third dimension *supply chain type* should be added to Fig. 5.1. For the sake of clarity, however, the need for industry-specific solutions is visualized in a separate figure (Fig. 5.2).

Software modules can be seen as some sort of “planning kit.” The users buy, install and integrate only those modules that are essential for their business. In most cases, not all modules of an APS provider are installed. Sometimes, but not often, components of different APS providers are combined.

The converse is also possible. Some APS providers do not offer software modules for all planning tasks. However, APS suppliers seem to be highly interested in providing complete solutions. As a result, further modules for supplier and customer collaboration and supply chain execution (as we will see later on) have been launched. Quite often, APS vendors bundle APS modules together with modules for ERP and CRM in order to provide a comprehensive supply chain suite. Thus, sometimes it may be hard to identify the *planning* modules of the suite (especially their functionality) and to verify the APS-structure described above when visiting the web pages of the respective software companies.

Software modules are not always implemented for the planning tasks they originally had been designed for. For example, a Master Planning module can

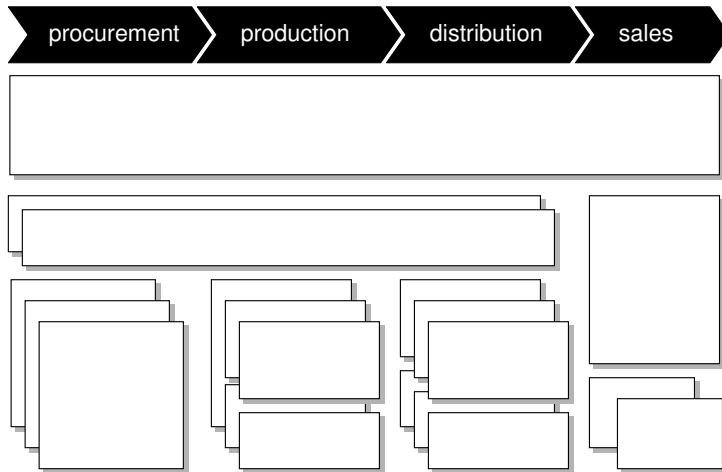


Fig. 5.2. Modules of APS for particular industries

be used for Distribution Planning. This happens if modelling features of the modules are quite similar and the same solution method can be applied to different types of problems.

Besides the already proposed software modules, additional software components are frequently supplied, which support the coordination of different software modules as well as the integration with other software systems, e.g. ERP systems or Data Warehouses (see Chap. 13).

However, preparing the technological capability to establish information links between different software modules is only the first step. The crucial question is what information should flow at which point in time. So the problem is to design and implement planning concepts that coordinate these software modules with respect to the objectives of the enterprise and supply chain as a whole, respectively, in the most effective manner. In Chap. 4 such planning concepts have been presented and it has been shown that they have to fit the particular planning requirements of different types of supply chains. Quite often, APS vendors provide solutions for particular industries, i.e. they arrange a set of software modules that are intended to serve a certain industry well. So far, however, “workflows” for particular industries are only seldom provided. Such workflows give some advice on how to establish the information flows between these modules so that they are well-integrated with respect to the peculiarities of the respective industry. This is achieved by rather general templates.

Also frequently offered are the tools for the integration (mostly using Internet technology) of different supply chain partners operating in different locations. These software components provide the necessary data for a supply chain-wide, long- and mid-term planning, and communicate the outcome of a central planning process to the respective de-central units. In most

cases, an alert system supports the interaction between central and de-central planning (see Sect. 4.1). Since Internet technology can be applied for various purposes, APS suppliers increasingly offer additional e-business tools, e. g. for the opening of virtual markets in order to purchase raw materials.

This book, however, concentrates on collaboration, not on market-based coordination. Market-based processes focus on pricing mechanisms to achieve coordination between two or more parties. Thus, they are of competitive nature. Collaboration or *Collaborative Planning*, however, places the emphasis on processes of cooperative nature as pursued in SCM.

Figure 5.3 shows the collaboration interfaces of an APS. Collaboration appears in two directions: collaboration with customers and collaboration with suppliers. From the view of a single member of the supply chain, collaboration is important on both ends of its SCP-matrix, the sales and the procurement side. The difference between the two types of collaboration is the divergent structure in the case of customer collaboration and the convergent structure in the case of supplier collaboration.

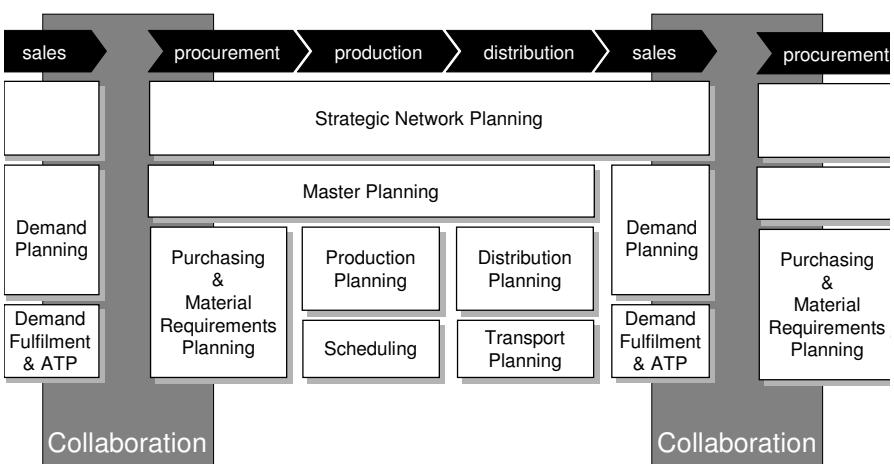


Fig. 5.3. Collaboration between APS

- One of the main applications of *Sales Collaboration* is the mid-term collaborative demand planning. In an iterative manner, forecasts are jointly generated. During this task, forecasts have to be coordinated and adjusted, e. g. by means of judgmental forecasting processes, as opposed to only aggregated. In shortage situations in particular, short-term collaboration may support ordinary ATP processes by providing additional information on alternative product configurations, delivery dates and prices.
- The task of mid-term *Procurement Collaboration* is to come to an agreement on procurement plans derived from master plans. Aggregated prod-

uct quantities have to be disaggregated and allocated to possible suppliers with respect to their capabilities. These capabilities can be evaluated and utilized efficiently in an iterative collaboration process. Thus, it is possible to generate procurement plans and delivery schedules that avoid material shortages.

As already shown in Sect. 4.1, Supply Chain Execution Systems (SCES) bridge the gap between preparing decisions in an APS and the final implementation of these decisions in practice (“execution”). Figure 5.4 (see e.g. Kahl (1999)) shows that software modules for supply chain execution also cover the supply chain processes “procurement,” “production,” “distribution” and “sales.” However, the planning tasks tackled there concern the execution, and thus comprise an even shorter-term planning horizon. For example, SCES deal with material handling, order transmission to suppliers, shop floor control, transportation execution (including tracking and tracing) and online response to customer requests. If necessary, they enrich the planning instructions of APS with further details (e.g. by human support), but mainly they monitor and control the execution of the decisions prepared by the APS. An *online* monitoring of the execution processes allows *real-time* reaction to unforeseen events.

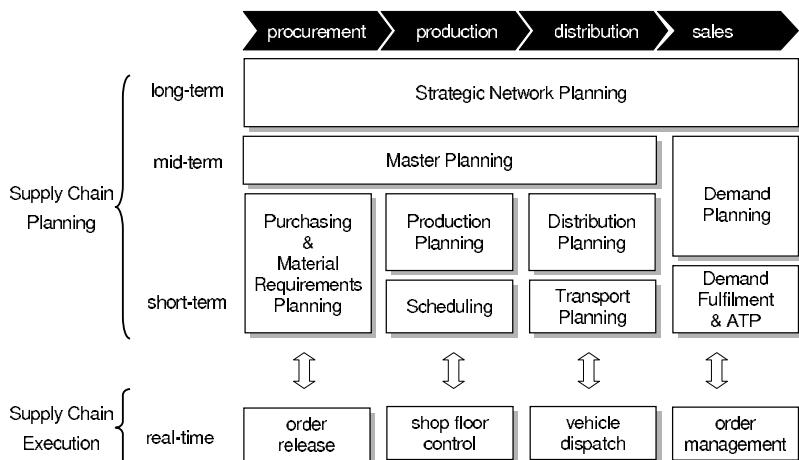


Fig. 5.4. Relation between APS and Supply Chain Execution Systems

SCES are closely coupled to APS by means of alert management systems, so-called Supply Chain Event Management (SCEM) systems. Thus, they are able to overcome the static planning intervals of traditional rolling horizon planning and allow for a reactive, event-driven planning. The borders between APS’ and SCES’ functionality cannot be clearly defined. For example, the order promising function may be part of both APS and SCES. Usually ATP

quantities are allocated to customer groups within an APS (see Chap. 9), whereas the online search for free ATP and real-time responses to customers are executed by an SCES. The search rules for ATP consumption may be defined in the APS (and sent to the SCES as directives) or may be customized directly within the SCES.

References

- Drexel, A.; Fleischmann, B.; Günther, H.-O.; Stadtler, H.; Tempelmeier, H. (1994) *Konzeptionelle Grundlagen kapazitätsorientierter PPS-Systeme*, in: Zeitschrift für betriebswirtschaftliche Forschung, Vol. 46, No. 12, 1022–1045
- Kahl, S.J. (1999) *What's the “value” of supply chain software?* in: Supply Chain Management Review, Vol. Winter 1999, 59–67
- Rohde, J.; Meyr, H.; Wagner, M. (2000) *Die Supply Chain Planning Matrix*, in: PPS-Management, Vol. 5, No. 1, 10–15
- Tempelmeier, H. (2001) *Supply Chain Planning with Advanced Planning Systems*, Proceedings of the 3rd Aegean International Conference on Design and Analysis of Manufacturing Systems, May 2001, Tinos Island, Greece

6 Strategic Network Planning

Mark Goetschalckx¹ and Bernhard Fleischmann²

¹ School of Industrial and Systems Engineering, Georgia Institute of Technology,
765 Ferst Drive, Atlanta, Georgia, USA

² University of Augsburg, Department of Production and Logistics,
Universitätsstraße 16, 86135 Augsburg, Germany

In this chapter we will focus on the long-term, strategic planning and design of the supply chain. During the strategic planning process an organization attempts to design a supply chain, which will enable this organization to maximize its economic performance over an extended period of time. Together with product research and development and marketing, the supply chain is one of the essential tools for a company to achieve their strategic business goals and practices. During the strategic planning process, companies identify their key products, customer markets for these products, core manufacturing processes, and suppliers of raw and intermediate materials. Virtually all organizations must redesign their supply chain from time to time to respond to changing market conditions, but the recent wave of mergers and acquisitions and the globalization of the economy have made this process even more frequent and important. For example, a company may wish to expand into a new geographical area where no infrastructure is currently in place, such as the expansion by electronics manufacturing companies into Eastern Europe after those countries adopted a market economy. Another company may wish to consolidate the duplicate distribution systems created by a merger or acquisition. Finally, strategic planning is not only used for expansion but also for retraction, such as when the United States Armed Forces developed a strategic plan for the base closings associated with the withdrawal from Western Europe.

Typically, the planning horizon for strategic planning ranges from three to ten years and the decisions involve the definition of customer and product zones, the definition of the stages in the manufacturing process, the establishment or closure of manufacturing and distribution facilities, and the installation of major manufacturing lines. The objective are most often financial objectives such as the maximization of the net present value (NPV) of profit or minimization of the NPV of costs, subject to customer service and budget constraints. The consequences of these decisions are manufacturing and distribution capacity and allocation of these capacities to products and customer zones. These capacities and allocations then become constraints in the master planning process. The master planning in turn determines the more detailed material flows and material storage for a number of smaller periods within a seasonal cycle.

Clearly, the decisions made during the strategic network planning have a major impact on the long-term profitability and competitive position of a corporation. But such far-reaching decisions typically have to be made based on data generated by very aggregate forecasts and economic trends. Demand for consumer goods in the developing nations of South America depends on the population data for that region, the global and local economic condition, and the profitability of serving that demand depends on the exchange rates during the planning period. As a consequence, corporations have become very much interested not only in the economic efficiency of their supply chain for the projected conditions, but also in the robustness and flexibility of their supply chain to adapt to changing and unanticipated conditions. Solution algorithms for designing systems when the data is not known with certainty belong to the class of stochastic optimization algorithms and simulation.

Planning models and decisions of the strategic network design have both interrelated spatial and temporal characteristics. For example, during an expansion into a new geographical area a company may decide to manufacture its products during the first two years in existing manufacturing facilities and to transport them to the new customer area. But starting in year three, when demand has grown sufficiently, the most economic production-distribution strategy may be to manufacture the products locally. This implies that the construction of the new manufacturing plant has to be started immediately. Many of the decisions made during strategic planning are of the go/no-go type. For example, the decision may be either to build a manufacturing plant in year three or not, but it is not possible to build half a plant. This type of decision is modeled with binary (integer) variables. The optimal solutions to these integer models are notoriously and provably hard to find. The solution algorithms and techniques to solve the models either to optimality or within a prescribed gap from optimal belong to the class of mixed integer programming (MIP) algorithms.

Finally, the strategic planning process is complicated by the fact that organizations execute strategic planning infrequently. A typical frequency may be during the creation of the next five-year corporate strategic plan. As a consequence, the people that performed the previous strategic planning have been promoted or left for other organizations. This implies that the new design team may have very little experience in model building and model solving.

To provide proper decision support for the strategic design of world-class supply chains, one has to recognize that supply chains have the following fundamental characteristics: they are holistic, global, and stochastic.

A holistic view of a supply chain does not focus exclusively on a single aspect of the supply chain performance such as inventory or direct labor cost, but takes an integrated and comprehensive view of the whole supply chain from the raw material suppliers, through the various transformation facilities and transportation channels, to the final customers. In addition, the evolution

of the supply chain over multiple years is considered. The "cradle to grave" approach is thus not only geographical but also temporal. Furthermore, it integrates the strategic capital budgeting decisions with the master planning. It considers purchasing, manufacturing, distribution, and transportation cost and capacities as well as the customer demand planning, which are geographically or organizationally dispersed components of a supply chain at a particular point in time. It also considers the full life cycle of different products, product portfolios, supply chain resources, and mergers and acquisitions that are temporally or organizationally dispersed components. The holistic view of the supply chain requires resolving the tradeoffs between the objectives and performance measures of the various agents and organizations in the supply chain.

Supply chains in virtually every industry are becoming increasingly global, which implies that materials are purchased, manufactured, and transported to the customers without consideration of national boundaries. While at the current time duties and tariffs still play an important role, the overall trend is towards elimination of these tariff and non-tariff trade barriers. Proper supply chain design must incorporate the conditions and aspects of international trade such as the various Incoterms (*international commercial terms*), duties and tariffs, not tariff-based trade barriers, and exchange rates. For highly specialized products the determination of the most advantageous transfer prices can also significantly impact the after-tax profits of the corporation, and must be addressed as part of the overall strategic supply chain design process (if allowed by the local taxation authorities).

The establishment of a facility as part of a global supply chain is a very important strategic decision. The economic life of a manufacturing facility or even a distribution facility can span several decades. The economic and market data over that time horizon are inherently not known with certainty at the time the decision is made to establish the facility. The supply chain configuration not only has to be efficient with respect to the expected conditions but also robust and flexible enough to adapt to the inevitable changes in these conditions. This implies that for every possible supply chain configuration there is not just one value of profit realization but rather a profit distribution that depends on the probabilities of the occurrence of various economic conditions. One way to quantify the robustness of the supply chain is to compute a measure of the variability of the financial performance, such as the standard deviation or absolute deviation. Robust supply chains will be characterized by relatively smaller variability measures. If we can assume that we know the probability distributions for the various economic conditions and parameters, the profit is said to be stochastic. A single combination of possible values of the economic conditions and parameters is called a scenario. The probability distributions of the conditions and parameters incorporate the chance of different underlying trends and the random variations around each trend. It is anticipated that a small number of underlying trends will be possible. A

corporation typically considers an optimistic, neutral or best-estimate, and pessimistic trend.

The problem is thus to design a supply chain configuration in a multi-period, stochastic with multiple scenarios, multi-country, multi-product, multi-echelon, multi-facility setting based on forecasted parameters and with bill of materials (BOM) flow conservation constraints. To a corporation, capital investments become more attractive if they have a higher expected return and/or if the variability of this return is smaller. The design objective is thus to maximize the difference of the expected value of the net present value (NPV) of the world-wide net cash flows (NCF) of the corporation over the planning horizon ($NPVNCF$) minus a compatible measure of the variability of the same $NPVNCF$ weighted by a risk-preference parameter α . This can be interpreted as α weighted combination of the multiple objectives of expected performance maximization and risk minimization. The risk-preference parameter α is non-negative number that may be different for each corporation. The higher the value of α the more risk averse the corporation is in its capital budgeting planning. At the current time, common practice is to design the supply chain to maximize the expected profit and then to evaluate the configuration with simulation or post-optimization to determine the variability as a measure for risk and robustness.

The complexity of this large-scale, holistic, global, and stochastic design problem far exceeds the capabilities and insight of even the most knowledgeable and experienced decision makers. To assist the decision makers in determining the most desirable supply chain configuration in an acceptable amount of time, the help of computers and software must be enlisted. A mathematical model of the supply chain has to be constructed combined with the development of efficient methods of determining highly desirable supply chain configurations.

Clearly, the proper execution of a strategic planning effort is a very challenging task. The decision support models must be comprehensive and cover both engineering and financial constraints, and often they are company or industry specific. The models require a large quantity and variety of data, which often must be forecasted with large degrees of uncertainty. The decisions are binary and thus even the deterministic MIP models would be very difficult to solve to optimality. But the data are typically not known with certainty, which indicates the use of stochastic optimization.

The remainder of this chapter will first present a semantic formulation of the strategic network planning problem. Next, the literature will be briefly reviewed. In Section 6.3, a general framework for the strategic network planning process will be presented and some advanced modelling issues will be discussed. Section 6.4 deals with properties of the corresponding modules in APS. Finally, conclusions and future directions will be presented.

6.1 Components of the Strategic Network Design Problem

The main objects in a strategic network design project are related to the different countries, planning periods, products, customers, vendors and suppliers, manufacturing and distribution facilities, and transportation assets. The two major types of design decisions are the 1) status of a particular facility or manufacturing line and relationships or allocations during a specific planning period and 2) the product flows and storage quantities (inventory) in the supply chain during a planning period. For example, if the binary variable y_{klt} equals one, this may indicate that a facility of type l is established at location or site k during time period t . Similarly, the binary variable w_{pksts} indicates if product p is assigned or allocated to a manufacturing plant at site k and of type l during period t in scenario s or not. The continuous variable x_{ijmpts} may indicate the product flow of product p from facility i to facility j using transportation channel m during time period t in scenario s .

The objective of the strategic supply chain design process is to maximize the long-term economic performance of the corporation. This objective has to be expressed in the financial performance measures familiar to corporate-level decision makers. For strategic supply chain configurations the primary performance measure is the NPV of the streams of the NCF. At the current time, a solution methodology does not exist if multiple objectives are combined, e.g. a linear combination of expected value and standard deviation (risk). So, supply chains are typically designed solely with an expected value objective and then evaluated with respect to more complicated objectives. Let cdf denote the capital discount factor and $E[\cdot]$ denote the expectation operator. Then the objective of the strategic supply chain design is

$$\max\{\mathbf{E}[NPVNCF]\} \quad (6.1)$$

$$NPVNCF = \sum_{t=1}^T NCF_t \cdot (1 + cdf)^{-t} = \sum_t \left(\sum_{c \in C} \frac{NCF_{ct}}{er_{ct}} \right) \cdot (1 + cdf)^{-t} \quad (6.2)$$

NCF_{ct} is the net cash flow for a country in the currency of the country during a particular time period (year) and it is equal to the sum of the net cash flows for all the facilities in operation or being established in that country during that time period. er_{ct} is the exchange rate for the currency of country c expressed in the currency of the home country. Also, time dependent discount factors cdf_t may be used resulting in a slightly modified objective function.

In the following formulas the subscripts t and s , indicating the planning time period and scenario, respectively, are omitted for notational simplicity. For each facility, there are four cost or revenue components: 1) revenue of finished products sold minus the cost of raw and component materials used, 2) fixed costs associated with the transformation process, and 3) variable costs associated with the transformation process, and 4) depreciation allowed

for the facility. The earnings before interest, taxes, and amortization are denoted by EBITA. The net revenue associated with acquiring or selling materials is based on the transfer price, the allocation of the transportation costs, and the allocation of the duties. See Choi (1997) for further details on the computation of the international accounting variables.

$$NCF_c = \sum_{k \in F_c} \left((1 - TaxRate_c) \cdot (EBITA_k - Interest_k) + Depreciation_k - Amortization_k \right) \quad (6.3)$$

$$EBITA_k = SalesRevenue_k - FixedCosts_k - VariableCosts_k - Depreciation_k \quad (6.4)$$

F_c is the set of facilities located in country c .

A particular strategic configuration of the supply chain will have a certain expected value and standard deviation of the $NPVNCF$. In case the probability distributions of the parameters are known, a classical risk analysis graph can be plotted where each candidate configuration is placed according to two dimensions: one axis representing the expected value and the other axis the variability or risk measure. Often the corporation does not know the value of the α parameter that corresponds to its risk preferences and is interested in identifying several alternative high-quality supply chain configurations for various values of α . The efficiency frontier is the collection of supply chain configurations that are not Pareto-dominated by any other configuration, i.e. for any efficient or non Pareto-dominated configuration, no configuration exists that has simultaneously a larger expected value and a smaller variability operator. For a given set or sample of supply chain configurations that are located in the risk analysis graph, the sample efficiency envelope (SEE) of those configurations can be determined by connecting efficient configurations. This SEE is an approximation of the efficiency frontier. The risk analysis graph for an industrial case with the standard deviation chosen as risk measure and including the SEE is shown in Fig. 6.1. The risk analysis graph is a very powerful communications tool with corporate executives since it displays in a concise manner the expected yield and risk of several possible candidates. It is the function of the strategic tools in the APS systems to perform all the calculations, optimizations, and simulations that are then synthesized into this graph.

The mean value problem (MVP) denotes a single scenario in which the best-guess, i.e. the expected value, of all the economic conditions and parameters is used. The location of the optimal configuration for the MVP is shown in the risk analysis graph. As is often the case in practice and also for the particular design project shown in Fig. 6.1 the above figure, the MVP configuration is not on the SEE and is dominated by other configurations and should

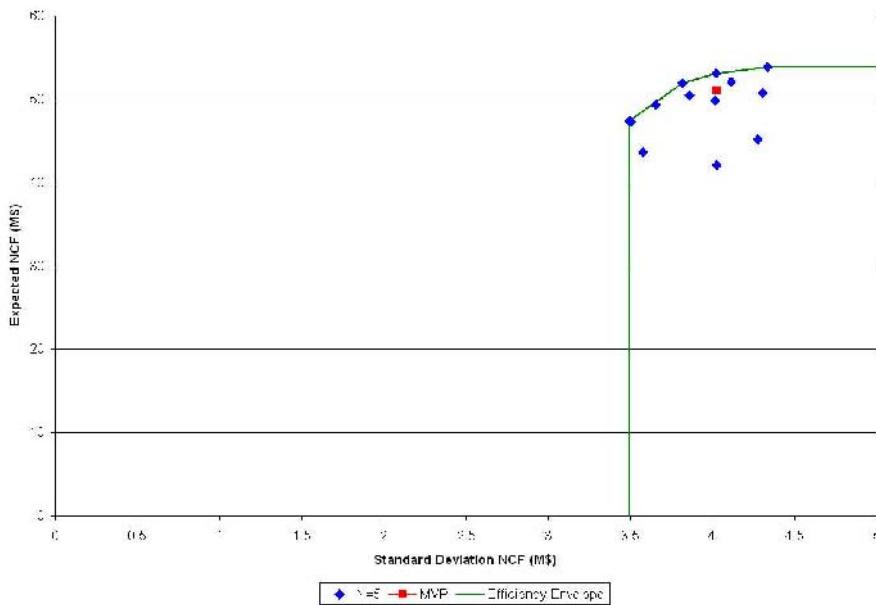


Fig. 6.1. Example of a Risk Analysis Graph for Strategic Supply Chain Configurations

not be selected. To avoid selecting such a non-efficient and non-robust supply chain configuration, one must perform extensive and systematic sensitivity analysis and stochastic evaluation.

The relationships between the different variables are usually fairly simple and typically linear. Four types of constraints are common in supply chain design models: conservation of flow, capacity, consistency or linkage constraints, and equality constraints used to compute intermediate quantities. Each type will next be discussed in further detail.

One type of conservation of flow constraints focuses on the material balance between different products, facilities, and transportation channels and are typically called bill of materials (BOM) constraints. They represent the fact that all material flow entering a facility or the total supply chain inevitably also must leave that facility or the supply chain, albeit in a different form. The general format of material balance constraints is

$$Inflow_t + Production_t = Outflow_t + Consumption_t \quad (6.5)$$

where $Inflow_t$ is the amount entering from all sources, $Production_t$ is the amount produced by transforming other products or supplied from external sources, $Outflow_t$ is the amount leaving to all destinations, and $Consumption_t$ is the amount removed by transformation into other products

or by customer demand. All quantities are always summed for a particular product and time period in a particular scenario.

A second type of constraints ensures that the model creates a feasible configuration by assigning capacities to different resources in the supply chain. For example, a supplier capacity relation will sum all the outgoing flows of a particular product to the different destination facilities and over all outgoing transportation channels to ensure it does not exceed the capacity of the supplier for that product during that period. These constraints also enforce consistency between the status of a facility or machine and the flow through that facility or machine. The general format of such constraints is given next for a single-product capacity constraint.

$$\sum_j Flow_{pkj} \leq Capacity_{pk} \cdot Status_k \quad \forall p \quad (6.6)$$

$Flow_{pkj}$ is the amount of material flow of product p from facility k to facility j , $Capacity_{pk}$ is the capacity of the facility k for the product p , and $Status_k$ is the status of the facility. All quantities are for a particular time period and in a particular scenario.

If capacities are limiting for a combination of products, then they are modeled using resources. Typical examples of resources are machine production hours and warehouse storage volume. The general format of such constraints is given next for a multiple-product capacity constraint.

$$\sum_{p \in P} ResourceRequirement_{rpk} \cdot Flow_{pk} \leq Capacity_{rk} \cdot Status_k \quad \forall r \quad (6.7)$$

$Capacity_{rk}$ is the capacity for a resource r in the facility k , the amount of resource required per unit of material flow is $ResourceRequirement_{rpk}$. All quantities are for a particular time period and in a particular scenario.

The third type of constraints ensures that flow in the supply chains is consistent with facility status and also specifies additional conditions on the configuration of the supply chain itself. An example of the first type is lower bound constraints on production of a product in a facility if any production of that product occurs. It should be noted that typically upper bounds are modeled using the capacity constraints described above. Examples of constraints on the configuration are limits on the number of facilities, either/or facility constraints which requires one out of a set of facilities to be established, or ordering of facilities which require one facility to be open before other facilities can be used. The general format of such constraints is as follows.

$$LowerBound_{pk} \cdot Assignment_{pk} \leq Flow_{pk} \quad (6.8)$$

$$Assignment_{pk} \leq Status_k \quad (6.9)$$

$$\sum_k Status_k \leq MaxCount \quad (6.10)$$

$Assignment_{pk}$ is a binary variable indicating if a product is assigned to be produced in a facility, $LowerBound_{pk}$ is the lower bound of production if this product is assigned to the facility, $MaxCount$ is the maximum number of facilities that can be established.

The fourth type of constraint is used to compute intermediate or derived variables and then force conditions on the value of those intermediate variables. One example is the EBITA for a country, which is computed as the sum of the EBITA of all the facilities in that country. Often there exists a lower bound on this country-wide EBITA. The computation of the derived variable is based on an equality constraint. The general format of such constraints is as follows.

$$NIBT_{ct} = \sum_{k \in F_c} (EBITA_{kt} - Interest_{kt}) \quad (6.11)$$

$$NIBT_{ct} \geq MinCountryIncome_{ct} \quad (6.12)$$

$NIBT_{ct}$ is the net income before taxes and $MinCountryIncome_{ct}$ is the minimal acceptable income in country c during period t , respectively.

As previously indicated, the parameters and economic conditions are not known with certainty at the time the strategic configuration decisions have to be made. If we assume that the probability distributions of all the parameters are known, then the design problem can be formulated as a two-stage stochastic programming problem. Because of the large number of random parameters each with their probability distribution, the total number of possible scenarios is infinite or extremely large. The scenarios can be used in two different ways. A deterministic equivalent problem (DEP) can be derived from the stochastic problem by explicitly including the objective function components and constraints for all or some of the scenarios multiplied by their respective probabilities. Solving the DEP yields a supply chain configuration that is optimal or near-optimal with respect to all the included scenarios and is called the robust configuration by some authors, see Kouvelis and Yu (1997). This is typically called scenario-based optimization. Scenarios can also be used in the evaluation of a given configuration. This is typically done with random sampling or simulation.

It should be noted that the optimal solution to the DEP has the best expected performance with respect to the scenarios included in the DEP. While including more scenarios is obviously one way to increase the accuracy of the expected performance of the configuration, this comes at significant increase in the computational burden. The selection of the scenarios to include is also a non-trivial problem. One would like to include all the possible scenarios that represent significant trends in the economic conditions. However, most

often there exist too many combinations of major trends, so that including all the corresponding scenarios would make the design problem computationally unsolvable. Rather than selecting specific scenarios, the sample average approximation (SAA) method creates scenarios through random sampling of the parameter probability distributions, Kleywegt et al. (1999).

While each decision variable and each constraint in itself is simple, the total number of variables and constraints creates very large problem instances. The creation and maintenance of the model formulation, data, and model solution requires significant information technology and computational resources. Typical comprehensive strategic supply chain design models may contain thousands of the binary variables and millions of the continuous variables in tens of thousands of constraints. Santoso et al. (2003) report solving a formulation with 1.25 million continuous variables for an industrial case. Papageorgiou et al. (2001) report that 3000 binary variables are present in a small illustrative example.

A schematic representation of didactic example of a single-period logistics system is given in Fig. 6.2. If there are multiple periods in the planning horizon, they would have a similar structure and typically would be displayed in different windows.

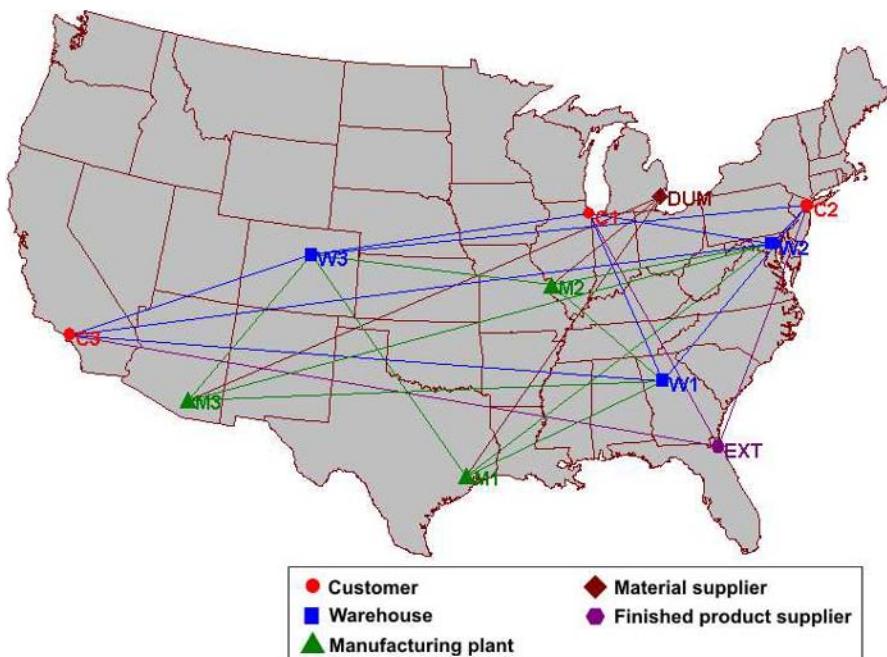


Fig. 6.2. Illustration of a Didactic Strategic Supply Chain Configuration

6.2 Review of Models in the Literature

A comprehensive review of all the models available or used in the design of supply chain systems is not possible within the confines of a single chapter. Because of its widespread application and significant financial impact, the strategic supply chain design problem has received a significant amount of attention in the research literature. Geoffrion and Powers (1995) provided a comprehensive review and evaluation of research. There exist several excellent reference books focusing on inventory models, Silver et al. (1998), on the inventory and transportation interactions, Tayur et al. (1999), and on the interactions between production planning and inventory, Graves et al. (1993). Fundamentally, international models have the same characteristics, variables, and constraints as single-country models but, in addition, they model exchange rates, tax rates, duties, tariffs, and local content laws. Vidal and Goetschalckx (1997) provide tables summarizing the features of strategic models for the design of domestic and global supply chain systems. A recent review of modeling and algorithms for the design of supply chain systems is given in Schmidt and Wilhelm (2000). Simchi-Levi et al. (2003) and Shapiro (2001) provide sections focusing on strategic design of supply chain systems.

Examples of recent detailed models for strategic supply chain design are given in Dogan and Goetschalckx (1999) for seasonal demands, Vidal and Goetschalckx (2001) for global supply chains with transfer prices, and Papa-georgiou (2001) for product portfolio selection.

Ballou and Masters (1993) surveyed developers and practitioners in the industry to determine the most important characteristics and the state-of-the-art in decision support systems for supply chain design. They found that model features and user friendliness were the most important features of the models and design packages. Ballou and Masters (1999) repeated the survey six years later and observed that advances in computer hardware and software had allowed real-world strategic supply chain systems design projects to be completed using mathematical models incorporated in commercial software packages. They reported that specialized and efficient algorithms have been developed to solve the spatial or geographical location aspect of supply chain systems, but that specialized or general-purpose simulation models are used for the temporal aspects such as inventory and production planning. Few models combine or integrate the spatial and temporal aspects of the supply chain. Based on a survey of active models and software packages, they found that the models are becoming more comprehensive and are beginning to include some tactical aspects. Global characteristics such as taxes, duties and tariffs, and exchange rates are included in only a few models. They report that linear programming (LP), MIP, and heuristics are the most commonly used techniques to find solutions. The practitioners responded with a large majority that modeling was used to configure their supply chain. In contrast with the 1993 result, in 1999 the practitioners ranked optimality of the solution as the most important characteristic of the software. The best features

of the models were their ability to represent the real-world system and to find an effective configuration. The worst features were difficulty in obtaining the necessary data, the complexity of using the model, and the poor treatment of inventory costs, especially in connection to customer service levels. The authors also observed that a consolidation trend is reducing the number of models and software packages available on the market.

6.3 Modelling Strategic Supply Chain Design

6.3.1 A Modelling Framework

As explained in Section 6.1, the strategic planning of a supply network must also consider the operations which are enabled by the strategic decisions on locations, capacities and products. Figure 6.3 shows the interdependence between the strategic and operational planning levels, where, for simplicity, no distinction is made between medium-term and short-term operational planning. The operational decisions concern all flows in the supply network, i.e. the quantities of various materials and products supplied, produced and distributed. The interdependence is

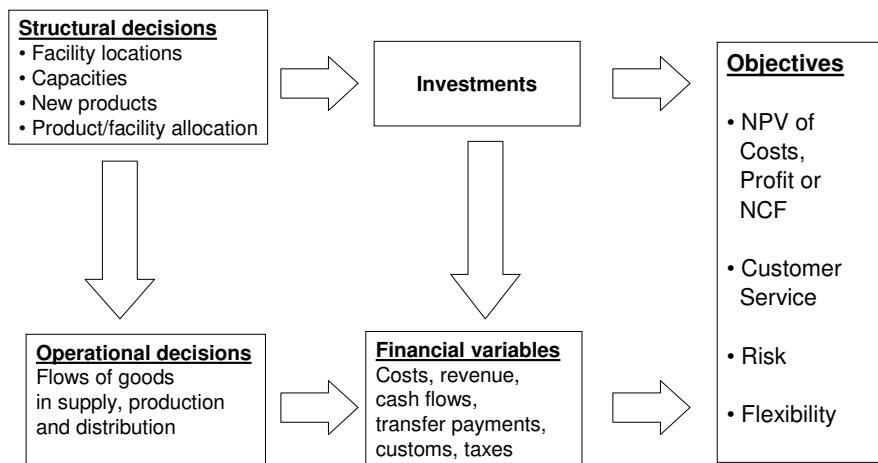


Fig. 6.3. Interdependence between strategic and operational levels

caused by the fact that both planning levels essentially influence the objectives of strategic planning. The financial objectives are affected directly by the strategic decisions on investments as well as by the yearly financial variables resulting from the operations. The latter are also influenced by the investments. For instance, the investment of a new machine can change the variable production cost significantly. Customer service clearly depends on

the operational performance characteristics, e.g. lead times and reliability, but also on strategic location decisions. For instance a new production site or distribution center will tend to improve the customer service in the respective country.

As explained in the introduction to this chapter, all these interrelations are not deterministic, but depend on unknown future data. An important objective is to minimize the *risk*, which can be expressed as the variability of some financial objective. A further criterion is the *flexibility* of a structural design, i.e. its ability to adapt to unanticipated changes of the environment. Some of these changes can be modeled by different scenarios, e.g. changes of the demand of certain products (*volume flexibility*). Other aspects of flexibility, however, are difficult to quantify. For instance, installing a general purpose machine allows the production of future products that are not conceived yet (*product flexibility*). Thus, this decision contributes to greater flexibility than installing a dedicated machine for existing products.

The strategic decisions determine a network design alternative, which can be evaluated by the various objective criteria in Fig. 6.3. The value of each criterion also depends on the scenario and on the (future) decisions on the appropriate operations. In addition, there often exist unquantifiable objectives and constraints formulated on the corporate level, e.g. the political stability of a country, where a new facility is installed. This multi-objective stochastic problem does not have an optimal solution. Optimization is only possible, if we restrict ourselves to a single objective, e.g. the expected value of the *NPVNCF* or a weighted sum of several objectives. In either case the stochastic design requires that the probabilities of the scenarios or the probability distributions of the model parameters are known.

In the general case of several, partly unquantifiable objectives and unknown probabilities, the strategic planning process iteratively runs through the following steps as shown in Fig. 6.4 (see Ratliff and Nulty, 1997):

Generate alternatives: Specifying objectives and scenarios and solving the optimization problems defined above provides various design alternatives. In order to keep the optimization tractable, an aggregate model should be used, which contains only a rough approximation of the operational level. Objectives that are not used in the current optimization can be considered in form of aspiration level constraints. Additional alternatives can be generated by intuition and managerial insight.

Evaluate alternatives: For any design alternative, the operations can be optimised using a more detailed operational model, that incorporates various scenarios. The main objective on this level is cost or profit, since the network configuration is considered given. A more detailed analysis can be obtained by simulating the operations. This allows the incorporation of additional *operational uncertainties*, e.g. the short-term variation of the demand or of the availability of a machine, resulting in the more accurate computation of performance measures such as service levels or flow times.

Benchmarking: The key performance indicators obtained in the evaluation step are compared to the best-practice standards of the respective industry. This provides an additional evaluation of the quality of a supply chain configuration and allows a rudimentary validation of the proposed configuration.

Select alternatives: Finally, the performance measures obtained in the previous steps and the consideration of additional non-quantifiable objectives can be used to eliminate inefficient and undesirable configurations. This is done based on discussions internal to the project team and by presentations to the final decision makers such as the board of directors. During this process, suggestions for the investigation of additional scenarios and objectives or modified alternatives may arise. This whole process may go through several iterations.

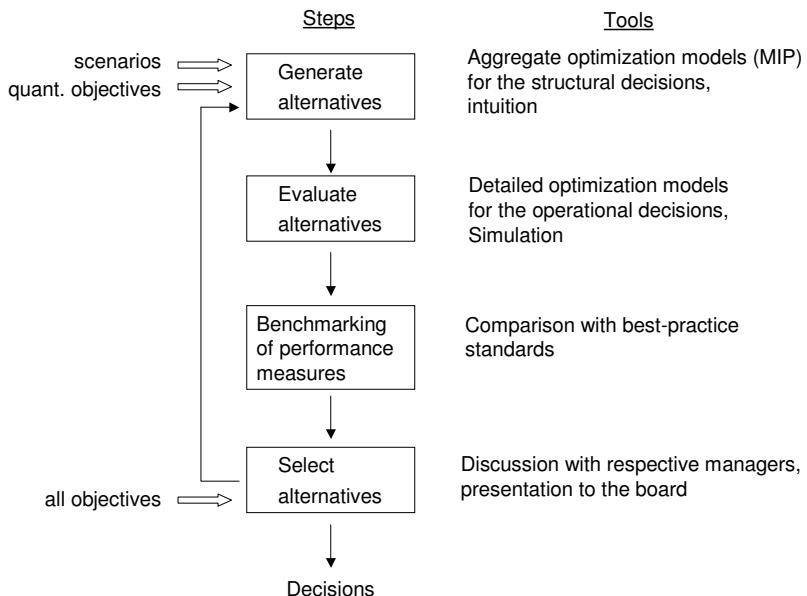


Fig. 6.4. Steps of the strategic network design

6.3.2 Modelling Aspects

In Section 6.1, the basic model for strategic network planning has been explained. The strategic decisions are expressed by binary variables indicating that a particular facility for certain products is established at a certain location and the operational model is a multi-commodity network flow problem. In the following, some important extensions of the basic model are discussed in more detail, based on models in literature. In particular we refer to the use of SNP models at a major German car manufacturer. These models extend

over a 12 year planning horizon and include the global production network with 10-14 factories for assembly and engines, about 40 product groups, a few engine and material types, and 5-10 regions of suppliers and sales markets. Henrich (2002) reports on the first phase of the development of strategic network planning models and their use within the strategic planning process. In a current project, refinements of the models are worked out.

Single period - several periods

Single-period models are appropriate for decisions on the immediate re-organisation of parts of a supply chain, e.g. new distribution centers or the investment of single machines in a factory. These models usually consider the operational level in more detail, e.g. the single transportation or production processes. Typical single period models are the classical production-distribution models (see Geoffrion and Powers, 1995).

The subdivision of the planning horizon in several periods, usually years, permits to model the development of a comprehensive national or global supply network from the present status into a future one. Moreover, the investment expenditures can be distributed over several years, as explained below, resulting in a dynamic cash flow and the corresponding NPV. However, the consideration of inventory carried from period to period, as it is usual in operational planning, has no importance for strategic planning, because the end-of-period inventory in a multi-period model is a seasonal inventory protecting against temporary undercapacity and hence its development over yearly periods is not a strategic issue. Modelling inventory in strategic planning is discussed below.

New products

In the planning models, production and distribution activities and the installation of necessary capacities are driven by the given (estimated) demand per product and year. However, the actual logic for new products is inverse: The demand for a particular new product is only created by the decision to launch this product starting from the year of the launch and developing according to its life cycle. Therefore, models with given demand can support the decisions whether or not to develop a new product and where to produce it, but not when to launch it. The latter decision requires a variable allocation of the life cycle demand to the years. This type of model has only been proposed by Popp (1983). In the automotive case, the years when new products go into serial production are fixed in advance for the 12-year horizon and life cycles of about 7 years. Products with very short life cycles, e.g. in the electronics industry, often are successors of other products with very similar functionality. In this case old and new products can be aggregated to one product group and share the same demand which removes the above difficulty.

Investment Expenditures

For the calculation of the NCF according to (6.3), it is noteworthy, that the expenditure of an investment is usually spread over several years before and after the time of installation. This can be modelled by means of the usual binary variables y_{it} for the decision on installing investment i in year t (see Popp, 1983): Let $I_{i,\delta}$ be the expenditure for investment i in the year δ before ($\delta < 0$) or after ($\delta \geq 0$) the installation. Then the contribution to the cashflow in year t is $\sum_{\tau} I_{i,t-\tau} y_{i\tau}$, where the summation index τ is the year of installation. This kind of modelling is important for the large investments for new products in the automotive industry and is currently implemented in the car manufacturer model, whereas the model reported by Henrich (2002) still contains pure cost objectives.

Inventories

The structural decisions may have significant impact on the inventories in the supply network. The *pipeline inventory* in an arc (i, j) with flowtime t_{ij} is $t_{ij}x_{ij}$, hence linear in the flow x_{ij} , as considered by Arntzen et al (1995) and Vidal and Goetschalckx (2001). However, the flow time t_{ij} , in particular the waiting time contained in it, may itself depend on the capacities and their utilization. This complex interrelation can only be considered in a detailed operational model.

Cycle stock is caused by a flow x_{ij} that occurs in intermittent batches. If the frequency is given, say f_{ij} batches in the considered period, the average cycle stock is $\frac{1}{2}x_{ij}/f_{ij}$ in both nodes i and j and can easily be incorporated into an SNP model (see Vidal and Goetschalckx, 2001). If (i, j) is a transport process that is carried out with the minimal frequency f_{ij} or whenever the vehicle capacity Q is reached, then the cycle stock becomes $\frac{1}{2}\min(x_{ij}/f_{ij}, Q)$. This non-convex term expresses a bundling effect of transport flows, but can only be modelled by means of a binary variable.

Seasonal stock is not contained in an SNP model with yearly periods. However, it depends on the structural decisions on capacities. This relationship is not considered in literature, maybe because most multi-period models contain end-of-period stocks which are misinterpreted as seasonal inventory. A simple linear approximation is possible: Let x_k the total production quantity in a certain year in a facility k with capacity C_k which may depend on the structural decisions. If $\gamma > 1$ is the seasonal factor, i.e. the peak demand over the average demand per week, no seasonal stock is necessary, as long as $\gamma x_k \leq C_k$. For $x_k = C_k$, i.e. 100% utilization, the average seasonal stock S_{100} can be calculated from the seasonality pattern according to the procedure in Sect. 2.4 (see Fig. 2.9). Linear interpolation for $C_k/\gamma < x_k < C_k$ leads to the seasonal stock

$$\max \left(0, \frac{\gamma x_k / C_k - 1}{\gamma - 1} S_{100} \right) \quad (6.13)$$

which is piecewise linear and convex in x_k .

Safety inventory is influenced by the structural decisions via the flow times and the number of stock points in the network. This relationships nonlinear and depends on many other factors such as the desired service level and the inventory policy. The consideration is only possible on the operational level. A recent survey of safety stock placement in a supply network is given by Graves and Willems (2003).

In the automotive case, only pipeline inventory of finished cars in transit is considered, which is significantly different for domestic and overseas shipping. Other types of inventory are of minor relevance, as production and a major part of materials supply is JIT controlled.

6.4 SNP Modules in Advanced Planning Systems

An SNP module in an APS has to take into account the operational planning level for the aggregate flows of goods along the supply chain, as shown in Fig. 6.3. Therefore, it can be used for Master Planning (see Chap. 8) as well and is, in some APS, identical with the Master Planning module. It always provides a modeling feature for a multi-commodity multi-period flow network, as explained in Section 6.1. This usually includes the end-of-period inventories as variables, which are only relevant for Master Planning, but not for SNP (see 6.3.2). In addition, an SNP module permits to model the strategic decisions on locations, capacities and investments by means of binary variables.

SNP modules contain an LP solver, which is able to determine the optimal flows in a given supply chain for a given objective, even for large networks and a large number of products and materials. However, the strategic decisions require a MIP solver, which is also available in the SNP module, but may require unacceptably long computation time for optimizing these decisions. Therefore, SNP modules also provide various heuristics which are usually proprietary and not published.

In contrast with other modules, the SNP module is characterized by a rather low data integration within the APS and with the ERP system. Therefore, it is often used as a stand-alone system. Current data of the stocks and of the availability of the machines are not required for SNP. Past demand data from the ERP data base can be useful, but they need to be manipulated for generating demand scenarios for a long-term horizon. Technical data of the machines, like processing times and flow times, can be taken from the ERP data as well. But a major part of the data required for SNP, such as data on new products, new markets and new machines, is not available in the ERP system. The same is true for data on investments, such as investment expenditures, depreciation and investment limitations.

The modeling tools that are available in the SNP module differ in the various APS. Some APS contain a modeling language for general LP and MIP models, which allows the formulation of various types of models as discussed

in the previous sections. However, it should be noted that this unlimited modeling flexibility is restricted by the computational solvability of the models with a general-purpose MIP solver. Specialized algorithms, which are mostly heuristics, can solve larger problems. However, they are based on predefined model structures and characteristics, such as the type of investment decision, the objective, and the type of fixed and variable costs. Other APS provide pre-formulated components of an SNP model, which describe typical production, warehousing and transportation processes. They allow the rapid assembly of a complex supply chain model, even using click-and-drag to construct a graphical representation on the screen and without LP/MIP knowledge. Of course, this entails a loss of flexibility in the models that can be formulated. But the resulting models are easy to understand and can be well explained to the decision makers. Thus, there is a trade-off between modelling flexibility, which requires thorough LP/MIP expertise, and the ease of building and understanding models.

The SNP modules with predefined model components, as well as the corresponding special-purpose algorithms, do not cover all the modeling aspects discussed in Section 6.1 and 6.3.2. In the automotive case, Henrich (2002) has first used a SNP module with predefined components, in order to convince the management of the usefulness of the model in the strategic planning process. Currently, extensions for investment models are worked out and tested by means of a general LP/MIP modeling language and solver. It is intended to transfer the model into an appropriate SNP module afterwards.

An SNP module provides the following main functions within the framework of the strategic planning process explained in Section 6.3.2 and 6.4:

- Generating alternatives,
- Evaluating alternatives,
- Administrating alternatives and scenarios,
- Reporting, visualizing and comparing results.

The last two functions are particularly important in the iterative strategic planning process involving large series of design alternatives and scenarios. These functions and possibly, modeling aids and special algorithms for network design, make up the essential advantages over a general LP/MIP software system.

6.5 Conclusions

The tradeoff between model solvability and model realism will always remain. The more realistic the model is the more resources have to be allocated for model development and validation, data collection and validation, model maintenance, and model solving. Since all models involve some level of abstraction, approximations, and assumptions, the results of the models should

always be interpreted carefully with common (engineering) sense. Different models with different levels of detail and realism are appropriate and useful at different stages of the design process. Systematically increasing the level of model complexity for the same problem and evaluating their solutions and their consistency provides a way to, at least partially, validate the models. The common thread among successful applications of model-based strategic supply chain design is the sustained effort of a group of highly specialized designers, who exploited the structure of the problem to generate a formulation of acceptable size and degree of realism and a solution algorithm that had an acceptable computation time. This model then became a strategic asset for the corporation that greatly increased the performance of the designed supply chain and speed of the design.

The more complex the real world system is, the more approximate any model will become. Models used to assist in strategic decision making are infamous for not capturing many of the real world factors and subjective influences (see Vidal and Goetschalckx (1996)). Such strategic models should only be used as decision support tools for the design team. A healthy skepticism with respect to the results of any model is required. Just because a computer model specifies a particular decision, does not imply that this is the best decision for the real world system.

To remain competitive global corporations need a methodology to evaluate and efficiently configure robust global logistics systems in a short amount of time. The methodology to design the best possible supply chain configuration for global logistics systems does not appear to exist in the current generation of APS systems. However, the research trend is towards an integration and combination of the features of the domestic and global models, which will allow the simultaneous optimization of facilities and production-distribution-inventory flows in a global logistics system. A second trend is towards the design of flexible and robust supply chains that are based on possible scenarios and stochastic data. Significant challenges exist with respect to determining near-optimal solutions for these models. Case studies have provided ample evidence that the use of such a model and solution methodology can yield significant savings for a corporation interested in expanding globally.

A drawback of the newer models and solution algorithms is the significant level of technical expertise required to achieve the fast solution times. A very important area of future research is the standardization and technology transfer process of these solution methodologies so that they can be more widely applied. Global corporations implement ERP systems and Business Data Warehouses at ever increasing rates, providing APS and decision-makers with the basic data and information necessary for Strategic Network Planning. It can be expected that models and methodologies currently available in APS will become more versatile in the near future and incorporate some of the features currently only discussed in the academic literature. This will

allow these global corporations to use this information in a timely fashion to significantly increase their profits and to remain competitive.

References

- Arntzen, B. C., Brown, G. G., Harrison, T. P., Trafton, L. L. (1995) *Global supply chain management at Digital Equipment Corporation*, Interfaces, Vol. 25, No. 1, 69–93
- Ballou, R. H., Masters, J. M. (1993) *Commercial software for locating warehouses and other facilities*, Journal of Business Logistics, Vol. 14, No. 2, 71–107
- Ballou, R. H., Masters, J. M. (1999) *Facility location commercial software survey*, Journal of Business Logistics, Vol. 20, No. 1, 215–233
- Choi, F. D. S. (1997) *International Accounting and Finance Handbook*, 2nd ed., Somerset, New Jersey
- Dogan, K., Goetschalckx, M. (1999) *A Primal Decomposition Method for the Integrated Design of Multi-Period Production-Distribution Systems*, IIE Transactions, Vol. 31, No. 11, 1027–1036
- Geoffrion, A. M., Powers, R. F. (1995) *20 Years of Strategic Distribution System Design: an Evolutionary Perspective*, Interfaces, Vol. 25, No. 5, 105–127
- Graves, S. C., Rinnooy Kan, A. H. G., Zipkin, P. H. (Eds.) (1993) *Handbook in Operations Research and Management Science, 4: Logistics of production and inventory*, Amsterdam
- Graves, S. C., Willems, S. P. (2003) *Supply Chain Design: Safety Stock Placement and Supply Chain Configuration*, Chapter 3 in: de Kok, A. G., Graves, S. C. (Eds.) *Handbooks in Operations Research and Management Science*, Vol. 11, Amsterdam et al.
- Henrich, P. (2002) *Strategische Gestaltung von Produktionssystemen in der Automobilindustrie*, Aachen
- Kleywegt, A. J., Shapiro, A., Homem-de-Mello, T. (2002) *The sample average approximation method for stochastic discrete optimization*, Siam Journal on Optimization, Vol. 12, No. 2, 479–502
- Kouvelis, P., Yu, G. (1997) *Robust Discrete Optimization and Its Application*, London
- Papageorgiou, L. G., Rotstein, G. E., Shah, N. (2001) *Strategic supply chain optimization for the pharmaceutical industries*, Industrial and Engineering Chemistry Research, Vol. 40, No. 1, 275–286
- Popp, W. (1983) *Strategische Planung für eine multinationale Unternehmung mit gemischt-ganzzahliger Programmierung*, OR Spektrum, Vol. 5, No. 1, 45–57
- Ratliff, H. D., Nulty, W. G. (1997) *Logistics composite modeling*, Chapter 2 in: Artiba, A., Elmaghhraby, S. E. (Eds.) *The planning and Scheduling of Production Systems*, London
- Santoso, T., Ahmed, S., Goetschalckx, M., Shapiro, A. (2003) *A stochastic programming approach for supply chain network design under uncertainty*, accepted for publication in the European Journal of Operational Research, E-Print available at www.optimization-online.org
- Schmidt, G., Wilhelm, W. E. (2000) *Strategic, tactical and operational decisions in multi-national logistics networks: a review and discussion of modeling issues*, International Journal of Production Research, Vol. 38, No. 7, 1501–1523

- Shapiro, J. F. (2001) *Modelling the Supply Chain*, Duxbury Press
- Silver, E., Pyke, D., Peterson, R. (1998) *Inventory management and production planning and scheduling*, 3rd ed., New York
- Simchi-Levi, E. D., Kaminsky, P., Simchi-Levi, E. (2003) *Managing the Supply Chain: The Definitive Guide for the Business Professional*, New York
- Tayur, S., Ganeshan, R., Magazine, M. (Eds.) (1999) *Quantitative models for supply chain management*, Boston, Massachusetts
- Vidal, C. J., Goetschalckx, M. (1996) *The role and limitations of quantitative techniques in the strategic design of global logistics systems*, CIBER Research Report 96-023, Georgia Institute of Technology. Accepted for publication in the Special Issue on manufacturing in a global economy of the *Journal of Technology and Forecasting and Social Change*
- Vidal, C. J., Goetschalckx, M. (1997) *Strategic production-distribution models: A critical review with emphasis on global supply chain models*, European Journal of Operational Research, Vol. 98, 1–18
- Vidal, C. J., Goetschalckx, M. (2001) *A Global supply chain model with transfer pricing and transportation cost allocation*, European Journal of Operational Research, Vol. 129, No. 1, 134–158

7 Demand Planning

Michael Wagner

Paul Hartmann AG, Corporate Materials Management, P.O.Box 1420,
89504 Heidenheim, Germany

Large benefits which are achieved by Supply Chain Management are accredited to the reduction of inventories, esp. to the decrement of safety stocks. While safety stocks are mainly influenced by uncertainty, it is appealing that most effort should be spent on the reduction of uncertainty. Two sources of uncertainty are known in supply chains:

- process uncertainty (e. g. unreliable production processes, fluctuating lead-times etc.) and
- demand uncertainty (difference in planned or estimated demand and actual sales).

The purpose of Demand Planning is to improve decisions affecting demand accuracy and the calculation of buffer or safety stocks to reach a predefined service level. All decisions in the whole supply chain should be based on already fixed (accepted) customer orders *and* planned sales or forecasts, the latter ones are determined in the Demand Planning process. Therefore, the performance of each supply chain entity depends on the quality of the demand plan. This also implies that these figures need to be the result of a collaborative effort.

The next section explains the different tasks of a demand planning process and shows which of the tasks should be supported by an APS.

7.1 A Demand Planning Framework

What are the results of Demand Planning and who needs it? This should be the first question in a Demand Planning implementation. Figure 7.1 shows all Demand Planning tasks and its planning horizon using the output of Demand Planning. For example, mid-term Master Planning will require demand forecasts for every product group, sales region and week and the safety stocks (as minimum stock levels) for every distribution centre or plant location. On the other hand, short-term replenishment decisions for finished products are based on daily forecasts for every product. Therefore, it is necessary to define the requirements of all planning tasks before deciding on Demand Planning tasks and their application. The structure of the forecasting part of Demand Planning depends heavily on the results one wants to get from it. Additionally the selection of forecasting methods requires knowledge of the corresponding forecasting horizon and the level of detail.

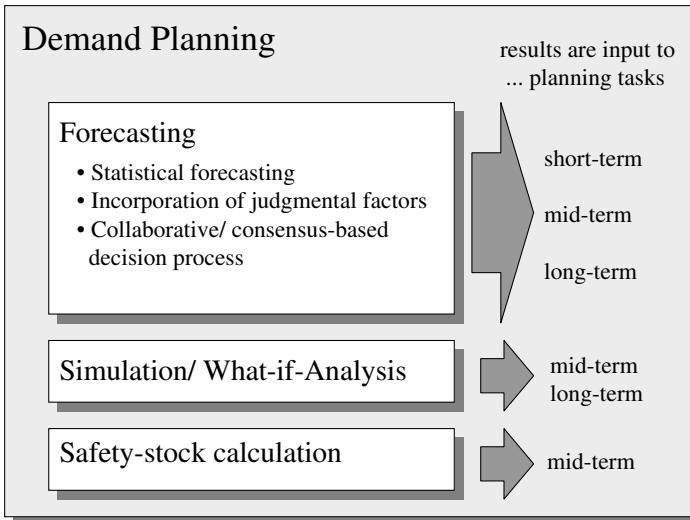


Fig. 7.1. Demand Planning tasks

The forecasting module of Demand Planning comprises the following three planning tools:

- Statistical forecasting uses sophisticated methods to create forecasts for a lot of items automatically. This might be the first step in the Demand Planning process and captures the main characteristics of the time series. A few statistical forecasting methods are described in the following Section.
- The second step uses the statistical forecasts and adds information to the time series which was not respected by the previous step. This input comprises information on promotions, marketing campaigns, change in the number of stores etc. and can be considered by manual corrections of the forecast or with the aid of the software tool. Here, the user provides the information on when the factor influences the forecast (e.g. promotion in sales region 3 in week 4) and the Demand Planning module calculates the corresponding quantities from former causal influences.
- As described above, the forecasting process has to be supported by a lot of supply chain members from different functional areas (sales, production, procurement etc.). Therefore, an efficient collaboration process is necessary to obtain a result which is accepted by all participants. The outcome of this process is a consensus-based forecast that is used for every planning step in the whole supply chain.

Forecasting, as described above, is not a real planning or decision process as it “only” aims at predicting the future as accurately as possible. But it does not influence the demand and therefore, for example, views the decisions on

promotions as being given. Hence, changing demand requires an additional module: simulation/what-if-analysis. This tool enables the user to view the consequences of different scenarios. This allows to plan promotions (when and where?), the shape of the life-cycle curve or decide on the point in time at which a new product will be launched.

So far, forecasting appears to be quite easy and seems to be a reliable tool to create reliable input to the processes distribution, production and procurement. But, as Nahmias (1997) argues in his textbook, the main characteristic of forecasts is that they are usually wrong! Therefore, each planning step which is based on Demand Planning data contains uncertainty to some extent. The difference between the production or distribution quantity (result of planning based on forecasts) and actual sales (customer orders) influences the service level of the whole supply chain. As this service level usually can not reach 100%, safety stocks are an adequate tool for improving customer service. The amount of safety stock required for reaching a desired service level is closely linked to forecasting, as the forecast error is input to the calculation.

Demand Planning means predicting future sales; therefore, it is necessary to incorporate all information available in a supply chain which could be relevant. But this information is often only specific and stored decentrally. For example, a sales person may “only” provide input to the forecasting process of the products and the sales region he/she is responsible for. All information pieces finally should add to a forecast which covers the whole demand being served by the supply chain. On the other hand one must be able to retrieve forecasts aggregated for special purposes, like demand figures aggregated to product groups and weeks for Master Planning. Therefore, the database of Demand Planning has to support at least the following three dimensions (see also Fig. 7.2) of aggregations and disaggregations:

- product dimension: product → product group → product family → product line;
- geographic dimension: customer → sales region → DC region/location;
- time dimension: different bucket size (days → weeks → years) and horizon.

The product dimension is structured in a hierarchy, relating products to product groups, groups to families, families to product lines etc. The product hierarchy ranges from the top level representing all products to the lowest product level on which a forecast is created. The geographic dimension is also hierarchically structured. It represents the structure of the market, for example regions, countries, industry branches, key accounts etc. The third dimension of forecasting is time. The time dimension is normally structured by years, quarters and months. In some cases it is necessary to go down to the week or even day. Forecast quantities can be attached to any intersection of product, geography and time.

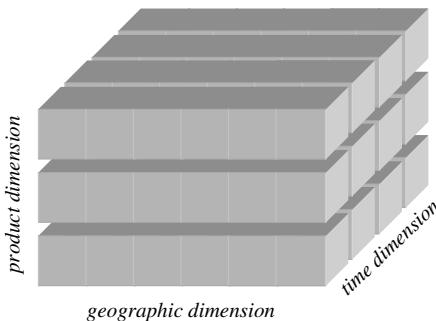


Fig. 7.2. Three-dimensional structure of demand planning data

Such three-dimensional databases grow very fast even for mid-sized companies. State-of-the-art database technology, as efficient relational systems or Data Warehouses plus modern OLAP (online analytical processing) tools, build the bottom line of a high-performance Demand Planning solution.

But, as many people work on one forecast it is evident that all this subjective information could be contradictory. This contradiction is either resolved by a hierarchical structure of forecast responsibilities or by a collaboration process among the concerned planners.

Responsibilities for demand planning tasks can be organized in process chains. For each process the characteristics (e.g. frequency of the process, peculiarities of the three dimensions) need to be defined during the implementation. Typical process chains for mid-term forecasting in the computer industry and in the food industry are described in the following:

Computer Industry

Sales forecast (dimensions: regions, finished product groups, months)

The sales force builds the sales forecast which estimates the quantities for the sales region they serve. This forecast is aggregated to product groups (e.g. high-end business PC), because the product which will be offered is determined afterwards.

Headquarter forecast (dimensions: all regions, finished products, months)

In the headquarter the sales forecast for the distinct regions is added up and refined according to single products. This is possible since the headquarter forecast is a collaborative process of representatives from sales, marketing, material management etc.

Component forecast (dimensions: all regions, components, months)

The forecast for the components which have to be procured from suppliers is not computed by applying the bill-of-materials (BOM) to the finished product forecast. This is not possible, since the BOM change very fast, because the life-cycle of certain components (e.g. hard disks) is shorter than the life-cycle of the computer. Therefore, the planner forecasts the

percentage to which a specific component is used in a finished product and derives the component forecast therefrom.

Food Industry

Sales forecast (dimensions: regions, finished products, weeks)

In the food industry the sales forecast is already detailed to single products and weeks. It also comprises additional demand which is expected to result from promotional activities (e.g. special prices for specific customers) in the sales regions.

Collaborative forecast (dimensions: all regions, finished products, weeks)

The sum of the sales forecasts is the basis for the collaborative forecast. It considers additional information on marketing activities (e.g. TV-spots), the launch of new products and the activities of competitors. The collaborative forecast is made by managers from sales, production, marketing and procurement.

Procurement forecast (dimensions: all regions, components, weeks)

The procurement forecast in the food industry can be calculated easily by applying the BOM to the collaborative forecast.

7.2 Statistical Forecasting Techniques

Forecasting methods were developed since the 1950's for business forecasting and at the same time for econometric purposes (e.g. unemployment rates etc.). The application in software modules makes it possible to create forecasts for a lot of items in a few seconds. Therefore, all leading APS vendors incorporate statistical forecasting procedures in their demand planning solution. Each of these methods tries to incorporate information on the history of a product/item in the forecasting process for future figures. But there exist two different basic approaches – time-series-analysis and causal models. The so-called *time-series-analysis* assumes that the demand follows a specific pattern. Therefore, the task of a forecasting method is to estimate the pattern from the history of observations. Future forecasts can then be calculated from using this estimated pattern. The advantage of those methods is that they only require past observations of demand. The following demand patterns are most common in time-series-analysis (see Silver et al. (1998) and also Fig. 7.3):

1. Level model: The demand x_t in a specific period t consists of the level a and random noise u_t which cannot be estimated by a forecasting method.

$$x_t = a + u_t \quad (7.1)$$

2. Trend model: The linear trend b is added to the level model's equation.

$$x_t = a + b \cdot t + u_t \quad (7.2)$$

3. Seasonal model: It is assumed that a fixed pattern repeats every T periods (cycle). Depending on the extent of cyclic oscillations a multiplicative or an additive relationship can be considered.

$$x_t = (a + b \cdot t) + c_t + u_t \text{ additive model,} \quad (7.3a)$$

$$x_t = (a + b \cdot t) \cdot c_t + u_t \text{ multiplicative model} \quad (7.3b)$$

where $c_t = c_{t-T} = c_{t-2T} = \dots$ are seasonal indices (coefficients).

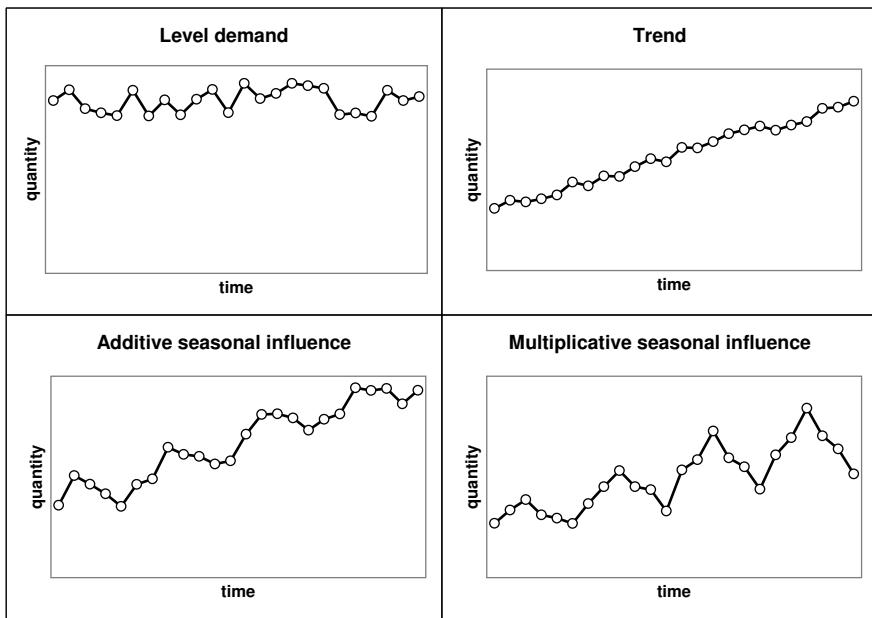


Fig. 7.3. Demand patterns

The second approach to statistical forecasting are *causal models*. They assume that the demand process is determined by some known factors. For example, the sales of ice cream might depend on the weather or temperature on a specific day. Therefore, the temperature is the so-called leading indicator for ice cream sales. If enough observations of sales and temperature are available for the item considered, then the underlying model can be estimated. For this example, the model might consist of some amount of independent demand z_0 and the temperature factor $z^1(t)$

$$x_t = z^0 + z^1(t) \cdot w_t + u_t \quad (7.4)$$

where w_t is the temperature on day t .

As for parameter estimation in causal models the demand history and one or more time-series with indicators are needed, the data requirements are much higher than for time-series analysis. Furthermore, practical experience shows that simple time-series models often produce better forecasts than complex causal models (see e.g. Silver et al. (1998, pp. 130)). These tend to interpret stochastic fluctuations (noise) as “structure” and therefore, introduce a systematic error in the model.

In the following three paragraphs the characteristics and the approach of the most frequently used forecasting methods are described. The first part introduces the forecasting techniques used for time-series models.

7.2.1 Moving Average and Smoothing Methods

As each demand history is distorted by random noise u_t , the accurate estimation of parameters for the model is a crucial task. Also, the parameters are not fix and might change over time. Therefore, it is necessary to estimate under consideration of actual observations *and* to incorporate enough past values to eliminate random fluctuations (conflicting goals!).

Simple Moving Average The simple moving average (MA) is used for forecasting items with level demand (see Sect. 7.1). The parameter estimate for the level \hat{a} is calculated by averaging the past n demand observations. This parameter serves as a forecast for all future periods, since the forecast \hat{x}_t is independent of time. According to simple statistics, the accuracy of the forecast will increase with the length n of the time-series considered, because the random deviations get less weight. But this is no more applicable if the level changes with time. Therefore, values between three and ten often lead to reasonable results for practical demand series. But the information provided by all former demands is lost according to this procedure.

Exponential Smoothing The need to cut the time-series is avoided by the exponential smoothing method, because it assigns different weights to all (!) observed demand data and incorporates them into the forecast. The weight for the observations is exponentially decreasing with the latest demand getting the highest weight. Therefore it is possible to stay abreast of changes in the demand pattern and to keep the information which was provided by older values. For the case of level demand the forecast for period $t + 1$ will be calculated according to the following equation:

$$\hat{x}_{t+1} = \hat{a}_t = \alpha \cdot x_t + \alpha(1 - \alpha) \cdot x_{t-1} + \alpha(1 - \alpha)^2 \cdot x_{t-2} + \dots \quad (7.5)$$

The parameter α is the smoothing constant, to which values between 0 and 1 can be assigned. For $\alpha = 0.2$ the weights in Table 7.1 are being used, if the forecast has to be made for period 1. Furthermore it is not necessary to

Table 7.1. Weights of past observations in exponential smoothing for $\alpha = 0.2$

period	0	-1	-2	-3	-4	...
weight	0.2	0.16	0.13	0.10	0.08	...

store the whole history of an item as (7.5) can be simplified. The only data which needs to be kept in the database are the latest forecast and the latest demand value.

Exponential smoothing for level demand patterns is easy to apply and requires little storage capacity. Therefore, it provides good forecasts for this kind of model and it also calculates reasonable forecasts for items which are influenced by high random fluctuations (Silver et al., 1998).

The exponential smoothing procedure for level demand can be extended for trend models and multiplicative seasonal models (see (7.2) and (7.3b)). The method for the trend model is known as Holt's procedure (see e.g. Nahmias (1997)). It smoothes both terms of the model, the level a and the trend component b with different smoothing constants α and β .

Winters introduced the seasonal model with exponential smoothing. A lot of lines of business are facing seasonal patterns, but don't incorporate it in forecasting procedures. For example, consider the manager of a shoe store, who wants to forecast sales for the next two weeks in daily buckets. As sales are usually higher on Saturdays than on Mondays, he has to take the weekly "season" into account. Winters' method is an efficient tool to forecast seasonal patterns, because it smoothes the estimates for the three parameters a , b and c . In contrast to the former two models the seasonal method needs far more data to initialize the parameters. For reliable estimates for the seasonal coefficients it is necessary to consider at least two cycles of demand history (e.g. two years). For more details on Winters' model see Chap. 26.

7.2.2 Regression Analysis

Where significant influence of some known factors is present, it seems to be straightforward to use causal models in the forecasting process. Regression analysis is the standard method for estimation of parameter values in causal models. Usually linear dependencies between the dependent variable x_t (e. g. the demand) and the leading factors (independent variables; e. g. temperature, expenditures for promotions etc.) are considered. Therefore, a multiple regression model can be formulated as follows (see e. g. Hanke and Reitsch (1995)):

$$x_t = z_0 + z_1 \cdot w_{1t} + z_2 \cdot w_{2t} + \dots \quad (7.6)$$

The ice cream model in (7.4) is called the simple regression model, as it only considers one leading indicator. Multiple linear regression uses the method of least squares to estimate model parameters (z_0, z_1, z_2, \dots) . This procedure

minimizes the sum of the squared difference between the actual demand and the forecast the model would produce. While exponential smoothing can consider all past observations, the regression method is applied to a predefined set of data. The drawbacks of such a procedure are the same as for the moving average model. Further, the weight of all considered values equals one and therefore the model cannot react flexibly to changes in the demand pattern.

As the data requirements of linear regression models are much higher than for simple time-series models, it is obvious that this effort is only paid back, if the models are used for aggregate mid-term or long-term forecasts or for a few important end products.

The following example shows the application of linear regression for the ice cream model: Assuming that the ice cream retailer observed the following demands and temperatures ($^{\circ}\text{C}$) over 10 days (Table 7.2) the linear regression

Table 7.2. Demand and temperature data for the ice cream example

period	1	2	3	4	5	6	7	8	9	10
actual demand	43	45	54	52	54	55	43	33	52	51
temperature ($^{\circ}\text{C}$)	15	17	19	16	21	22	18	15	19	18

will calculate the equation

$$\text{demand } x_t = 8.24 + 2.22 \cdot w_{1t} \quad (7.7)$$

with w_{1t} being the temperature on day t . Using (7.7) one can determine the forecasts (model value) which the model would have produced (see Table 7.3). But, for this it is necessary to be able to estimate the temperature reliably. Figure 7.4 shows the data and the resulting forecasts for the ice cream model.

Table 7.3. Example forecasts using the linear regression model

period	1	2	3	4	5	6	7	8	9	10
model value	42	46	50	44	55	57	48	42	50	48

7.2.3 ARIMA/Box-Jenkins-method

While both model-types described above assume statistical independence of demand values in different periods, the autoregressive integrated moving-average (ARIMA) models explicitly consider dependent demands. Therefore, these methods don't make assumptions about the underlying demand pattern, but compose a function from different building-blocks which fits the

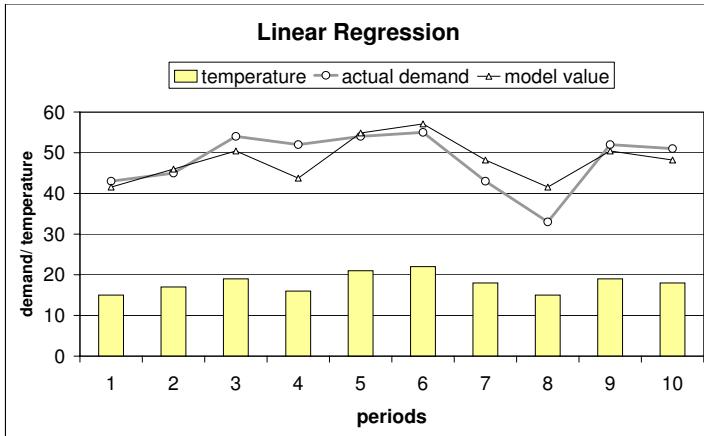


Fig. 7.4. Linear regression: results for the ice cream model

observed data best. The model function is found by iteratively executing the following three steps (see Hanke and Reitsch (1995)):

1. *Model identification:* An appropriate ARIMA model is selected by comparing the autocorrelation of the theoretical distributions and the observed autocorrelation. Autocorrelation states the existence of correlation between the actual demand and observations somewhere back in time.
2. *Model estimation:* As in regression models the parameters of the forecasting function have to be estimated in ARIMA models. Therefore, the procedure searches for values which minimize the mean squared error of the model.
3. *Testing of the model:* If the error term is purely random and independent, then the model is assumed to be reliable.

As one can imagine from the above procedure, the Box-Jenkins-method requires much experience and significant input from the demand planner. Further, the initial estimation of the model should be based on at least 50 observations of demand. Therefore, ARIMA models might be suitable only for some important A-class items or for mid-term aggregate forecasts. But, if ARIMA models are utilized, the quality should be better than for simple time-series models or even causal models.

7.3 Incorporation of Judgmental Factors

Implementing an efficient demand planning process usually comes along with supporting the human planner with software tools. When it comes to statistical methods and their application one typical question arises: How is the

software able to make better forecasts than a human planner with years of experience in demand planning? The simple answer is that mathematical methods are unbiased. Empirical studies (see e.g. Makridakis et al. (1997)) give evidence, that bias is the main reason why myopic statistical methods often produce better result. But that's only half of the truth, because information on specific events or changes (e.g. promotional activities, customer feedback on new products etc.) can lead to significant changes in demand patterns which might not be considered in standard time-series models. Therefore, it is necessary to combine the advantages of both worlds in an integrated demand planning process.

For example, consider the demand planning process of a company selling mineral water. In such an environment the regular demand can be forecasted by a seasonal time-series model quite accurately. But, the demand series are distorted by occasional additional demand due to promotional activities in some retail outlets. This effect can be estimated by the sales force responsible for the promotion, while the base line is forecasted by a fitting statistical model.

Integration of statistical and judgmental forecasting is only reasonable, if information inherent in a statistical forecast is not considered in the judgmental process. In this case the information would be double counted and therefore the demand would be overestimated.

In the following we describe some methods on how to integrate statistical forecasting and *structured* judgment. Non-structured judgment is often applied by demand planners, if they check the figures produced by a decision-support tool and “tune” the values using their sure instinct. But, for integration purposes it is necessary to structure judgment. Detailed process definitions and guidelines create a framework for such a structured judgment. Armstrong and Collopy in Wright and Goodwin (1998) describe the following five procedures for the integration:

- Revised judgmental forecasts:

The first step in this procedure is made by demand planners, who create judgmental forecasts based on the knowledge of relevant data (e.g. historical data, causal factors etc.). Afterwards they are confronted with forecasts which are calculated using statistical methods. Then the planners have the possibility to revise their initial estimate incorporating the new information. But, there is no predefined percentage to which extent each of the components has to be considered in the final forecast. This procedure often leads to more accurate forecasts than simple judgment not aided by statistical methods. Furthermore, it has the advantage that it leaves the control over the demand planning process to the human planner.

- Combined forecasts:

As the above procedure assigned variable weights to the two forecasts, it is evident that these values are often biased or influenced by political

means. A more formal procedure is assured by combining the two values according to a predefined weighting scheme. Even if equal weights are assigned to judgmental forecasts and statistical forecasts, better results are possible.

- Revised extrapolation forecasts:

Modifying statistical forecasts manually to take specific domain knowledge of the planner into account is common practice in a lot of companies. But, the revision process has to be structured accordingly. This means that the judgmental modification has to be based on predefined triggers (e.g. promotions, weather etc.).

- Rule-based forecasts:

Rule-based forecasts are also based on statistical forecasts. But, the selection or combination of different forecasting methods is supported by structured judgments of experts. The rules used for the selection are derived from the specific knowledge of the experts or on past research. They are based on characteristics of time-series or on causal factors. Rule-based forecasting improves simple extrapolation methods especially, if the series have low variability and low uncertainty.

- Econometric forecasts:

Regression models are referred to as econometric forecasting methods, if the model selection process and the definition of causal variables is provided by structured judgment. Improvements are reported especially, if this procedure is applied to long-range forecasts. As bias could have much impact on the result of econometric forecasts, it is advisable to give the judgmental process a very rigid structure.

Structured judgment needs to be supported by detailed feedback mechanisms which show the planner the quality of his input. Therefore, error reports have to differentiate between the quality of (automatic) statistical forecasting and judgmental forecasts.

7.4 Additional Features

The implementation of demand planning needs to address the specific requirements of the partners working together in the supply chain. Therefore, it is essential to understand the time-series under consideration. Depending on the characteristics of the series and the dimensions (time, product, geography) of the forecast different forecasting procedures have to be utilized. Some of the implementation issues which have to be taken into account are described in the following paragraphs. In the last part of this Section additional functionality for specific demand planning tasks is introduced. These are not necessary in each branch of business, but they would be helpful if those problems are present.

7.4.1 Sporadic Demand

We call a time-series sporadic (intermittent), if no demand is observed in quite a lot of periods. Those demand patterns especially occur for replacement parts or if only a small part of the demand quantity is forecasted; for example the demand for jeans in a specific size on one day in a specific store might be sporadic. The usage of common statistical forecasting methods would produce large errors for those items. Additional judgmental forecasting would not increase the quality, because the occurrence of periods with no demand is usually pure random and therefore not predictable. Furthermore, sporadic demand often occurs for a large amount of C-class items, for which it would be appreciable to get forecasts with low costs and low time effort for human planners.

Therefore, efficient procedures for automatic calculation of forecasts for sporadic demand items were developed. These methods try to forecast the two components “occurrence of period with positive demand” and “quantity of demand” separately. For example, Croston’s method (see Silver et al. (1998) or Tempelmeier (1999)) determines the time between two transactions (demand periods) and the amount of the transaction. The update of the components can then be done by simple exponential smoothing methods. Significant reduction of the observed error is possible, if the sporadic demand process has no specific influence which causes the intermittent demand pattern. For example, the frequent occurrence of stockouts in a retail outlet could produce a time-series that implies sporadic demand.

7.4.2 Lost Sales vs. Backorders

Forecasts are usually based on the demand history of an item. But, while industrial customers (B2B) often accept backorders, if the product is not available, the consumer (B2C) won’t. Therefore, the amount of observed sales equals the amount of demand in the backorder case, but in the lost-sales case the sales figures might underestimate the real demand. For forecasting purposes the demand time-series is needed and therefore, must be calculated from the observed sales figures. This problem frequently occurs, if forecasts for the point-of-sales (retailers, outlets) should be calculated.

There are two generally different solution approaches for the problem of forecasting in presence of lost sales: The first one tries to calculate a virtual demand history which is based on the sales history and the information on stock-outs. The forecasts can then be computed on the basis of the virtual demand history. This approach delivers good results, if the number of stock-outs is quite low. An alternative solution to the lost-sales problem is the usage of sophisticated statistical methods which consider the observed sales as a censored sample of the demand sample (see e.g. Nahmias (1997)). For these methods it is necessary to know the inventory management processes which were/are applied for the products under consideration.

7.4.3 Measuring Forecast Accuracy and Triggering Exceptions

Why do we measure forecast accuracy? First of all, it is *not* a tool for the controller to check the quality of the demand planner's work. It is rather a building block in the demand planning process. The demand planner might check whether the statistical method is appropriate for the time-series, whether additional human judgment pays back or whether it is useful to incorporate information on promotions. In all cases a criterion is needed for the evaluation of his decisions. But, there are many ways to get the appropriate forecast accuracy.

All accuracy measures are based on the forecast error e_t . It is defined as the difference between the forecasted quantity \hat{x}_t and the actual quantity x_t : $e_t = \hat{x}_t - x_t$. This value of the forecast error is influenced by the following parameters:

- The time delta between forecast and actuals: Forecasting is aiming at providing information about future shipments, sales etc. Normally, it is easier to tell the nearer future than the future that is far away. Thus, the forecast accuracy strongly depends on the time between the forecast creation and the time period that is being forecasted. For example, consider a forecast for the sales volume in June this year. The sales forecast for the month of June that has been created in March normally has a lower accuracy than the forecast that has been created in May.
- The forecast granularity: The level of aggregation also has a strong impact on the forecast accuracy. Take sales forecast again as an example: It is easier to forecast the total sales volume for all products, for all geographic areas and for a complete fiscal year, than to forecast on a weekly basis low level product groups for all sales regions individually. Thus, the forecast accuracy normally decreases if the forecast granularity increases.

There are many methods to compute the forecast accuracy, based on the forecasting error e_t . Each measure is calculated for a fixed horizon n (in the past) which has to be defined by the planner. If the horizon is short, then the value reacts fast to deviations from the average, but then it also might fluctuate heavily due to random demand variations. The following three measures (see e.g. Silver et al. (1998)) are most common in practice and also in demand planning software:

$$\text{mean squared error (MSE)} = \frac{1}{n} \sum_{i=1}^n e_i^2 \quad (7.8)$$

$$\text{mean absolute deviation (MAD)} = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (7.9)$$

$$\text{mean absolute percentage error (MAPE)} = \left[\frac{1}{n} \sum_{i=1}^n \left| \frac{e_i}{x_i} \right| \right] \cdot 100 \quad (7.10)$$

The MSE is the variance of the forecast error in the time horizon under consideration. In the Linear Regression forecasting procedure the MSE is used as the objective function which is minimized. As the error is squared in the formula, large deviations are weighted more heavily than small errors. Whereas the MAD uses linear weights for the calculation of the forecast accuracy. Further, the meaning of the MAD is easier to interpret, as it can be compared with the demand quantity observed. The main drawback of the two measures above is the lack of comparability. The values of MSE and MAD are absolute quantities and therefore, cannot be benchmarked against other products with higher or lower average demand. But, the computation of the MAPE standardizes the value based on the observed demand quantities x_i . The result is a percentage-value for the forecast accuracy which is comparable to other products.

The measures described above allow detailed analysis of the past, but they need to be discussed from the beginning each time they are calculated. In demand planning tools for some 100 or 1000 items one wants an automatic “interpretation” of the forecast error and therefore, might need an alert or triggering system. This system should raise an alert, if the statistical forecasting procedure no more fits to the time-series or if the sales office didn’t provide the information on a sales promotion. Such an alert system can be triggered by thresholds which are based on one of the measures for the forecast accuracy. These thresholds are defined by the demand planner and updated under his responsibility. Besides the threshold technique some other triggering mechanisms have been developed which all are based on the forecast accuracy measured by MSE or MAD.

7.4.4 Model Selection and Parameter Estimation

The selection of a forecasting model and the estimation of necessary parameters¹ are issues which are raised in the implementation phase of demand planning or during the update of forecasting parameters. This update should be made more or less regularly (e.g. every year) but not too often, as this would result in too much nervousness. APS often provide some kind of automatic model selection and parameter estimation. Thereto, the user only has to define the time-horizon on which the calculation should be based. The system then searches all available statistical forecasting procedures and parameter combinations and selects the one which produces the best forecast accuracy in the specified time-segment. As a result the user gets a list with the forecasting method and the corresponding parameters for each product/item he should implement. Therefore the demand planner doesn’t have to check if a model fits the time-series under consideration (e.g. “Are the sales figures really seasonal or does the system only interpret random fluctuations?”) and can use the toolset of statistical methods like a black box.

¹ so-called pick-the-best option

But, practical experience shows that the long-term performance is better and more robust, if only 1-3 forecasting methods with equal parameter settings for a group of products are applied. This follows from the following drawbacks of the described automatic selection:

- The time-horizon should cover enough periods to get statistically significant results. But often the history of time-series is relatively short when demand planning is introduced first.
- The criterion for the evaluation is mostly one of the forecast accuracy measures described above. However, those values don't tell you anything about the robustness of the models' results.
- For the selection procedure three distinct time-segments are necessary:
 - In the first segment the models components are initialized. For example for Winters seasonal model 2-3 full seasonal cycles (e. g. years) are necessary to calculate initial values for the seasonal coefficients.
 - The second segment of the time-series history is used to optimize the parameter values. Therefore, the parameters are changed stepwise in the corresponding range (grid-search) and the forecast accuracy is measured.
 - The optimized parameters are used to get forecasts for the third time-segment which is also evaluated using the forecast accuracy. This accuracy value is then the criterion for the selection of the best forecasting model.

The setting of the length of each of the time-segments has significant influence on the result of the model selection. Mostly the user has no possibility to change those settings or even can not view the settings in the software.

Therefore, the automatic model selection can guide an experienced demand planner while searching for the appropriate proceeding. But, it is not suggestive to use it as a black box.

7.4.5 Life-Cycle-Management and Phase-in/Phase-out

In quite a lot of innovative businesses, like the computer industry, the life-cycles of certain components or products were reduced to less than a year. For example, high-tech firms offer up to three generations of a hard-disk every year. As common statistical forecasting procedures require significant demand history, it would take the whole life-cycle until useful results are gathered. But, since new products replace old products with almost the same functionality, it is plausible to reuse some information on the demand curve for the next generation.

Two main approaches are known in practice: The first one indexes the complete time-series and determines the life-cycle-factor which has to be multiplied with the average demand to get the quantity for a specific period in

the life-cycle (*life-cycle-management*). This method is able to stay abreast of arbitrary types of life-cycles. The only information needed for the application for new products is the length of the cycle and the estimated average demand. These two values are adapted continuously when observed demand data gets available during the “life” of the product.

The second approach (*phasing method*) divides the whole life-cycle in three phases. The “phase-in” describes the launch of a new product and is characterized by the increase of the demand according to a certain percentage (linear growth). Afterwards the series follows a constant demand pattern, as considered for the statistical forecasting procedures. During the “phase-out” the demand decrease along a specific percentage until the end of the life-cycle of the product. The only data necessary for the phasing model are the lengths of the phases and the in-/decrease-percentages.

For successful application of the above models it is necessary that the APS provides the functionality to build a “model library”. In this database life-cycles or phasing models are stored for each product group under consideration. Mostly only one life-cycle exists for the whole product group and this model is updated every time a life-cycle ends.

7.4.6 Safety Stocks

Most APS-providers complement their Demand Planning module with the functionality for safety stock calculation. This is intuitive since the forecast error is one of the major factors influencing the amount of stock which is necessary to reach a specific service level. The calculation of safety stocks is quite complex, as there exist many different formulas each for a specific problem setting. The demand planner’s task hereby is to check whether the prerequisites are met in his application. While this Chapter cannot provide a fully detailed overview on safety stocks and inventory management, we want to focus on the functionality which can be found in most APS. For further information the reader is referred to one of the inventory management books by Silver et al. (1998) or Nahmias (1997).

Most software tools offer safety stock calculations for “single-stage inventory systems”. This means that it is assumed that there exists only one single stocking point from which the demand is served. Multi-stage or multi-echelon systems (e. g. distribution chains with DC- and retailer-inventories) on the other hand have the possibility to store safety stocks on more than one stage.

For single-stage systems the amount of necessary safety stock ss is generally determined by the product of the standard deviation of the forecast error during the risk time σ_R and the safety factor k :

$$\text{safety stock } ss = k \cdot \sigma_R \quad (7.11)$$

Assuming that the variance of the forecast error in the future is the same as in the past, σ_R can be calculated by multiplying the standard deviation of

the forecast error² σ_e with the square root of the risk time \sqrt{R} . The length of the risk time depends on the inventory management system. The following two systems have to be distinguished:

- Periodic review system:

In such an environment the inventory position is reviewed only every t time periods (review interval). Each time the inventory is reviewed, an order is triggered and sent to the supplying entity (e.g. the production department, the supplier). The delivery is assumed to be available after the replenishment lead-time L . Therefore, the risk time equals the sum of the review interval and the replenishment lead-time: $R = L + t$.

- Continuous review system:

In continuous review systems the point in time at which an order is released is triggered by a predefined reorder point. If the inventory position falls below the reorder point, an order of a specific quantity q is released. The risk time in a continuous review system equals only the replenishment lead-time L : $R = L$.

But that is only half of the safety stock formula. The safety factor k represents all other determinants of the safety stock. In the following the determinants and some of their values are explained:

- Service level:

For the service level quite a lot of definitions exist. The most common ones are the following:

- cycle- or α -service level:

α is defined as the fraction of periods in which no stock-out occurs.

Therefore, the safety stock has to ensure the probability (which fits the companies business objectives) of no stock-out during the replenishment cycle;

- fill rate (β -service level): The fill rate is the order quantity of a product which can be met directly from stock;

- order fill rate: While the fill rate considers the smallest unit of measurement of a product, the order fill rate counts *complete* customer orders served from stock.

- Review interval or order quantity:

In periodic review systems the review interval is fixed and the order quantities depend on the estimated demand in an order cycle. For continuous review systems the opposite applies, as the order quantity is fixed and the length of the order cycle depends on the demand. But, if the demand is approximately level, both parameters can be converted in each other by the simple relation:

order quantity $q = \text{demand } d \cdot \text{cycle length } t$.

The required parameter can be calculated by minimizing the ordering

² calculated from past time-series

costs and the holding costs for the lot-sizing stock. This computation can be made by applying the well-known economic order quantity (EOQ)-formula (e.g. Silver et al. (1998)).

- Demand distribution function:

The distribution function of the observed demand is usually approximated by a standard distribution known from statistics. One of the most common distribution functions is the normal distribution. The distribution parameters (mean and variance) can easily be calculated from a sample of demands from the historic time-series.

All these parameters have to be combined in a formula which stays abreast the requirements of the business under consideration. Now, it should be clear that an APS-tool can only provide safety stock calculations if specific assumptions are met. But, if all parameters are user-definable the software can cover a wide range of different settings. Therefore, it is necessary to transfer the inventory management rules which are applied in the company to the standard parameters which are needed in the software. And that is the challenge of the demand planner.

As the above part on safety stock calculation should only give a short impression on the complexity of inventory management, the inspired reader can gather more information in one of the inventory management books listed at the end of this Chapter.

References

- Hanke, J. E.; Reitsch, A. G. (1995) *Business forecasting*, 5th ed., Englewood Cliffs et al.
- Makridakis, S. G.; Wheelwright, S. C.; Hyndman, R. J. (1997) *Forecasting: methods and applications*, 3rd ed., New York et al.
- Nahmias, S. (1997) *Production and operations analysis*, 3rd ed., Chicago et al.
- Silver, E. A.; Pyke, D. F.; Peterson, R. (1998) *Inventory management and production planning and scheduling*, 3rd ed., New York et al.
- Tempelmeier, H. (1999) *Material-Logistik: Modelle und Algorithmen für die Produktionsplanung und -steuerung und das Supply Chain Management*, 4th ed., Berlin et al.
- Wright, G.; Goodwin, P. (1998) *Forecasting with judgment*, New York et al.

8 Master Planning

Jens Rohde¹ and Michael Wagner²

¹ Darmstadt University of Technology, Department of Production & Supply Chain Management, Hochschulstraße 1, 64289 Darmstadt, Germany

² Paul Hartmann AG, Corporate Materials Management, P.O.Box 1420, 89504 Heidenheim, Germany

The main purpose of *Master Planning* is to synchronize the flow of materials along the entire supply chain. Master Planning supports mid-term decisions on the efficient utilization of production, transport, supply capacities, seasonal stock as well as on the balancing of supply and demand. As a result of this synchronization, production and distribution entities are able to reduce their inventory levels. Without centralized Master Planning, larger buffers are required in order to ensure a continuous flow of material. Coordinated master plans provide the ability to reduce these safety buffers by decreasing the variance of production and distribution quantities.

To synchronize the flow of materials effectively it is important to decide how available capacities of each entity of the supply chain will be used. As Master Planning covers mid-term decisions (see Chap. 4), it is necessary to consider at least one seasonal cycle to be able to balance all demand peaks. The decisions on production and transport quantities need to be addressed simultaneously while minimizing total costs for inventory, overtime, production and transportation.

The results of Master Planning are targets/instructions for Production Planning and Scheduling, Distribution and Transport Planning as well as Purchasing and Material Requirements Planning. For example, the Production Planning and Scheduling module has to consider the amount of planned stock at the end of each master planning period and the reserved capacity up to the planning horizon. The use of specific transportation lines and capacities are examples of instructions for Distribution and Transport Planning. Section 8.1 will illustrate the Master Planning decisions and results in detail.

However, it is not possible and not recommended to perform optimization on detailed data. Master Planning needs the aggregation of products and materials to product groups and material groups, respectively, and concentration on bottleneck resources. Not only can a reduction in data be achieved, but the uncertainty in mid-term data and the model's complexity can also be reduced. Model building including the aggregation and disaggregation processes is discussed in Sect. 8.2.

A master plan should be generated centrally and updated periodically. These tasks can be divided into several steps as described in Sect. 8.3.

8.1 The Decision Situation

Based on demand data from the Demand Planning module, Master Planning has to create an aggregated production and distribution plan for all supply chain entities. It is important to account for the available capacity and dependencies between the different production and distribution stages. Such a capacitated plan for the entire supply chain leads to a synchronized flow of materials without creating large buffers between these entities.

To make use of the Master Planning module it is necessary that production and transport quantities can be split and produced in different periods. Furthermore, intermediates and products should be stockable (at least for several periods) to be able to balance capacities by building up inventories. Because Master Planning is a deterministic planning module, reasonable results can only be expected for production processes having low output variances.

The following options have to be evaluated if bottlenecks on production resources occur:

- produce in earlier periods while increasing seasonal stock,
- produce at alternative sites with higher production and/or transport costs,
- produce in alternative production modes with higher production costs,
- buy products from a vendor with higher costs than your own manufacturing costs and
- work overtime to fulfil the given demand with increased production costs and possible additional fixed costs.

It is also possible that a bottleneck occurs on transportation lines. In this case the following alternatives have to be taken into consideration:

- produce and ship in earlier periods while increasing seasonal stock in a distribution centre,
- distribute products using alternative transportation modes with different capacities and costs and
- deliver to customers from another distribution centre.

In order to solve these problems optimally, one must consider the supply chain as a whole and generate a solution with a centralized view while considering all relevant costs and constraints. Otherwise, de-central approaches lead to bottlenecks at other locations and suboptimal solutions.

To generate feasible targets, a concept of anticipation is necessary. This concept should predict the (aggregate) outcomes of lower levels' decision-making procedures that result from given targets as precisely as possible within the context of Master Planning. This prediction should be less complex than performing complete planning runs at the lower planning levels. A simple example for anticipation is to reduce the periods' production capacity

by a fixed amount to consider setup times expected from lot wise production. However, in production processes with varying setup times dependent on the product mix, this concept may not be accurate enough; therefore, more appropriate solutions for anticipating setup decisions have to be found (for further information see Schneeweiss (2003); approaches for accurate anticipation of lot-sizing and scheduling decisions are given, for example, in Rohde (2004) and Stadtler (1988)).

The following paragraph introduces a small example to depict this decision situation. It will be used to illustrate the decisions of Master Planning and show the effects, results and the data used.

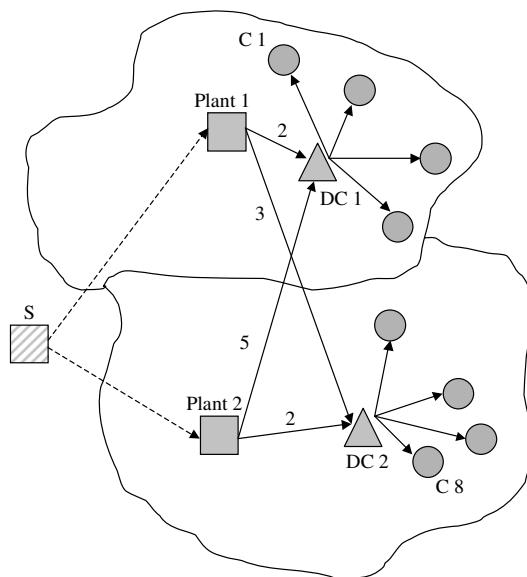


Fig. 8.1. Example of a supply chain

The example supply chain in Fig. 8.1 has two production sites (Plant 1 and Plant 2) as well as two distribution centres (DC 1 and DC 2). Two different products are produced in each plant in a single-stage process. The customers are supplied from their local *distribution centre* (DC), which usually receives products from the nearest production site. However, it is possible to receive products from the plant in the other region. Such a delivery will increase transport costs per unit. Inventory for finished products is exclusively located at the DCs. The regular production capacity of each production unit is 80 hours per week (two shifts, five days). It is possible to extend this capacity by working overtime.

In introducing a third production unit (supplier S), a multi-stage production problem and, in this example, a common capacity restriction for the

production of parts, has to be considered. In the remainder of this chapter the third production unit will not be regarded.

8.1.1 Planning Horizon and Periods

The planning horizon is characterized by the interval of time for which plans are generated. It is important to select a planning horizon that covers at least one seasonal cycle. Otherwise, there would be no possibility to balance capacities throughout a season, and hence, peaks in demand would possibly not be covered. If, for example, demand peaks were to occur in the last quarter of a year, and only half a year were considered, it might not be possible to balance this peak during planning of the second half (see following simplified example, Tables 8.1–8.3). Often, the planning horizon for Master Planning covers twelve months.

Table 8.1. Seasonal demand peak

Quarter	1	2	3	4
Demand	12	13	10	45
Capacity available	20	20	20	20

Table 8.1 shows the quarterly demand and the available capacity. Producing one part takes one capacity unit. If it is not possible to extend capacity, a plan with a horizon of two quarters would lead to the infeasible plan shown in Table 8.2. Considering a whole seasonal cycle (in this case four quarters), a feasible plan can be derived (see Table 8.3).

Table 8.2. Infeasible solution

Quarter	1	2	3	4
Demand	12	13	10	45
Capacity available	20	20	20	20
Capacity used	12	13	20	35

Table 8.3. Feasible solution

Quarter	1	2	3	4
Demand	12	13	10	45
Capacity available	20	20	20	20
Capacity used	20	20	20	20

As we have already seen in the previous example, the planning horizon is divided into several periods, the so-called *time buckets*. The length of these periods (often a week or month) must be chosen carefully with respect to the lead-times at every stage of the supply chain. In *bucket-oriented* Master Planning, the lead-time of each process¹ that uses potential bottleneck resources is usually defined as one time bucket or an integer multiple. A potential bottleneck resource might cause delays (waiting times) due to a high utilization rate. It is also possible that one might neglect the lead-time of some activities performed on non-bottleneck resources. Then, a part has to be produced in the same bucket as its predecessor and successor, respectively. This imprecision may lead to instructions that may not be disaggregated into feasible schedules at the lower planning levels, but then, shorter planned lead-times are possible. On the other hand, using one bucket or an integer multiple regularly leads to more appropriate instructions, but artificially extends the planned lead-times.

Shorter time buckets result in a more accurate representation of the decision situation and the lead-time modelling, but imply a higher complexity for the planning problem. Higher complexity, inaccuracy of data in future periods and the increasing expenditure for collecting data emphasize the trade-off between accuracy and complexity. Only the use of big time buckets allows for the planning of quantities, but not for individual orders or product units.

Another possibility is the use of varying lengths for different periods; that is, the first periods are represented by shorter time buckets to enable more exact planning on current data. The more one reaches the planning horizon, the bigger the chosen time buckets are. However, this approach poses problems with the modelling of lead-time offsets between production and distribution stages (see also Chap. 4).

To work on current data, it is necessary to update the master plan at discrete intervals of time. Thus, new and more reliable demand forecasts as well as known customer orders are considered in the new planning run. During the *frozen horizon* (see also Chap. 4), the master plan is implemented. Looking several periods ahead is necessary to be able to balance demand and capacities as already mentioned.

8.1.2 Decisions

Master Planning has to deal with the trade-off between costs for inventories, production, transports and capacity extension. The corresponding quantities that are produced, moved or stored need to be determined in the master planning process.

Production quantities (for each time bucket and product group) are mainly determined by the production costs and the available capacity. Capacity ex-

¹ Here, a process is an aggregation of several successive production and transportation activities (see also Sect. 8.2.3).

tensions have to be modelled as decision variables in Master Planning if production quantities also depend on these enhancements. Not only production capacity, but also transport capacity on the links between plants, warehouses and customers needs to be planned in Master Planning. Decisions on setups and changeovers are taken into account in Master Planning only if lot-sizes usually cover more than a period's demand. Otherwise, the decision is left to Production Planning and Scheduling, and setup times are anticipated in Master Planning.

While transport capacities only set a frame for the quantities that can be carried from A to B, the decision on the transport quantity (for every product group and time bucket) also needs to be addressed. Generally, linear transport costs are considered in mid-term planning. Hence, it is only possible to determine the quantities, not the detailed loading of single transportation means. This is to be done in Distribution and Transport Planning (see Chap. 12).

If production and transport quantities are determined, the stock levels are known. Inventory variables are used to account for inventory holding costs. Decision variables in the *example* are:

- production quantities for every product, period and plant,
- transport quantities on every transportation link from plant to DC, for every product and period,
- ending inventory level for every product, period and DC and
- overtime for every plant in every period.

8.1.3 Objectives

As described in the previous section, a model for Master Planning has to respect several restrictions when *minimizing total costs*. The costs affecting the objective function depend on the decision situation. In Master Planning they do not have to be as precise as, for example, within accounting systems; they are only incorporated to find out the most economical decision(s). A simple example may clarify this. If two products share a common bottleneck, it is only necessary to know which of the two products has the least inventory cost per capacity unit used. This will be the product to stock first, irrespective of “correct costs,” as long as the relation between the costs remains valid.

In most master planning settings, products can be stored at each production site and DC, respectively. Therefore, the inventory holding costs (e.g. for working capital, handling) have to be part of the objective function. Furthermore, the ability of extending capacity has to be taken into account. The corresponding costs need to be considered in the objective function. Also, variable production costs may differ between production sites, and thus, are part of the master planning process. If lot-sizing decisions should be made in Master Planning, it is necessary to incorporate costs for setups as well.

The different prices of the suppliers have to be considered in the objective function if Master Planning models are extended to optimize supply decisions.

Every stage of the production-distribution network is connected to other entities of the supply chain by transportation links, which are associated with transport costs. Usually, only variable linear cost rates for each transportation link and an adequate lead-time offset are considered in mid-term Master Planning.

The objective function of the *example* minimizes to the sum of

- production costs,
- inventory holding costs,
- additional costs for using overtime and
- transport costs.

8.1.4 Data

Master Planning receives data from different systems and modules. The forecast data, which describe the demand of each product (group) in each period in the planning horizon, are a result of Demand Planning.

Capacities need to be incorporated for each potential bottleneck resource (e.g. machines, warehouses, transportation). Transport capacities need not to be modelled if a company engages a third-party logistics provider who ensures an availability of 100%. But if capacity has to be extended on condition that cost rates increase, this additional amount of capacity and the respective cost rates have to be considered. For the calculation of necessary capacity, production efficiency and production coefficients have to be part of the model.

The BOMs of all products (groups) form the basis of the material flows within the model and provide the information on input-output coefficients. For every storage node (e.g. warehouse, work-in-process inventory) minimum (e.g. safety stocks and estimated lot-sizing stocks) and maximum stock levels (e.g. shelf lives) need to be defined for each product (group).

Additionally, all cost elements mentioned above are input to the model. Data for the *example* are

- forecasts for each sales region and product in every period,
- available regular capacity for each plant (machine) and period,
- maximum overtime in each plant,
- the production efficiency of products produced at specific plants (e.g. in tons of finished products per hour),
- the current stock levels at each DC and for every product and
- the minimum stock levels at each DC and for every product.

8.1.5 Results

The results of Master Planning are the optimized values of decision variables, which are instructions to other planning modules. Some decision variables have only planning character and are never (directly) implemented as they are determined in other modules in more detail (e.g. production quantities are planned in Production Planning and Scheduling).

Therefore, the most important results are the planned capacity usage (in each bucket for every resource (group) and transportation link) and the amount of seasonal stock at the end of each time bucket. Both cannot be determined in the short-term planning modules because they need to be calculated under the consideration of an entire seasonal cycle. Production capacities are input to Production Planning and Scheduling, and seasonal stock (possibly plus additional other stock targets), at the end of each Master Planning bucket, provide minimum stock levels in detailed scheduling.

Capacity extensions need to be decided during the frozen period as they often cannot be influenced in the short term. The same applies to procurement decisions for special materials with long lead-times or those that are purchased on the basis of a basic agreement.

Results in the *example* are

- seasonal stock, which is the difference between the minimum stock and the planned inventory level, for every product, period and DC and
- amount of overtime for every plant in every period that should be reserved.

8.2 Model Building

In most APS, Master Planning is described by a *Linear Programming* (LP) model with continuous variables. However, some constraints (including binary and integer variables, respectively) imply to convert the LP model to a more complex *Mixed Integer Programming* (MIP) model. Solution approaches for LP and MIP models are described in Chap. 27. In this section we will illustrate the steps of building a Master Planning model, and we will illustrate how complexity depends on the decisions modelled. Furthermore, it will be explained how complexity can be reduced by aggregation and how penalty costs should be used for finding (feasible) solutions.

Although it is not possible to give a comprehensive survey of all possible decisions, this chapter will show the dependence between complexity and most common decisions. In contrast to a perfect representation of reality, Master Planning needs a degree of standardization (i.e. constraints to be modelled, objectives, etc.), at least for a line of business. Thus, it is possible to use a Master Planning module that fits after adjusting parameters (i.e. costs, BOMs and routings, regular capacities, etc.), as opposed to after building new mathematical models and implementing new solvers.

8.2.1 Modelling Approach

Figure 8.2 shows a general approach for building a supply chain model that can be applied to most APS.

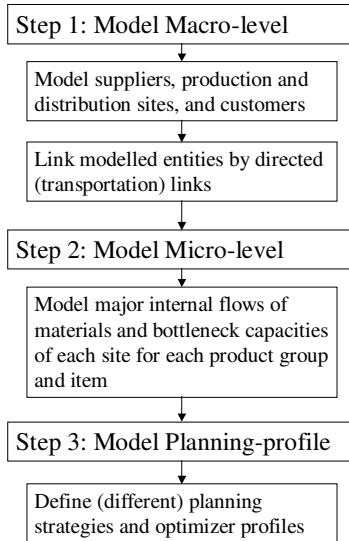


Fig. 8.2. Building a supply chain model

Step 1: Model Macro-level

In the first step, key-customers, key-suppliers, and production and distribution sites of the supply chain are modelled. These entities are connected by directed transportation links. In some APS, transportation links are modelled as entities and not as directed connections². The result of this step is a general network of supply chain entities.

In our example (Fig. 8.1) the two plants (Plant 1 and Plant 2) and the distribution centres (DC 1 and DC 2) are modelled; that is, their locations and possibly their types (e.g. production entity and distribution entity) are determined. Afterwards, the key-supplier (S) and the key-customers (C 1, ..., C 8) are specified. The supplier (S) does not represent a potential bottleneck. Hence, he should not explicitly be modelled in this step. The customers represent the demand of products of the supply chain. Finally, the transportation links are modelled. If no transportation constraints are applied, transportation links represent a simple lead-time offset between two stages.

² For example, i2 Technologies' Supply Chain Planner (see also Chaps. 18, 21 and 23).

Step 2: Model Micro-level

Each entity of the supply chain can be modelled in more detail in the second step, if required. All resource groups that could turn out to become a bottleneck should be modelled for each entity and transportation link. The internal flows of material and the capacities of potential bottlenecks are defined for each product group and item (group). The dependence between product and item groups is modelled by defining input and output materials for each process. Table 8.4 shows selected features that can be modelled in APS.

Table 8.4. Selected model features of Master Planning in APS

Process	Parameter	Characteristics
Procurement	purchase costs	linear piecewise linear
Production	production costs	linear
	production quantities	piecewise linear continuous
	capacity	semi-continuous regular capacity enhanced capacity with linear costs per extra unit
	capacity requirements	fixed or linear
Storage	inventory costs	linear
	capacity	regular capacity enhanced capacity with linear costs per extra unit
Distribution	transport costs	linear
	transport quantities	piecewise linear continuous
	capacity	integer partially integer regular capacity enhanced capacity with linear costs per extra unit
Sales	lateness	maximum lateness linear penalty costs
	not fulfilled	linear penalty costs

In our example, capacities and costs are modelled for each entity and transportation link. Plant 1 has a regular capacity of 80 units per time bucket. Each unit of overtime costs 5 monetary units (MU) without fixed costs, and producing one part of Product 1 or Product 2 costs 4 MU. Plant 2 is equally structured except that linear production costs amount to 5 MU for each unit produced. Then, the internal structure of each distribution centre is modelled.

The storage capacity of each distribution centre is limited. DC1 has linear storage costs for Product 1 of 3 MU per product unit per time bucket and for Product 2 of 2 MU per product unit per time bucket. DC2 has linear storage costs of 4 and 3 MU for one unit of Product 1 and 2, respectively, per time bucket. Finally, the transportation links are modelled on the micro-level. All transportation links are uncapacitated. Transport costs from Plant 1 to DC1 are linear with 2 MU per unit of both Products 1 and 2, and to DC2 costs of 3 MU per product unit are incurred. The transport from Plant 2 to DC1 and 2 costs 5 and 2 MU per product unit, respectively. Transportation from distribution centres to customers is not relevant.

Step 3: Model Planning-profile

The last step is to define a planning-profile. Defining the planning-profile includes the definition of resource calendars, planning strategies for heuristic approaches and profiles for optimizers. Planning strategies could include how a first feasible solution is generated and how improvements are obtained. Optimizer profiles could include different weights for parts of the objective function (inventory costs, transport costs, etc.). For example, an optimizer profile that forces production output could be chosen within a growing market. One way to force production output could be setting a lower weight on penalties for capacity enhancements and a higher weight on penalties for unfulfilled demand.

The example of this chapter should be solved by Linear Programming. The only objective is to minimize total costs resulting from inventories, production, overtime and transportation. A planning-profile would instruct the Master Planning module to use an LP-solver without special weights (or with equal weights) for the different parts of the objective function.

8.2.2 Model Complexity

Model complexity and optimization run time are (strongly) correlated. For this reason, it is important to know which decisions lead to which complexity of the model. Thus, it is possible to decide on the trade-off between accuracy and run time. The more accurate a model should be, the more the decisions are to be mapped. But this implies increased run time and expenditure for collecting data. The following paragraphs show the correlation between decisions described in this chapter and a model's complexity.

The main *quantity decisions* that have to be taken into consideration in a Master Planning model are production and transport quantities. For these quantities integer values are mostly negligible at this aggregation level. Mainly, they are used to reserve capacity on potential bottleneck resources. Because these are rough capacity bookings, it is justifiable to abstract from integer values. If different production or transportation modes can be used partially, additional quantity decisions for each mode, product and period are

necessary. Other important quantity decisions are stock levels. They result from corresponding production and transport quantities as well as stock levels of the previous period.

Capacity decisions occur only if it is possible not to utilize complete regular capacity or to enhance capacity of certain supply chain entities. One aspect of enhancing regular capacity is working overtime. This implies a new decision on the amount of overtime in each period for each resource. Additional costs have to be gathered. Binary decisions have to be made if extra shifts are introduced in certain periods (and for certain resources) to take fixed costs for a shift into account (e.g. personnel costs for a complete shift). Thereby, the problem is much harder. Performance adjustments of machines usually lead to non-linear optimization models. Computational efforts, and thus, solvability of such models, decrease sharply.

Decisions concerning production and transportation processes are, for example, decisions about the usage of alternative routings. Such additional decisions and more data will increase the model's complexity. However, if it is not possible to split production and transport quantities to different resources – e.g. supply customers from at least one distribution centre – process modes have to be considered. In contrast to the quantity decision on different modes as described above, additional binary decisions on the chosen process mode have to be made.

8.2.3 Aggregation and Disaggregation

Another way to reduce complexity of the model is *aggregation*. Aggregation is the reasonable grouping and consolidation of time, decision variables and data to achieve complexity reduction for the model and the amount of data (Stadtler, 1988, p. 80). The accuracy of data can be enhanced by less variance within an aggregated group, and higher planning levels are unburdened from detailed information.

Furthermore, inaccuracy increases in future periods. This inaccuracy, e.g. in case of the demand of product groups, can be balanced by reasonable aggregation if forecast errors of products within a group are not totally correlated. Therefore, capacity requirements for aggregated product groups (as a result of Master Planning) are more accurate, even for future periods.

Aggregation of time, decision variables and data will be depicted in the following text. Regularly, these alternatives are used simultaneously.

Aggregation of Time

Aggregation of time is the consolidation of several smaller periods to one large period. It is not reasonable to perform Master Planning, for example, in daily time buckets. Collecting data that are adequate enough for such small time buckets for one year in the future, which is mostly the planning horizon in mid-term planning, is nearly inoperable. Therefore, Master Planning is

regularly performed in weekly or monthly time buckets. If different intervals of time buckets are used on different planning levels, the disaggregation process raises the problem of giving targets for periods in the dependent planning level that do not correspond to the end of a time bucket in the upper level. To resolve this problem, varying planning horizons on lower planning levels can be chosen (see Fig. 8.3).

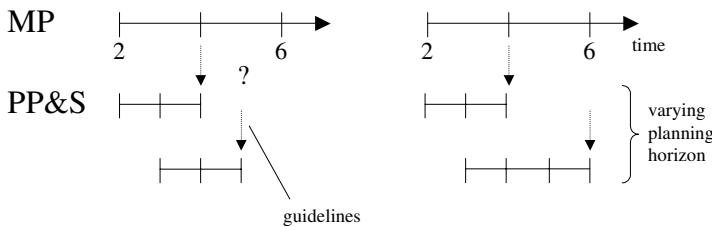


Fig. 8.3. Aggregation of time

Aggregation of Decision Variables

Generally, aggregation of decision variables refers to the consolidation of production quantities. In the case of Master Planning, transport quantities also have to be aggregated. Bitran et al. (1982) suggest aggregating products with similar production costs, inventory costs and seasonal demand to product types. Products with similar setup costs and *identical* BOMs are aggregated to product families. A main problem in Master Planning, which is not regarded by the authors, is the aggregation of products in a multi-stage production process with *non-identical* BOMs. The similarity of BOMs *and* transportation lines is very important. But the question of what similarity means remains unanswered. Figure 8.4 illustrates the problem of aggregating BOMs. Products P 1 and P 2 are aggregated to product type P 1/2 with the average quotas of demand of P 1/2 of $\frac{1}{4}$ for P 1 and $\frac{3}{4}$ for P 2. Parts A and B are aggregated to part type AB. The aggregated BOM for P 1/2 shows that one part of P 1/2 needs one part of type AB and $\frac{1}{4}$ part of part C (caused by the average quota for demand). Producing one part of type AB means producing one part of A and two parts of B. The problem is to determine a coefficient for the need for type AB in product P 3 (an aggregation procedure for a sequence of operations is discussed in Stadtler (1996) and Stadtler (1998)). Shapiro (1999) remarks with his 80/20-rule that in most practical cases about 20% of the products with the lowest revenues regularly make the main product variety. Thus, these products can be aggregated to fewer groups while those with high revenues should be aggregated very carefully and selectively.

It is important to perform an aggregation with respect to the decisions that have to be made. If setup costs are negligible for a certain supply chain,

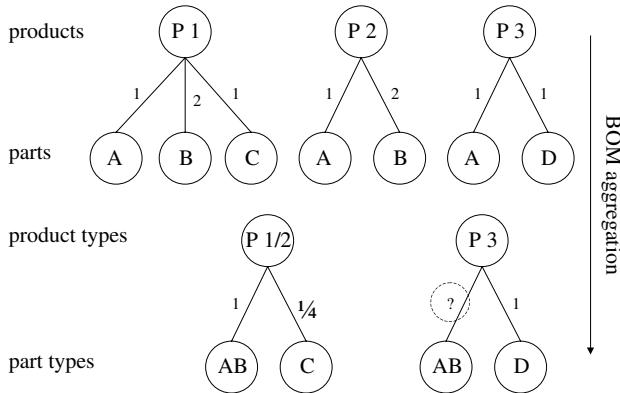


Fig. 8.4. BOM aggregation (following Stadtler (1988, p. 90))

it does not make sense to build product groups with respect to similar setup costs. No product characteristic, important for a Master Planning decision, should be lost within the aggregation process.

Aggregation of Data

The aggregation of data is the grouping of, for example, production capacities, transport capacities, inventory capacities, purchasing bounds and demand data. Demand data are derived from the Demand Planning module and have to be aggregated with respect to the aggregation of products. Particularly, aggregating resources to resource groups cannot be done without considering product aggregation. There should be as few interdependencies as possible between combinations of products and resources. Transport capacities, especially in Master Planning, have to be considered in addition to production and inventory capacities. Due to the various interdependencies between decision variables and data, these aggregations should be done simultaneously.

8.2.4 Relations to Short-Term Planning Modules

Master Planning interacts with all short-term planning modules by sending instructions and receiving reactions (see Fig. 8.5). Furthermore, master plans provide valuable input for collaboration modules and strategic planning tasks such as mid-term purchasing plans or average capacity utilization in different scenarios (see also Chaps. 13 and 14).

Instructions can be classified as primal and dual instructions (Stadtler, 1988, p. 129). The first type directly influences the decision space of the base-level model (here: the short-term planning modules) by providing constraints such as available capacity and target inventory at the end of a period. The

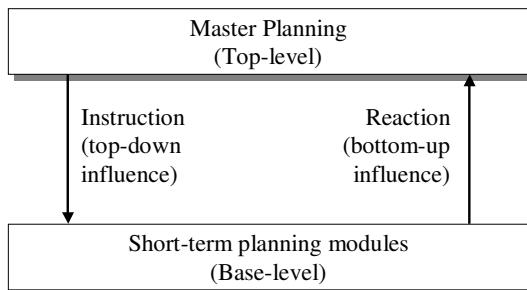


Fig. 8.5. Instruction and reaction in Master Planning (following Schneeweiss (2003, p. 17))

second one influences the objective function of the base-level model by setting cost parameters.

After a planning run for the short-term modules is performed, Master Planning is able to receive feedback/reactions from the base-level. Instructions that lead to infeasibilities have to be eliminated or weakened. By changing some selected Master Planning parameters, e. g. maximum capacity available, elimination of particular infeasibilities can be achieved.

To avoid a multitude of instruction/reaction loops, an anticipated model of the base-level should predict the outcome of the planning run of this level according to given instructions (see also Sect. 8.1).

Finally, ex-post feedback, gathered after executing the short-term plans, provides input, e. g. current inventory levels or durable changes in availability of capacity, to the Master Planning module.

As part of model building, the coupling parameters, i. e. instructions, reactions and ex-post feedback parameters, have to be defined (see also Chap. 4). Additionally, the type of the coupling relations (e. g. minimum/maximum requirements, equality) and the points of time in which to transfer the coupling parameters have to be assigned (Stadtler, 1988, pp. 129–138). To build the anticipated base-level model, the main influences of the short-term planning decisions within Master Planning have to be identified. For example, lot-sizes and setup-times resulting from Production Planning & Scheduling might be anticipated.

8.2.5 Using Penalty Costs

A model's solution is guided by the costs chosen within the objective function. By introducing certain costs that exceed the relevant costs for decisions (see Sect. 8.1.3), these decisions are penalized. Normally, relevant costs for decisions differ from costs used for accounting, e. g. only variable production costs are considered without depreciations on resources or apportionments of indirect costs. *Penalty costs* are used to represent constraints that are not explicitly modelled. Master Planning has to fulfil all demand requests in time.

To avoid infeasible plans, it can be necessary to penalize unfulfilled demand. Similarly, if setup times are not explicitly considered, the loss of time on a bottleneck resource can be penalized by costs correlated to this loss of time.

To be able to interpret the costs of the objective function correctly, it is important to separate the costs according to accounting and penalty costs. Regularly, penalty costs exceed other cost parameters by a very high amount. To obtain the “regular” costs of a master plan, not only the penalized costs of a solution, this separation is indispensable. Among others, the following penalties can be inserted in the objective function:

- setup costs to penalize the loss of time on bottleneck resources (if not explicitly modelled),
- costs for unfulfilled demand and late deliveries of finished products and parts,
- costs for enhancing capacity (especially overtime) to penalize its use explicitly,
- additional production costs for certain sites to penalize, for example, minor quality and
- penalty costs for excessive inventory of customer specific products.

8.3 Generating a Plan

This section illustrates which steps have to be performed to generate a master plan (see Fig. 8.6) and how to use Master Planning effectively.

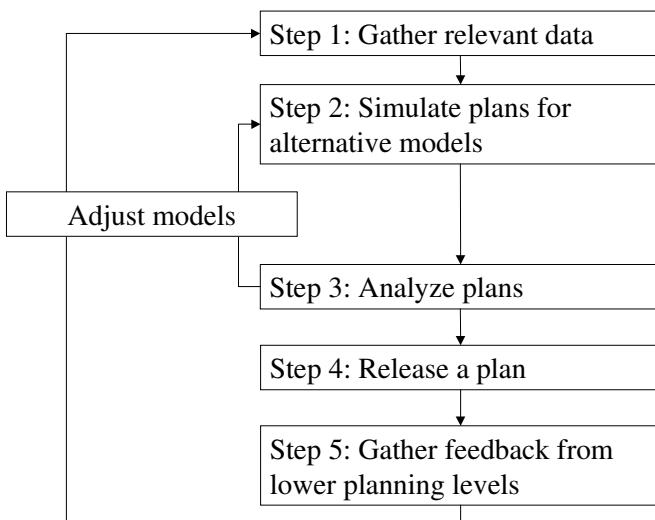


Fig. 8.6. Steps in Master Planning

As already mentioned, the master plan is updated successively, e.g. in weeks or months. Thus, new and accurate information such as actual stock levels and new demand data are taken into consideration. It is necessary to gather all relevant data before performing a new planning run (see Sect. 8.1). This can be a hard task as data are mostly kept in different systems throughout the entire supply chain. However, to obtain accurate plans this task has to be done very seriously. To minimize expenditure in gathering data, a high degree of automation to execute this process is recommended. For the previous example, the parameters described in Sect. 8.2.1 and the demand data shown in Table 8.5 have been gathered.

Table 8.5. Demand data

Period		1	2	3	4	5
Product 1	Sales region 1	20	30	30	50	40
Product 2	Sales region 1	50	40	60	40	50
Product 1	Sales region 2	40	30	30	40	40
Product 2	Sales region 2	50	40	50	60	50

Most APS provide the possibility to simulate alternatives. Several models can be built to verify, for example, different supply chain configurations or samples for shifts. Furthermore, this simulation can be used to reduce the number of decisions that have to be made. For example, *dual values* of decision variables (see Chap. 27) can be used after analyzing the plans to derive actions for enhancing regular capacity.

The master plan does not necessarily need to be the outcome of an LP or MIP solver. These outcomes are often unreproducible for human planners. Thus, insufficient acceptance of automatically generated plans can be observed. Following the ideas of the OPT philosophy (Goldratt, 1988), an alternative approach comprises four steps:

1. Generate an unconstrained supply chain model, disregarding all purchasing, production and distribution capacities.
2. Find the optimal solution of the model.
3. Analyse the solution regarding overloaded capacities. Stop, if no use of capacity exceeds upper (or lower) limits.
4. Select the essential resources of the supply chain of those exceeding capacity limits. Take actions to adjust the violated capacity constraints and insert those fixed capacities into the supply chain model. Proceed with step 2.

If each capacity violation is eliminated, the iterative solution matches the solution that is generated by an optimization of the constrained model. Due to the successive generation of the optimal solution, the acceptance of the decision makers increases by providing a better understanding of the system

and the model. In contrast to a constrained one-step optimization, alternative capacity adjustments can be discussed and included. It must be pointed out that such an iterative approach requires more time and staff than a constrained optimization, particularly if capacity usage increases. It seems to be advisable to include some known actions to eliminate well known capacity violations in the base supply chain model.

In the next step one has to decide which plan of the simulated alternatives should be released. If this is done manually, subjective estimations influence the decision. On the other hand, influences of not explicitly modelled knowledge (e.g. about important customers) are prevented by an automated decision.

Table 8.6. Production quantities

Period		1	2	3	4	5
Product 1	Plant 1	30	30	30	50	40
Product 2	Plant 1	50	50	50	40	50
Product 1	Plant 2	30	30	30	40	40
Product 2	Plant 2	50	40	50	60	50

Table 8.6 shows the planned production quantities of our example. The transport quantities correspond to the production quantities, except for transport of Product 1 from Plant 1 to the DCs in the first period. Twenty units of Product 1 produced in Plant 1 are delivered to DC1, while 10 are delivered to DC 2 to meet the demand of Sales region 2, even though transport costs are higher. Seasonal stock is only built in the second period for Product 2 in DC 1, amounting to 10 product units. Overtime is necessary for both plants in periods four and five. Plant 1 utilizes 10 time units of overtime each, while Plant 2 utilizes 20 and 10 time units, respectively. The costs for the five periods' planning horizon are shown in Table 8.7.

Table 8.7. Costs for five periods

Production	3,780.00
Overtime	250.00
Transportation	1,690.00
Inventory	20.00

Having forwarded the master plan's instructions to the de-central decision units, detailed plans are generated. The results of these plans have to be gathered to derive important hints for model adjustments. For example, if setups considered by a fixed estimate per period result in infeasible de-central plans, it is necessary to change this amount.

The mid-term purchasing quantities for raw materials from supplier S in our example (see Fig. 8.1) can be derived from the mid-term production plans for plants P1 and P2. These purchasing quantities are input for a collaborative procurement planning process (see Chap. 14). The joint plan of supplier S and plants P1 and P2 then serves as adjusted material constraint. Once an adequate master plan has been generated, decisions of the first period(s) are frozen and the process of rolling schedules is continued.

References

- Bitran, G. R.; Haas, E. A.; Hax, A. C. (1982) *Hierarchical production planning: A two-stage system*, in: Operations Research, Vol. 30, 232–251
- Goldratt, E. M. (1988) *Computerized shop floor scheduling*, International Journal of Production Research, Vol. 26, 433–455
- Rohde, J. (2004) *Hierarchical supply chain planning using artificial neural networks to anticipate base-level outcomes*, accepted for publication in the special issue on Soft Computing of OR Spectrum, 2004
- Schneeweiss, C. (2003) *Distributed decision making*, 2nd ed., Berlin et al.
- Shapiro, J. F. (1999) *Bottom-up vs. top-down approaches to supply chain modeling*, in: Tayur, S.; Ganeshan, R.; Magazine, M. (Eds.) Quantitative models for supply chain management, Dordrecht, The Netherlands, 737–760
- Stadtler, H. (1988) *Hierarchische Produktionsplanung bei losweiser Fertigung*, Heidelberg
- Stadtler, H. (1996) *Mixed integer programming model formulations for dynamic multi-item multi-level capacitated lotsizing*, in: European Journal of Operational Research, Vol. 94, 561–581
- Stadtler, H. (1998) *Hauptproduktionsprogrammplanung in einem kapazitätsorientierten PPS-System*, in: Wildemann, H. (Ed.) *Innovationen in der Produktionswirtschaft – Produkte, Prozesse, Planung und Steuerung*, München, 163–192

9 Demand Fulfilment and ATP

Christoph Kilger¹ and Lorenz Schneeweiss²

¹ j & m Management Consulting AG, Kaiserringforum, Willy-Brandt-Platz 5, 68161 Mannheim, Germany

² Nestlé GLOBE Center Europe GmbH, Rohnestrassse 4, 60523 Frankfurt/Main, Germany

The planning process that determines how the actual customer demand is fulfilled is called *demand fulfilment*. The demand fulfilment process determines the first promise date for customer orders and – thus – strongly influences the order lead-time and the on time delivery.¹ In today's competitive markets it is important to generate fast and reliable order promises in order to retain customers and increase market share. This holds particularly true in an e-business environment: Orders are entered online in the e-business front end, and the customer expects to receive a reliable due date within a short time period.

Further, e-business solutions have to support online inquiries where the customer requests a reliable due date without committing the order.

The fast generation of reliable order promises gets more complex as

- the number of products increases,
- products are configured during the ordering process,
- the average product life cycles get shorter,
- the number of customers increases,
- flexible pricing policies are being introduced and
- demand variations increase and get less predictable.

The traditional approach of order promising is to search for inventory and to quote orders against it; if there is no inventory available, orders are quoted against the production lead-time. This procedure may result in non-feasible quotes, because a quote against the supply lead-time may violate other constraints, e. g. available capacity or material supply.

Modern demand fulfilment solutions based on the planning capabilities of APS employ more sophisticated order promising procedures, in order

1. to improve the on time delivery by generating reliable quotes,
2. to reduce the number of missed business opportunities by searching more effectively for a feasible quote and
3. to increase revenue and profitability by increasing the average sales price.

¹ In the following, we use the terms *order promising* and *order quoting* synonymously, as well as the terms *promise* and *quote*.

In the following section, the principles of APS-based demand fulfilment solutions are described and the basic notion of ATP (available-to-promise) is introduced. Section 9.2 introduces the concept of allocation planning, resulting in allocated ATP (AATP). Section 9.3 illustrates the AATP-based order promising process by means of examples.

An early reference describing the concept of ATP and the improvement of the customer service level by ATP based on the master production schedule is Schwendinger (1979). In Fischer (2001) a comprehensive overview of ATP related work is given, and planning tasks related to ATP are classified. Most work so far about ATP concepts is described in the context of concrete APS like APO (Dickersbach, 2003) and i2 (i2 Technologies, 2000). Fleischmann and Meyr (2003) investigate the theoretical foundations of demand fulfillment and ATP, and discuss the generation of ATP and order promising strategies based on linear and mixed integer programming models.

9.1 Available-to-Promise (ATP)

9.1.1 The Role of Master Planning

The main target of the demand fulfilment process is to generate fast *and* reliable order promises to the customer and shield production and purchasing against infeasibility. The quality of the order promises is measured by the on time delivery KPI as introduced in Chap. 2. Using the traditional approach – quoting orders against inventory and supply lead-time – often will result in order promises that are not feasible, decreasing the on time delivery.

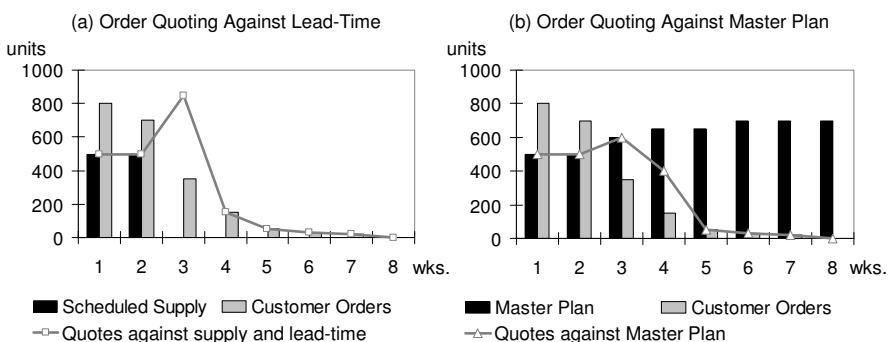


Fig. 9.1. Demand wave beyond the standard lead-time

Figure 9.1 illustrates this by means of a simple example. Consider a material constrained industry like the high-tech industry, and let us assume that a specific component has a standard lead-time of two weeks. There are receipts from the suppliers scheduled for the next two weeks and – as we assume

a material constrained industry – no additional supply will be available for the next two weeks. The volume of new orders that need to be promised for week 1 and week 2 is exceeding the volume of the scheduled receipts. Figure 9.1 (a) illustrates this situation. Standard MRP logic is to schedule all new orders against the scheduled receipts and – if not all orders can be satisfied by that – against the standard lead-time (in our example two weeks). In other words, MRP assumes infinite supply beyond the standard lead-time and creates supply recommendations based on the order backlog. The grey line in Fig. 9.1 (a) shows the quotes created by the MRP logic. In week 3 (i. e. after the standard lead-time) all orders are scheduled that cannot be quoted against the scheduled receipts. It is quite clear that the fulfilment of this “demand wave” will not be feasible, as the available supply will most probably not increase by 100% from one week to the next week.

The master planning process (see Chap. 8) has the task to create a plan for the complete supply chain, including purchasing decisions. Thus, master planning creates a plan for future supply from the suppliers even beyond the already existing scheduled receipts.² The idea of APS-based demand fulfilment is to use this information to create reliable order quotes. Figure 9.1 (b) shows the master plan and the orders quoted against the master plan. For week 3 the master plan reflects the constraints of the suppliers, anticipating a slight increase of the supply volume that is considered to be feasible. As orders are quoted against the master plan the unrealistic assumption of infinite supply beyond the standard lead-time is obsolete – resulting in more reliable order promises.

In most APS – and also ERP systems – the master plan quantities that form the basis for order promising are called *available-to-promise (ATP)*. ATP is the result of a synchronized supply and capacity plan and represents actual and future availability of supply and capacity that can be used to accept new customer orders on.

Figure 9.2 summarizes the role of master planning. The master planning process is based on the demand plan that reflects the capability of the market to create demand. During the master planning process all material, capacity and lead-time constraints of the supply chain are applied to the demand plan, resulting in a feasible master plan. This plan is the common basis for the supply processes (supply recommendations for purchasing, production and distribution) and the order promising process (available-to-promise quantities). By that, supply is synchronized with order promising, resulting in reliable order quotes. As a consequence the on time delivery is improved.

Please note that the on time delivery KPI is mainly influenced by the quality of the master planning model to reflect the reality in a sufficiently accurate way. ATP based on an accurate master planning model guarantees almost 100% on time delivery which is only influenced by supply deviations on the supply side and unexpected capacity problems, e.g. in production, on

² For week 1 and 2 the master plan reflects the scheduled receipts.

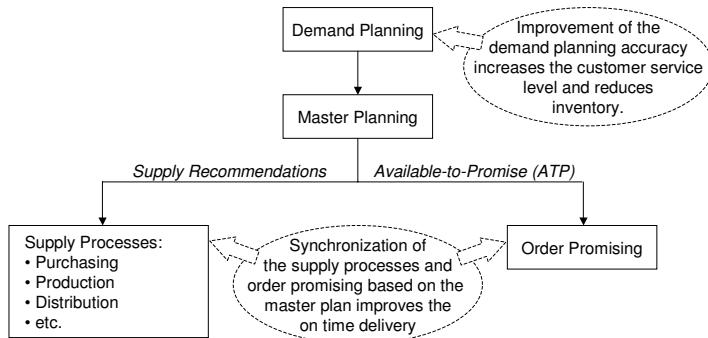


Fig. 9.2. Master Planning as the common basis of supply chain processes and order promising

the capacity side. Apart from on time delivery the delivery performance KPI plays an important role in demand fulfilment as it reflects how fast the supply chain is able to fulfil a customer order. Delivery performance in contrast to on time delivery depends mainly on the forecast accuracy of the demand plan and the ability of the supply chain to satisfy the demand signal of the demand plan. The master planning process is responsible to create a feasible supply plan based on the demand plan. If the demand plan does not mirror future orders very well the probability is low that there is ATP available when a new customer order requests for it. In this case, the customer order receives a late, but reliable promise and the delivery performance is affected. If new customer orders come in as planned by the demand plan and the master planning process was able to generate a feasible supply plan from the demand plan, then supply matches the demand, the number of inventory turns increases and orders receive a good promise within a short lead-time.

ATP can be structured along three dimensions: the product dimension, the time dimension, and the customer or market dimension. The granularity of ATP along the product dimension is described in Sect. 9.1.2, the granularity of ATP along the time dimension is described in Sect. 9.1.3. Structuring ATP by customer or market is called *allocation of ATP* and is outlined in Sect. 9.2.

9.1.2 Granularity of ATP along the Product Dimension

Principally ATP can be represented on any stage of the supply chain, e. g. finished goods, components, or raw materials. The decision where to represent ATP best for a certain business is strongly linked with the location of the decoupling point in that particular supply chain (Fleischmann and Meyr, 2003). The decoupling point separates the forecast-driven parts of a supply chain from the order-driven parts. Typically, a safety stock is held at the decoupling point to account for forecast errors. Fig. 9.3 shows the location of the

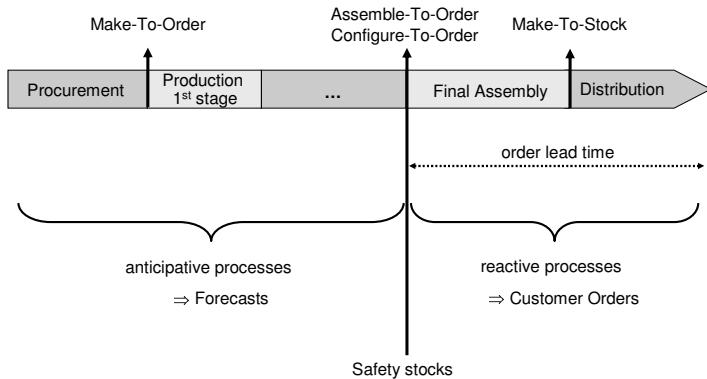


Fig. 9.3. Decoupling point in make-to-stock, assemble/configure-to-order, and make-to-order business environments (adapted from Fleischmann and Meyr (2003))

decoupling point for make-to-stock, assemble/configure-to-order, and make-to-order business environments. (Refer also to Chap. 3 for an introduction into the concept of decoupling point.)

The decision on which level ATP is being represented – i. e. where the decoupling point is defined – is also influenced by the forecast accuracy that can be achieved on that level of the supply chain and on the targeted service level.

- The forecast accuracy at the decoupling point must have a sufficient quality to procure, produce and distribute products based on this forecast. The more the decoupling point is moved to the left (upstream), the easier it is to forecast the demand on that level. Forecast accuracy on raw material level is usually higher than on finished goods level.
- The service level to the customer – represented by the order lead time and the ability of the supply chain to offer a good promise for customer orders – plays an important role in the competition with other supply chains. The service level can be increased by moving the decoupling point to the right (downstream), as the time needed to ship a product gets shorter.

Make-to-Stock

In a make-to-stock environment (see Chap. 3) the standard way to represent ATP is on finished goods level (e. g. actual end products, articles, etc. that are to be sold or aggregated product groups). The complete production process in a make-to-stock business is forecast-driven. Further, parts of the distribution processes can be forecast-driven (for example if products are transported to regional distribution centers, refer to Chap. 12). From there customer orders are served with a short lead time. The promise would be given under

consideration of availability of finished goods ATP and transportation times. Examples for make-to-stock industries are consumer packaged goods, food and beverages, and retail. In some make-to-stock industries the decoupling point even moves with a seasonal pattern in the distribution network.³

Assemble/Configure-to-Order

In an assemble-to-order environment, all components are produced and/or procured driven by the forecast. Only final assembly is order-driven (see Fig. 9.3). Usually, there are some (or many) configuration options the customer can choose from (e.g. color, technical options, country specific options like power plug), and the actual configuration is determined only at order entry time. This is called *configure-to-order*. In an assemble/configure-to-order environment the forecast is created on finished products or product group level; the forecast is then transformed by master planning into a supply plan on component level. For this, the bill-of-materials of the finished products or specific *planning bill-of-materials* for the product groups are exploded and lead-times and capacity usage are considered. ATP is represented on component level based on the master plan. Upon customer order entry, the bill-of-material of the (configured) product is exploded, component request dates are derived from the customer requested date, and component availability is checked for all ATP-relevant components. The latest availability date of all ATP-relevant components determines the quote for the complete order; all ATP consumptions are then synchronized according to the final quote, and lead times for assembly and transportation are added. This scheme is also called *multi-level ATP* (Dickersbach, 2003).

For configurable products there exists no deterministic bill-of-material representation for final products or product groups. Thus, the distribution of demand for the configuration options must be planned explicitly. E.g. consider a color option with three possible values “red”, “blue” and “green”, the demand for the three options may be distributed as follows: “red” 60 %, “blue” 15 % and “green” 25 %. Based on this distribution of the demand for the configuration options and the forecast on product group level, master planning provides a supply plan on component level, that is then represented as ATP. Consider the computer industry as an example (see Chap. 21 for further details). From a limited number of components – e.g. disk drives, processors, controllers, memory – a huge number of configurations can be made. An order consumes ATP from the base configuration of the computer (motherboard, housing, power supply, key board, etc.), and from all compo-

³ In the tyre industry the decoupling point is usually located at the central DC. At the start of the winter tyre business (in Western Europe usually in September/October), the demand for winter tyres is at peak and exceeds the handling capacity of the central DC. Therefore the decoupling point for winter tyre business is moved from the central DC to the regional DCs for that time frame.

nents that were configured by the customer, e. g. speed of processor, size of disk drive and memory.

Make-to-Order

Make-to-order environments are similar to assemble-to-order, but the decoupling point is located more upstream. In a make-to-order environment procurement is driven by forecast, and production, final assembly and distribution is driven by customer orders (see Fig. 9.3). Finished products and components are either customer-specific or there are so many different variants that their demand cannot be forecasted with a high accuracy. Besides material availability, the required capacity is typically an important constraint for the fulfillment of customer orders. Thus, ATP in a make-to-order environment is representing (a) the availability of raw material (see description of multi-level ATP above) and (b) the availability of capacity. For this purpose, specific ATP sources are formed representing the capacity of a specific kind that is available for promising customer orders. The capacity ATP is either represented in the demand fulfillment module of the APS on an aggregated level (resource groups), or the production planning and scheduling module is used to generate a promise. In the first case, capacity is treated like a component, and the availability of that “capacity component” is checked as described above for assemble-to-order and configure-to-order environments. In the second case, the customer order is forwarded to the production planning and scheduling module of the APS, is inserted into the current production plan, and the completion date of the order is returned to the customer as promised date. This concept is also called *capable-to-promise* (CTP) (Dickersbach, 2003).⁴

With capable-to-promise, the production process is simulated for the new customer order. This simulation may involve all subordinate production levels (multi-level production). Both, material and capacity availability are checked, resulting in a highly accurate order promise. A further advantage of CTP is that planned production orders are created upon order promising directly, and have only to be changed later if orders have to be replanned (due to material or capacity shortages or additional demand with higher priority). In complex production environments with many levels in the bill-of-materials and complex capacity constraints including setup constraints, CTP does not lead to an optimal production plan and schedule. The reason is the order-by-order planning scheme applied by CTP, often leading to poor schedules.

Tab. 9.1 summarizes the ATP granularity of the three manufacturing environments. Please note that the consumption of ATP does not mean that a certain supply represented by ATP is reserved for a certain customer order. ATP is a concept that allows a customer order to enter the planning sphere

⁴ Note that capable-to-promise can also be applied to assemble-to-order environments.

Table 9.1. ATP granularity of different manufacturing environments

Manufacturing environm.	Order lead-time	ATP granularity
make-to-stock	transportation time	product groups, finished goods
assemble-to-order, configure-to-order	assembly lead-time, transportation time	assembly capacity, components
make-to-order	production lead-time, transportation time	production capacity, components

of a supply chain to a certain date (promise date) so it can be delivered on time. The detailed material and capacity assignment for a customer order is only done in detailed scheduling and execution to keep the flexibility for optimization.

9.1.3 Granularity of ATP Along the Time Dimension

ATP is maintained in discrete time buckets. These correspond to the time buckets of the master planning process which might be different (usually more granular) from the forecasting buckets. Typically a forecast is given in weekly or monthly buckets whereas the supply is planned in daily or weekly patterns. Therefore ATP is usually represented in the same granularity as the supply given by the master plan. Orders are quoted by consuming ATP from a particular time bucket. Today, the granularity of the master plan is usually a compromise between the needed level of detail to offer accurate promises and the performance of an APS. The higher the level of detail the more exact a master plan has to be calculated and the more time buckets have to be searched for ATP to generate a promise. Therefore the time axis within the allocation planning horizon is typically more detailed in the near term and less detailed in the mid to long term.

9.2 Allocated ATP

9.2.1 Motivation

A supply chain (or a part of a supply chain) operates either in the supply constrained mode or in the demand constrained mode. If material and/or capacity are bottlenecks, then there is “open” demand that cannot be fulfilled. The supply chain supplies less finished goods than the customers request and operates in the *supply constrained mode*. If demand is the bottleneck, then all demand can be matched by supply. The supply chain operates in the *demand constrained mode*.

In the demand constrained mode the supply chain is able to generate excess supply that is not requested and will – most probably – not be consumed

by the customers. The capability of a supply chain to produce excess supply is an indicator for inefficiencies in the supply chain (refer to Chap. 15). A supply chain is working more profitable if it is “operated on the edge” (Sharma, 1997) by removing all inefficiencies, e.g. excess capacity, excess assets and excess expenses. Thus, even in the demand constrained mode, supply should be limited in order to make the supply chain more profitable. As a consequence the supply chain moves towards the supply constrained mode. To summarize, a supply chain is either in the supply constrained mode or should move towards the edge of this mode.

In the supply constrained mode, if orders are promised on a *first-come-first-served policy*, all orders are treated the same without taking the profitability of the order, the importance of the customer and the fact whether the order was forecasted or not, into account. As a consequence the profitability of the business, the relationship to the customers and the performance of the supply chain may be jeopardized.

A good example of how business can be optimized by using more sophisticated order promising policies is given by the international airlines (Smith et al., 1992). Airlines keep a specific fraction of the business and the first class seats open even if more economy customers are requesting seats than the total number of economy seats. For each flight, some of the business class and first class seats are *allocated* to the business and first class passengers based on the forecasted passenger numbers for that flight. Only a short time before the flight departs the allocations are released and passengers are “upgraded” to the next higher class. By that, airlines achieve a higher average sales price for the available seats and strengthen the relationship to their important customers, the business class and first class passengers.

APS apply the same principle by allocating ATP quantities to customer groups or sales channels in order to optimize the overall business performance. A classification scheme is defined that is used to classify customer orders. Typically, the order classes are structured in a hierarchy. The ATP quantities are allocated to the order classes according to predefined business rules. These allocations represent the right to consume ATP. Fig. 9.4 shows the principal connection between allocation and ATP. When an order is entered, the order promising process checks the allocations for the corresponding order class. If allocated ATP is available, ATP can be consumed and the order is quoted accordingly. Otherwise, the order promising process searches for other options to satisfy the order, e.g. by checking ATP in earlier time buckets, by consuming ATP from other order classes (if that is allowed by the business rules defined) or by looking for ATP on alternate products.

As indicated in Fig. 9.4 the time buckets for the allocations and the actual ATP quantities may differ. Allocations must be carefully controlled and regularly adjusted by human planners (as described for the “airline” example above). Otherwise the order lead time for some order classes will deteriorate while the ATP buckets for other order classes remain full as they are not

consumed as anticipated. Thus, it is helpful to provide allocations in a larger granularity, e.g. weeks or months than the actual ATP quantities, in order to support the manual control and adjustment processes.

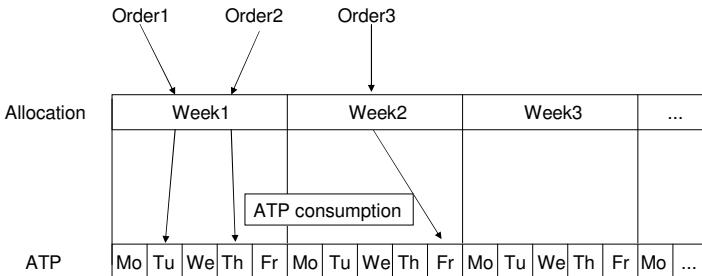


Fig. 9.4. The allocation controls the access to the ATP

The allocation of ATP to order classes can be exploited to increase the revenue and profitability of the business. For example, the average selling price may be increased by allocating supply to customers that are willing to pay premium prices, instead of giving supply away to any customer on a first-come-first-served basis. Traditional ATP mechanisms without allocation rules have to break commitments that have been given to other customers in order to be able to quote an order of a key customer or an order with a higher margin. It is obvious that this business policy has a negative impact on the on time delivery and deteriorates the relationship to other customers.

9.2.2 The Customer Hierarchy

In order to allocate supply to customers a *model of the customer structure* and a *forecast of the future customer demand* is required. The model of the customer structure should be aligned with the geographic dimension in demand planning (see Chap. 7), as demand planning is structuring the forecast in terms of the geographic dimension. Hence, the customer structure forms a hierarchy similar to the geographic dimension in demand planning. Figure 9.5 shows an example of a customer hierarchy.

In the first step the forecast quantities for each customer (or customer group, resp.) are aggregated to the root of the hierarchy. This number gives the total forecast for that specific product (or product group). The total forecast is transferred to master planning, and master planning checks whether it is feasible to fulfil the total forecast considering the supply constraints. In our example, the total forecast is 1400, and we assume that master planning can confirm only 1200 to be feasible.

In the second step the total feasible quantity according to the master plan is allocated from the top down to the leaves of the customer hierarchy. This allocation process for our example is visualized in Fig. 9.6 (the quantities

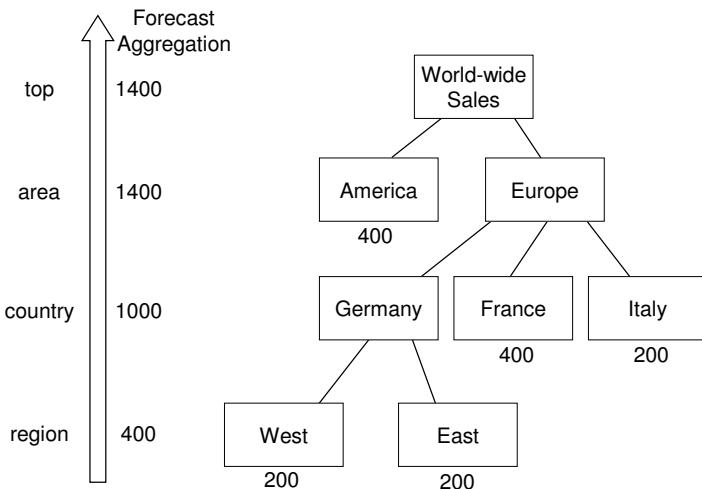


Fig. 9.5. Sales forecast aggregated along the customer hierarchy

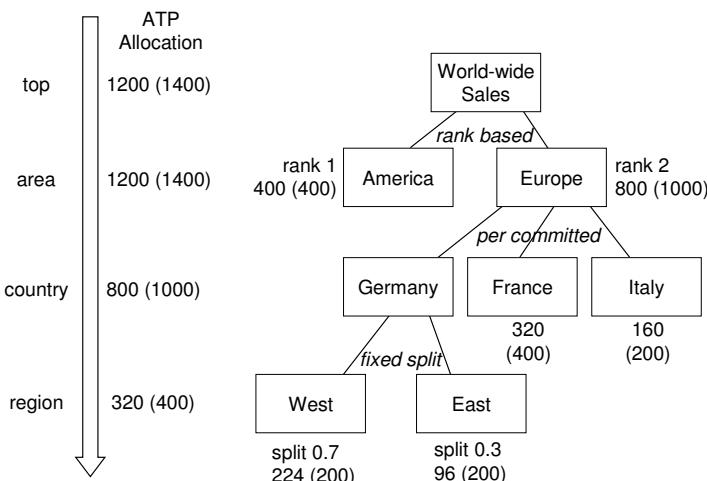


Fig. 9.6. Allocation of ATP in the customer hierarchy

in parentheses indicate the original forecast for this customer group). The allocation of the master plan quantities to the nodes of the customer hierarchy is controlled by allocation rules. In our example we have used three different allocation rules:

- *Rank based:* U.S. customers receive a higher priority (rank 1) compared to customers in Europe (rank 2). Thus, the available quantity for the U.S. and European customers is allocated to the U.S. first up to the original forecast for that area. A rank-based allocation policy may be helpful to

support sales to a specific market, e. g. if the development of that market is in an early stage.

- *Per committed:* The available quantity is allocated to the nodes of the customer hierarchy according to the forecast the customers have committed to. In our example Germany and France have forecasted 400 each, and Italy has forecasted 200, making 1000 in total. However, for this group of customers, only 800 is available. The quantity of 800 is split according to the fraction of the original forecast, i. e. Germany and France receive 40% each (320), and Italy receives 20% (160). The per committed allocation policy is well suited if each customer group shall get a fair share allocation according to what has been forecasted by that customer group.
- *Fixed split:* The fixed split allocation policy applies predefined split factors to distribute the feasible quantity to the customer groups. In our example, the customers in the Western part of Germany receive 70% of the available quantity, the customers in East Germany 30%. Please note that the resulting quantities are independent of the individual forecast of the customer groups. (But it does dependent on the total forecast of these customer groups.)

In addition to these allocation rules a portion of the available quantity can be retained at every level of the customer hierarchy. These retained quantities are consumed based on a first-come-first-served policy. Retained ATP can be used to account for potential variations of the actual demand related to the forecasted demand. For example, if 25% of the total quantity available for European customers is retained at the customer group Europe, 200 would be available on a first-come-first-served basis for all European customers, and only 600 would be allocated to German, French and Italian customers as defined by the corresponding allocation rules.

The allocations are the basis for generating order quotes. Thus, the allocations are an important information for the sales force before making commitments to their customers. Further, the APS keeps track of the consumptions due to already quoted orders. The total allocated quantities and the already consumed quantities give a good indication whether the order volume matches the forecast. If orders and forecast do not match, some allocations are being consumed too fast, whereas others remain unconsumed. This can be given as an early warning to the supply chain that the market behaves differently than forecasted – and appropriate action can be taken. For example, sales can setup a sales push initiative to generate additional demand to consume the planned ATP, and the supply operations can adjust production plans accordingly.

9.2.3 Allocation Planning

The process that assigns the overall ATP quantities received from master planning to the nodes of the customer hierarchy is called *allocation planning*.

Allocation planning is executed directly after a new master plan is created – which normally takes place once a week. Thus, once a week the adjusted forecast is transformed into ATP by master planning and allocated to the customer hierarchy.

In addition to that, the allocations are updated on a daily basis in order to reflect changes in the constraints of the supply chain. For example if the supplier of some key component announces a delay of a scheduled delivery this may impact the capability of the supply chain to fulfil orders and – because of that – should be reflected in the ATP as soon as the information is available in the APS.

The planning horizon of allocation planning cannot be longer than the planning horizon of master planning, as no ATP is available beyond the master planning horizon. The master planning process covers usually six to twelve months. However, in many cases it is not necessary to maintain allocations over six months or more. For example, in the computer industry, 90% of the orders are placed three weeks prior to the customer requested delivery date. Thus, dependent on the lead-time from order entry date to the customer requested delivery date a shorter planning horizon for allocation planning can be chosen compared to master planning. In the computer industry for example, a three-months horizon for allocation planning is sufficient.

9.3 Order Promising

Order promising is the core of the demand fulfilment process. The goal is to create reliable promises for the customer orders in a short time. The quality of the order promising process is measured by the on time delivery and the delivery performance.

The *on time delivery* KPI is described in detail in Chap. 2; it measures the percentage of the orders that are fulfilled as promised (based on the first promise given). Thus, to achieve a high on time delivery it is important to generate reliable promises. A promise is *reliable* if the customer can trust the ability of the supply chain to fulfil the order as promised, i.e. if the customer receives the promised product in the promised quantity at the promised date. A supply chain that is able to consistently generate reliable promises over a long time period gets a competitive advantage over supply chains with a lower on time delivery.

Besides the on time delivery, the actual order promising lead-time is an important aspect of the customer service level. In recent times a large variety of order entry paths have been created – and more are evolving. For example, in the computer industry, customers can order products at an authorized dealer, at a reseller, in department stores and consumer markets or directly at the manufacturer by telephone or via the Internet. By that, the probability that peak situations occur where large number of orders are entered at the same time, increases. A typical value in the computer industry is 1000 orders

per hour in peak load situations. Thus, the order promising process must be able to generate (reliable) quotes for individual orders in a very short time, e.g. down to milliseconds. Otherwise, customers wait for their promises and – while waiting – may change their mind and order online over the Internet or by telephone at a competitor.

On the one hand the online order promising process offers advantages in responsiveness and performance. On the other hand it is not possible that promises can be reviewed by order management before they reach the customer. Especially in allocation situations it might be possible to react on customer orders by trading allocations between customer structures and thereby generating better promises in order not to lose a customer. In other industries it might be sufficient to issue promises for customer orders only once or twice a day. Therefore the promises can be balanced in the allocation situation before they are issued to the customer.

9.3.1 ATP Search Procedure

The general ATP-based order promising process works as follows: First, the order promising process searches for ATP according to a set of search rules that are described below. If ATP is found, it is reduced accordingly and a quote for the order is generated. If ATP can only be found for a portion of the ordered quantity and partial fulfillment of the order is allowed, a quote for the partial order quantity is generated. If no ATP can be found, no quote is generated, and the order must be either rejected or confirmed manually at the end of the allocation planning horizon. Note that if no ATP can be found for an order, the supply chain will not be able to fulfil the order within the allocation planning horizon.

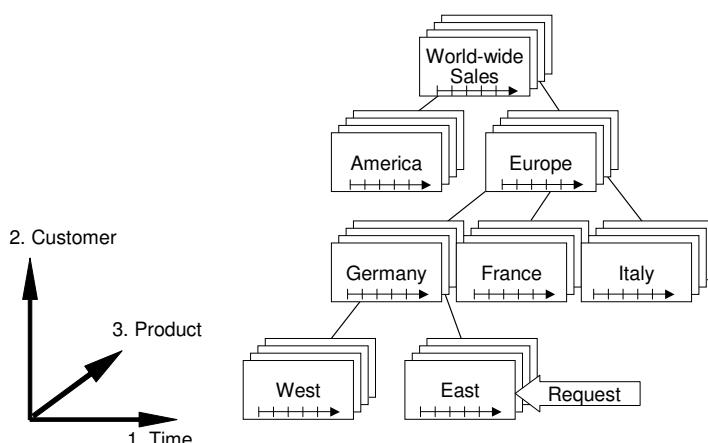


Fig. 9.7. Three dimensions of ATP search paths

ATP is searched along three dimensions: the time dimension, the customer dimension and the product dimension. Figure 9.7 illustrates the three dimensions of the ATP search paths. The following *search rules* are applied (for simplicity we assume that the ATP is on finished goods level; the search rules are similar for ATP on product group level and component level):

1. The leaf node in the customer hierarchy, to which the customer belongs, the product being requested by the order and the time bucket containing the customer requested date are determined. The ATP at this point is consumed – if available.
2. If ATP is not sufficient, then the time dimension is searched back in time for additional ATP (still at the leaf node in the customer hierarchy and at the product requested by the order); all ATP found up to a predefined number of time buckets back in time is consumed. Note that if ATP is consumed from time buckets earlier than the time bucket containing the customer requested date, the order is pre-built, and inventory is created.
3. If ATP is still not sufficient, steps 1. and 2. are repeated for the next higher node (parent node) in the customer hierarchy (searching for retained ATP quantities), then for the next higher and so on up to the root of the customer hierarchy.
4. If ATP is still not sufficient, steps 1. to 3. are repeated for all alternate products that may substitute the original product requested by the order.
5. If ATP is still not sufficient, steps 1. to 4. are repeated, but instead of searching backward in time, ATP is searched forward in time, up to a predefined number of time buckets. Note that by searching ATP forward in time, the order will be promised late.

The set of search rules described above is only one example of an ATP search strategy. In fact, an ATP search strategy may consist of any meaningful combination and sequence of the following search rule types:

- *Search for Product Availability:* This is the standard ATP search for a product including future receipts and constraints.
- *Search for Allocated ATP:* ATP is searched for along the customer dimension.
- *Search for Forecasted Quantities:* The creation of a quote for an order is based on forecasted quantities. The forecasted quantities in general are not customer specific.
- *Search for Component Availability:* For complex production processes and bill-of-material structures a multi-level ATP search for component availability is performed.
- *Capable-to-Promise:* ATP is dynamically generated by invoking the production planning and scheduling module.
- *Perform Substitution:* If no ATP can be found for a given product in a given location this type of rule allows to search for (a) the same product in another location, or (b) another product in the same location, or (c)

another product in other locations. This so-called *rule-based ATP* search requires the maintenance of lists of alternate products and/or locations and a rule to define the sequence in which the product and/or location substitutions are to take place.

Fleischmann and Meyr (2003) describe linear and mixed integer programming models for order promising in make-to-order, assemble-to-order and make-to-stock supply chains. In the following, we illustrate the ATP search procedure by means of a simple example.

9.3.2 ATP Consumption by Example

Let us assume an order is received for 300 units from a customer in East Germany, with a customer requested date in week 4. The ATP situation for East Germany is depicted in Fig. 9.8. First, the ATP is checked for the customer group East Germany for week 4, then for week 3 and for week 2. (We assume that the ATP search procedure is allowed to consume ATP 2 weeks back in time.) The ATP that is found along that search path is 10 in week 4, 60 in week 3 and 50 in week 2, 120 in total (see Fig. 9.8).

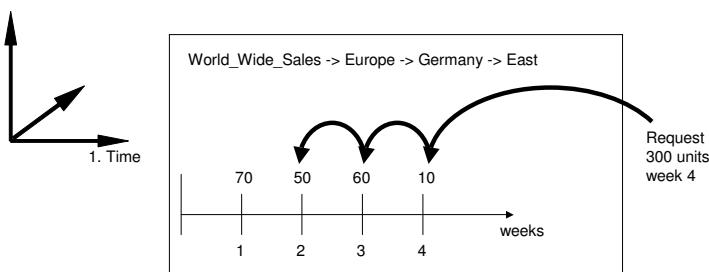


Fig. 9.8. Consumption of ATP along the time dimension

As the ATP search procedure may not consume ATP from a time bucket that is more than two weeks prior to the customer requested date, 180 units of the requested quantity is still open after the first step. In the second step, ATP is searched along the customer dimension. We assume for this example that there is ATP in the next higher node in the customer hierarchy that is Germany as shown in Fig. 9.9, but no ATP in the next higher nodes, i.e. Europe and World-wide Sales. From the ATP allocated at Germany, another 120 units can be consumed in weeks 4, 3 and 2, resulting in a total promised quantity of 240. 60 units are still open, as the requested quantity is 300 units.

In the next step, the ATP search algorithm looks for alternate products as shown in Fig. 9.10. The alternates are sorted by priority. First, the alternate with the highest priority is considered, and the same steps are applied as for the original product, i.e. first search back in time and second search up

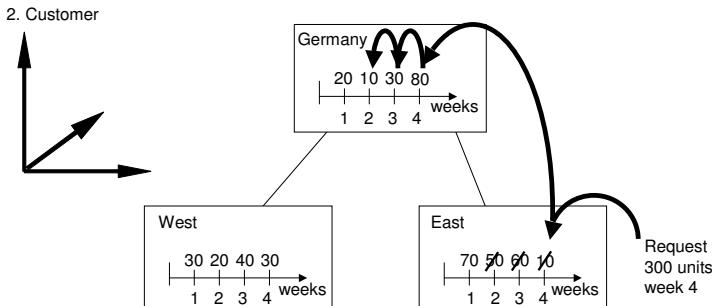


Fig. 9.9. Consumption of ATP along the customer dimension

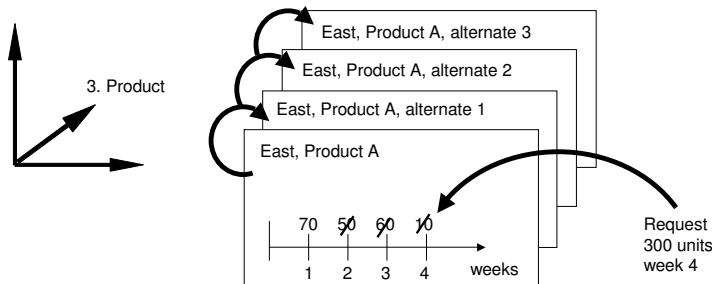


Fig. 9.10. Consumption of ATP along the product dimension

the customer hierarchy. Then, these steps are applied to the alternate with second highest priority and so on.

References

- Dickersbach, J. (2003) *Supply Chain Management with APO*, Heidelberg
- Fischer, M. (2001) *Available-to-Promise: Aufgaben und Verfahren im Rahmen des Supply Chain Management*, S. Roderer Verlag, Regensburg
- Fleischmann, B.; Meyr, H. (2003) *Customer Orientation in Advanced Planning Systems*, in: Dyckhoff, H.; Lackes, R.; Reese, J. (Eds.): *Supply Chain Management and Reverse Logistics*, Berlin et al., pp. 297-321
- i2 Technologies (2000) *RHYTHM Demand Fulfillment Concept Manual*, Part No. 4.2.1-DOC-ITM-DFCM-ALL, Irving, Texas
- Schwendinger, J. (1979) *Master Production Scheduling's Available-to-Promise*, in: APICS Conference Proceedings, pp. 316-330
- Sharma, K. (1997) *Operating on the Edge*, private conversation
- Smith, B.; Leimkuhler, J.; Darrow, R.; Samuels, J. (1992) *Yield Management at American Airlines*, *Interfaces*, Vol. 22, pp. 8-31,

10 Production Planning and Scheduling

Hartmut Stadtler

Darmstadt University of Technology, Department of Production & Supply Chain Management, Hochschulstraße 1, 64289 Darmstadt, Germany

Assuming that the master plan has been generated, we can now derive detailed plans for the different plants and production units. In the following we will describe the underlying decision situation (Sect. 10.1) and outline how to proceed from a model to a solution (Sect. 10.2). Some of these steps will be presented in greater detail, namely model building (Sect. 10.3) and updating a production schedule (Sect. 10.4). Whether Production Planning and Scheduling should be done by a single planning level or by a two-level planning hierarchy largely depends on the production type of the shop floor. This issue will be discussed together with limitations of solution methods in Sect. 10.5.

10.1 Description of the Decision Situation

Production Planning and Scheduling aims at generating detailed production schedules for the shop floor over a relatively short interval of time. A *production schedule* indicates for each order to be executed within the planning interval its start and completion times on the resources required for processing. Hence, a production schedule also specifies the sequence of orders on a given resource. A production schedule may be visualized by a gantt-chart (see Fig. 10.4).

The *planning interval* for Production Planning and Scheduling varies from one day to a few weeks depending on the industrial sector. Its “correct” length depends on several factors: On the one hand it should at least cover an interval of time corresponding to the largest throughput time of an order within the production unit. On the other hand the planning interval is limited by the availability of known customer orders or reliable demand forecasts. Obviously, sequencing orders on individual resources is useful only if these plans are “reasonably” stable, i.e. if they are not subject to frequent changes due to unexpected events like changing order quantities or disruptions.

For some production types (like a job shop) Production Planning and Scheduling requires sequencing and scheduling of orders on potential bottlenecks. For other production types (like group technology) an automated, bucket-oriented capacity check for a set of orders to be processed by a group within the next time bucket(s) will suffice. Sequencing of orders may then be performed manually by the group itself.

Planning tasks can and should be done decentrally, utilizing the expertise of the staff at each location and its current knowledge of the state of the shop floor (e.g. the availability of personnel). Readers interested in the daily business of a planner and scheduler and resultant requirements for decision support are referred to McKay and Wiers (2004).

The master plan sets the frame within which Production Planning and Scheduling at the decentralized decision units can be performed. Corresponding directives usually are:

- the amount of overtime or additional shifts to be used,
- the availability of items from upstream units in the supply chain at different points in time,
- purchase agreements concerning input materials from suppliers – not being part of our supply chain.

Furthermore, directives will be given by the master plan due to its extended view over the supply chain and the longer planning interval. As directives we might have

- the amount of seasonal stock of different items to be built up by the end of the planning horizon (for production units facing a make-to-stock policy),
- given due dates for orders to be delivered to the next downstream unit in the supply chain (which may be the subsequent production stage, a shipper or the final customer).

10.2 How to Proceed from a Model to a Production Schedule

The general procedure leading from a model of the shop floor to a production schedule will be described briefly by the following six steps (see Fig. 10.1).

Step 1: Model building

A model of the shop floor has to capture the specific properties of the production process and the corresponding flows of materials in a detail that allows to generate feasible plans at minimum costs.

Only a subset of all existing resources on the shop floor – namely those which might turn out to become a bottleneck – will have to be modelled explicitly, since the output rate of a system is limited only by these potential bottlenecks. Details on model building are presented in Sect. 10.3.

Step 2: Extracting required data

Production Planning and Scheduling utilizes data from

- an ERP system,
- Master Planning and
- Demand Planning.

Only a subset of the data available in these modules will be used in Production Planning and Scheduling. Therefore, it is necessary to specify which data will actually be required to model a given production unit (see step 2 in Fig. 10.1).

Step 3: Generating a set of assumptions (a scenario)

In addition to the data received from sources like the ERP system, Master Planning and Demand Planning the decision-maker at the plant or production unit level may have some further knowledge or expectations about the current and future situation on the shop floor not available in other places (software modules). Also, there may be several options with respect to available capacity (e.g. due to flexible shift arrangements).

Therefore, the decision-maker must have the ability to modify data and thereby to set up a certain scenario (step 3, Fig. 10.1: A dotted frame indicates that this step has to be performed by the decision-maker and is optional).

Step 4: Generating a (initial) production schedule

Next, a (initial) production schedule will be generated for a given scenario, automatically (step 4, Fig. 10.1). This may be done either by a two-level planning hierarchy or in one step (for more details see Sect. 10.4).

Step 5: Analysis of the production schedule and interactive modifications

If there is a bucket-oriented upper planning level then this production plan may be analyzed first before a detailed schedule is generated (step 5, Fig. 10.1). Especially, if the production plan is infeasible, the decision-maker may indicate some course of action interactively to balance capacities (like the introduction of overtime or the specification of a different routing). This may be easier than modifying a detailed sequence of operations on individual resources (lower planning level). Infeasibilities – like exceeding an order's due date or an overload of a resource – are shown as *alerts* (see Sect. 13.1).

Also, a solution generated for a scenario may be improved by incorporating the experience and knowledge of the decision-maker, interactively. However, to provide real decision support, the number of necessary modifications should be limited.

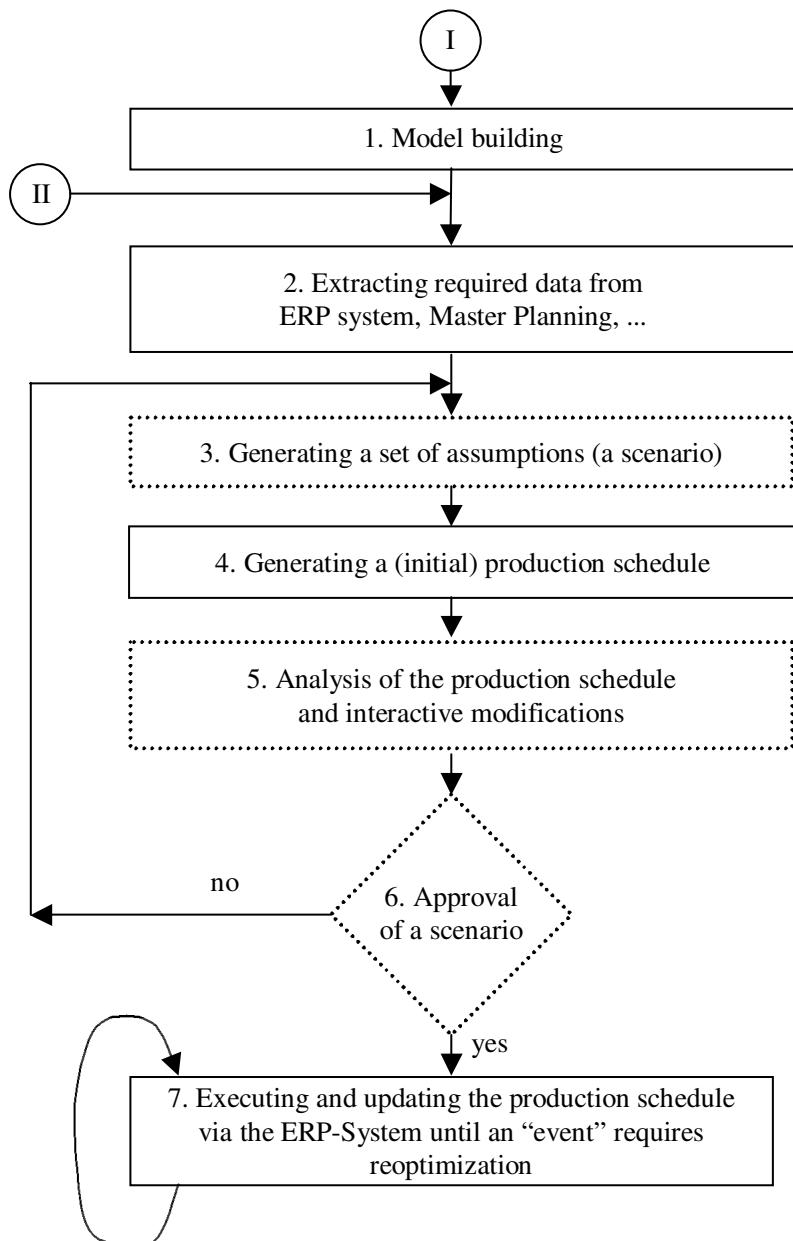


Fig. 10.1. General procedure for production scheduling

Step 6: Approval of a scenario

Once the decision-maker is sure of having evaluated all available alternatives, he / she will choose the most promising production schedule relating to a scenario.

Step 7: Executing and updating the production schedule

The production schedule selected will be transferred to

- the MRP module to explode the plan (Chap. 11),
- the ERP system to execute the plan and
- the Transport Planning module for finding vehicle loadings once customer orders have been completed.

The MRP module performs the explosion of all planned activities on bottleneck resources to those materials that are produced on non-bottleneck resources or those to be purchased from suppliers. Furthermore, required materials will be reserved for certain orders.

The schedule will be executed up to a point in time where an event signals that a revision of the production schedule seems advisable (loop II; Fig. 10.1). This may be an event like a new order coming in, a breakdown of a machine or a certain point in time where a given part of the schedule has been executed (for more details on updating a production schedule see Sect. 10.4).

Changing the model of the plant is less frequent (loop I; Fig. 10.1). If the structure remains unaltered and only quantities are affected (like the number of machines within a machine group or some new variants of known products), then the model can be updated automatically via the data that is downloaded from the ERP system. However, for major changes, like the introduction of a new production stage with new properties, a manual adaptation of the model by an expert is advisable.

We will now describe the task of modelling the production process on a shop floor in greater detail.

10.3 Model Building

A *model* of the shop floor has to incorporate all the necessary details of the production process for determining (customer) order completion times, the input required from materials and from potential bottleneck resources. The time grid of a production schedule is either very small (e.g. hours) or even continuous.

10.3.1 Level of Detail

The model can be restricted to operations to be performed on (potential) bottlenecks, since only these restrict the output of the shop floor.

Since Production Planning and Scheduling is (currently) not intended for controlling the shop floor (which is left to the ERP system) some details of the shop floor – like control points monitoring the current status of orders – can be omitted.

All processing steps to be executed on non-bottleneck resources in between two consecutive activities modelled explicitly are only represented by a *fixed lead-time offset*. This recommendation is no contradiction to the well-known statement that Advanced Planning yields lead-times as a *result of planning* and not as an *a priori* given constant. Here, the lead-time offset will consist only of processing and transportation times on preceding non-bottleneck resources, since waiting times would not exist.

The model can be defined by the associated data. We discriminate between structural data and situation-dependent data.

Structural data consists of

- locations,
- parts,
- bill of materials,
- routings and associated operating instructions,
- (production) resources,
- specification of suppliers,
- setup matrices and
- timetables (calendars).

In a large supply chain with many plants at different locations it may be advantageous to attribute all the data to a specific location. Consequently, a part can be discriminated by its production location even if it is the same in the eyes of the customer.

The bill of materials is usually described on a single-level basis (stored in a materials file). There, each part number is linked only to the part numbers of its immediate predecessor components. A complete bill of materials for a given part may be constructed easily on a computer by connecting the single-level representations.

The resource consumption per item can be obtained from the routings and operating instructions. Both the number of items per order as well as the resource consumption per item are required for sequencing and scheduling of individual orders. Hence, a combination of the two representations called *Production Process Model* (abbreviated by PPM) concept is appealing.

As an example the PPMs in Fig. 10.2 describe the two-stage production of ketchup bottles of a specific size and brand. The first PPM represents production of the liquid – including cleaning the tub, stirring the ingredients and waiting to be filled up in bottles. Once the liquid is ready it has to be

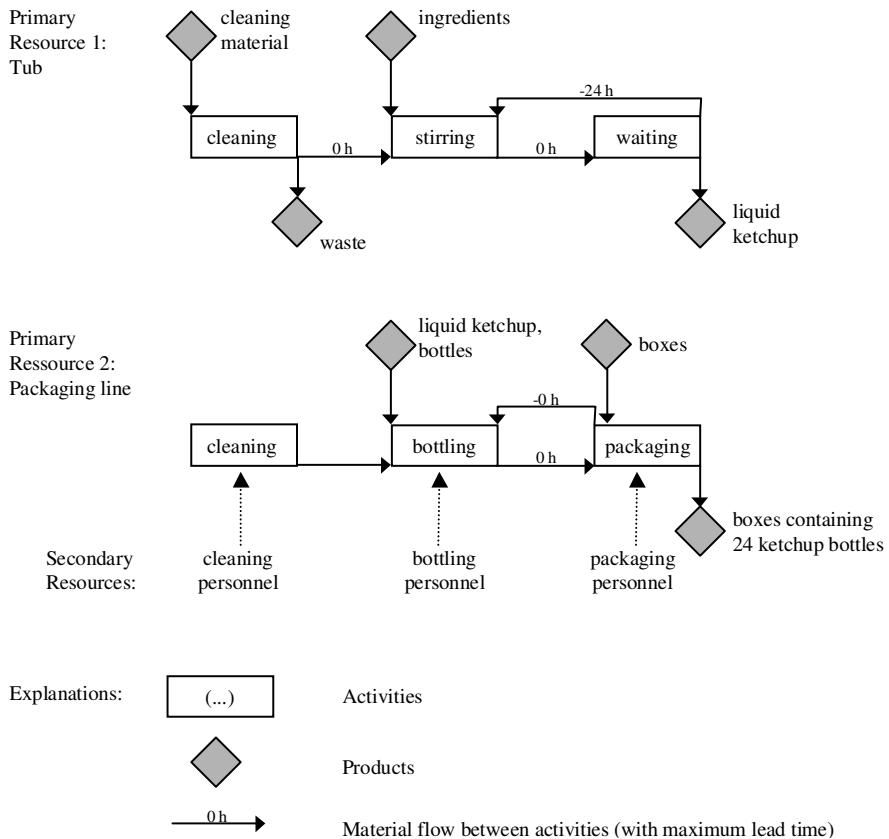


Fig. 10.2. A Production Process Model (PPM) for a two stage ketchup production

bottled within the next 24 hours. The liquid can be used in bottles of different sizes. For each size there will be an individual PPM. Also the liquid ketchup can be used up for different bottle sizes simultaneously.

A PPM is made up of at least one *operation* while each operation consists of one or several *activities*. An operation is always associated with one primary resource (like a tub). Secondary resources – like personnel – can also be attributed to an activity.

Activities may require some input material and can yield some material as an output. Surely, it has to be specified, at which point in time an input material is needed and when an output material is available. The technical sequence of activities within an operation – also called precedence relationships – can be represented by arcs. Like in project planning activities can be linked by

- end-start, end-end, start-end and start-start relationships together with
- maximal and minimal time distances.

This allows a very precise modelling of timing restrictions between activities including the parallel execution of activities (overlapping activities).

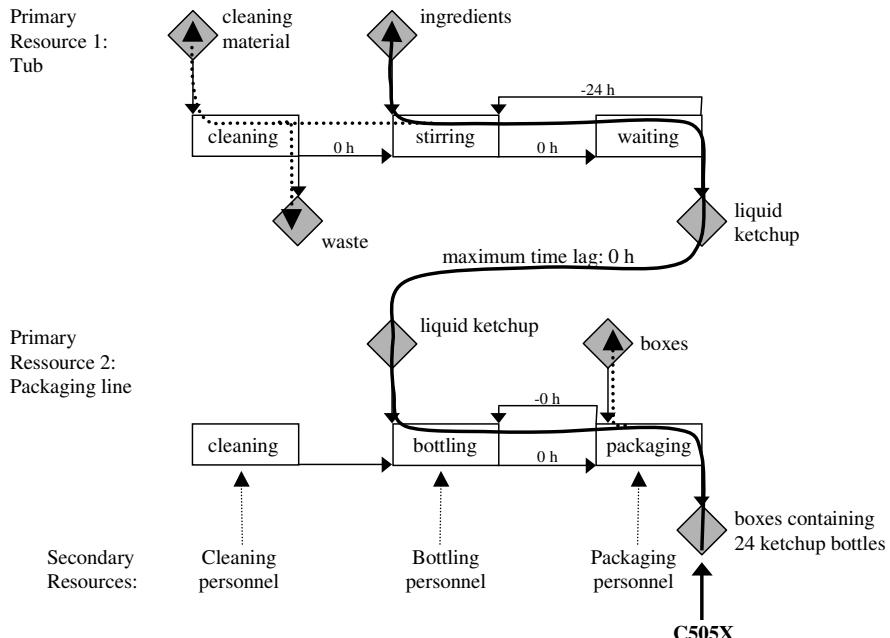


Fig. 10.3. Pegging: Linking two Production Process Models (PPMs)

The timing as well as the resource and material requirements of a (customer) order may be derived by linking the associated PPMs by the so-called *pegging arcs* (bolt and dotted arcs in Fig. 10.3). Pegging arcs connect the input material (node) of one PPM with the respective output material (node) of the predecessor (upstream) PPM. Consequently, exploding an order (see order C505X in Fig. 10.3) and the corresponding PPMs, starting with the final production stage, yields information about resource and material consumption within respective time windows. These time windows may be used directly when generating a feasible schedule (see also Vollmann et al. (1997, pp. 804)).

PPMs may be stored and updated solely within an APS. This option allows to take into account more details - like timing restrictions - than are usually stored and maintained in an ERP system. On the other hand operating instructions and routings are also kept in an ERP system. As one can imagine this may give rise to inconsistencies. Hence, some APS vendors propose to take (only) the data from the ERP system and to transfer the BOM, operating instructions and routings to the APS whenever a new production schedule will be created. From these so called *runtime objects* are created

resulting in the PPMs needed for an APS. Instead of using runtime objects also flat (ASCII) files may be used (like in the case study described in Chap. 23).

The (factory) calendar indicates breaks and other interruptions of working hours of resources. Another information included will be whether a plant (or resource) is operated in one, two or three shifts. Usually Advanced Planning Systems offer several typical calendars to choose from.

Situation-dependent data varies with the current situation on the shop floor. It consists of

- initial inventories, including work-in-process,
- setup state of resources and
- set of orders to be processed within a given interval of time.

Operational procedures to be specified by the user may consist of

- lot-sizing rules,
- priority rules or
- choice of routings.

Although rules for building lot-sizes should ideally be based on the actual production situation – like utilization of resources and associated costs – Advanced Planning Systems often require to input some (simple) rules a priori. Such rules may be a fixed lot-size, a minimum lot-size or a lot-size with a given time between orders. Software packages might either offer to pick a rule from a given set of rules or to programme it in a high level programming language. Rules for determining sequences of orders on a certain resource are handled in a similar fashion (for more details on priority rules see Silver et al. (1998, pp. 676)).

If alternative routings exist to perform a production order then one should expect that the system chooses the best one in the course of generating a production schedule. However, we experienced that the user has to pick one “preferred” routing. Sometimes alternative routings are input as a ranked list. Only if a preferred routing leads to infeasibilities the solver will try the second best routing, then the third best etc.

10.3.2 Objectives

Last but not least objectives will have to be specified. These guide the search for a good – hopefully near optimal – solution. As objectives to choose from within Production Planning and Scheduling we observed mainly time oriented objectives like minimizing the

- makespan,
- sum of lateness,
- maximum lateness,

- sum of throughput times and
- sum of setup times.

Three objectives referring to costs should be mentioned, too, namely the minimization of the sum of

- variable production costs,
- setup costs and
- penalty costs.

Although the degree of freedom to influence costs at this planning level is rather limited one can imagine that the choice of different routings, e. g. declaring an order to be a standard or a rush order, should be evaluated in monetary terms, too.

Penalty costs may be included in the objective function, if *soft constraints* have been modelled (e. g. fulfilling a planned due date for a make-to-stock order).

If the decision-maker wants to pursue several of the above objectives, an “ideal” solution, where each objective is at its optimum, usually does not exist. Then a compromise solution is looked for. One such approach is to build a weighted sum of the above individual objectives. This combined objective function can be handled like a single objective, and hence, the same solution methods can be applied (for more details on multi-objective programming see Tamiz (1996)).

10.3.3 Representation of Solutions

There are several options for representing a model’s solution, namely the detailed production schedule. It may simply be a list of activities with its start and completion times on the resources assigned to it. This may be appropriate for transferring results to other modules.

A decision-maker usually prefers a *gantt-chart* of the production schedule (see Fig. 10.4). This can be accomplished by a gantt-chart showing all the resources of the plant in parallel over a certain interval of time. Alternatively one might concentrate on a specific customer order and its schedule over respective production stages. Likewise, one can focus attention on one single resource and its schedule over time.

If the decision-maker is allowed to change the production schedule interactively – e. g. by shifting an operation to another (alternative) resource – a gantt-chart with all resources in parallel is the most appropriate.

Now we will point our attention to the options of updating an existing production schedule.

10.4 Updating Production Schedules

Production Planning and Scheduling assumes that all data is known with certainty, i.e. the decision situation is deterministic. Although this is an ideal assumption, it may be justified for a certain interval of time. To cope with uncertainty – like unplanned variations of production rates or unexpected downtime of resources – software tools allow monitoring deviations from our assumptions taking place on the shop floor immediately, resulting in updated expected completion times of the orders. Whether these changes are that large that a reoptimized schedule is required will be based on the decision-maker's judgment. Current software tools will enhance this judgment by providing extensive generation and testing facilities of alternative scenarios (also called *simulation*) before a schedule is actually delivered to the shop floor (see also steps three to five; Fig. 10.1).

Another feature to be mentioned here is a two step planning procedure – also called *incremental planning*. Assume that a new order comes in. If it falls into the planning horizon of Production Planning and Scheduling the activities of this new customer order may be inserted into the *given sequence* of orders on the required resources. Time gaps are searched for in the existing schedule such that only minor adjustments in the timing of orders result. If feasibility of the schedule can be maintained a planned due date for the new customer order can be derived and sent back to the customer.

Since this (preliminary) schedule may be improved by a different sequence of orders, reoptimization is considered from time to time, aiming at new sequences with reduced costs.

The following example will illustrate this case. Assume there are four orders that have to be scheduled on a certain machine with given due dates and the objective is to minimize the sum of sequence dependent setup times. Then the optimal sequence will be A-B-C-D (see Fig. 10.4). The current time is 100 (time units). Processing times for all orders are identical (one time unit). Sequence dependent setup times are either 0, 1/3, 2/3 or 1 time unit.

Table 10.1. Data: Due dates

Order	A	B	C	D
Due Dates	102	104	107	108

After having started processing order A, we are asked to check whether a new order E can be accepted with due date 107. Assuming that *preemption* is not allowed (i.e. interrupting the execution of an order already started in order to produce another (rush) order), we can check the insertion of job

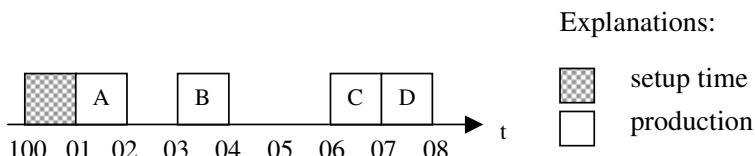
Table 10.2. Data: Matrix of setup times

to	A	B	C	D	E
A	1	0	1	1	1
B	1	1	0	1	$\frac{2}{3}$
C	1	1	1	0	$\frac{1}{3}$
D	1	1	$\frac{1}{3}$	1	1
E	1	1	$\frac{2}{3}$	1	1

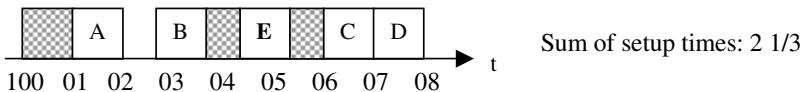
E in the existing sequence directly after finishing orders A, B, C or D (see Fig. 10.5). Since there is a positive setup time between order A and E this sub-sequence will not be feasible since it violates the due date of order B. Three feasible schedules can be identified, where alternative c has the least sum of setup times. Hence, a due date for order E of 107 can be accepted (assuming that order E is worth the additional setup time of one time unit).

Once reoptimization of the sequence can be executed, a new feasible schedule – including order E – will be generated reducing the sum of setup times by $1/3$ (see Fig. 10.6).

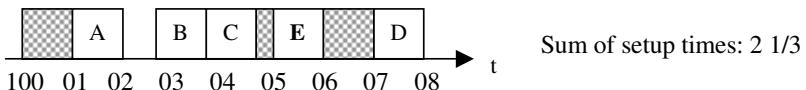
Generating new sequences of orders is time consuming and usually will result in some *nervousness*. We discriminate nervousness due to changes regarding the start times of operations as well as changes in the amount to be produced when comparing an actual plan with the previous one. Nervousness can lead to additional efforts on the shop floor – e. g. earlier deliveries of some input materials may be necessary which has to be checked with suppliers. In order to reduce nervousness usually the “next few orders” on a resource may be *firmed* or *fixed*, i.e. their schedule is fixed and will not be part of the re-optimization. All orders with a start time falling within a given interval of time – named *frozen horizon* – will be firmed.

**Fig. 10.4.** Gantt-chart for four orders on one machine with due dates and sequence dependent setup times

Alternative a)



Alternative b)



Alternative c)

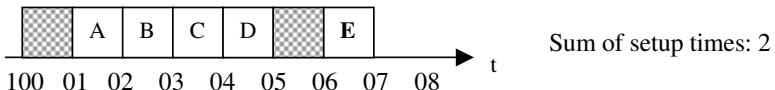


Fig. 10.5. Generating a due date for the new customer order E

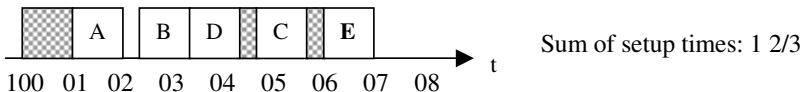


Fig. 10.6. Reoptimized schedule

10.5 Number of Planning Levels and Limitations

10.5.1 Planning Levels for Production Planning and Scheduling

As has been stated above, software modules for Production Planning and Scheduling allow to generate production schedules either within a single planning level or by a two-level planning hierarchy. Subsequently, we will discuss the pros and cons of these two approaches.

Drexel et al. (1994) advocate that the question of decomposing Production Planning and Scheduling depends on the production type given by the production process and the repetition of operations (see Chap. 3 for a definition). There may be several production units within one plant each corresponding to a specific production type to best serve the needs of the supply chain. Two well-known production types are process organization and flow lines.

In *process organization* there are a great number of machines of similar functionality within a shop and there are usually many alternative routings for a given order. An end product usually requires many operations in a multi-stage production process. Demands for certain operations may be combined

to a lot-size in order to reduce setup costs and setup times. Usually many lot-sizes (orders) have to be processed within the planning interval (e.g. the next eight to sixteen) weeks.

In order to reduce the computational burden and to provide effective decision support the overall decision problem is divided into two (hierarchical) planning levels. The upper planning level is based on time buckets of days or weeks, while resources of similar functionality are grouped in resource groups. These big time buckets allow to refrain from sequencing. Consequently, lot-size decisions and capacity loading will be much easier. Given the structure of the solution provided by the upper planning level, the lower planning level will perform the assignment of orders to individual resources (e.g. machines) belonging to a resource group as well as the sequencing. The separation of the planning task into two planning levels requires some slack capacity or flexibility with respect to the routing of orders.

For (automated) *flow lines* with sequence dependent setup times a separation into two planning levels is inadequate. On the one hand a planning level utilizing big time buckets is not suited to model sequence dependent setup costs and times. On the other hand sequencing and lot-sizing decisions cannot be separated here, because the utilization of flow lines usually is very high and different products (lot-sizes) have to compete for the scarce resource. Luckily, there are usually only one to three production stages and only a few dozen products (or product families) to consider, so Production Planning and Scheduling can be executed in a single planning level.

In the following some definitions and examples illustrating the pros and cons of the two approaches will be provided.

A *time bucket* is called *big*, if an operation started within a time bucket has to be finished by the end of the time bucket. The corresponding model is named a big bucket model. Hence, the planning logic assumes that the setup state of a resource is not preserved from one period to the next. Usually, more than one setup will take place within a big time bucket of a resource (see Fig. 10.7).

In a model with *small* time buckets the setup state of a resource can be preserved. Hence, the solution of a model with small time buckets may incur less setup times and costs than the solution of a model with big time buckets (see operation B in Fig. 10.8). Usually, at most one setup can take place within a small time bucket of a resource (a further example is given by Haase (1994, p. 20)).

An aggregation of resources to resource groups automatically leads to a big bucket model, because the setup state of an individual resource as well as the assignment of operations to individual resources is no longer known.

However, although a feasible big bucket oriented production plan exists, there may be no feasible disaggregation into a production schedule on respective resources. This can occur in cases such as

- sequence dependent setup times,

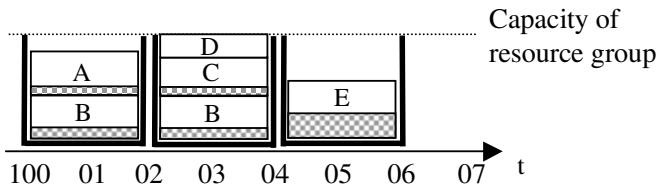


Fig. 10.7. A big bucket model

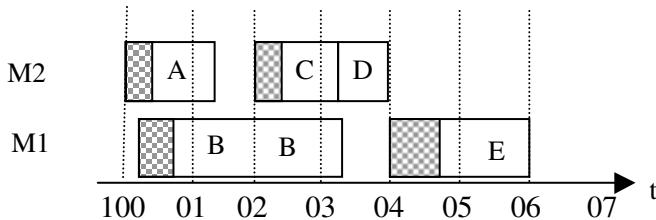


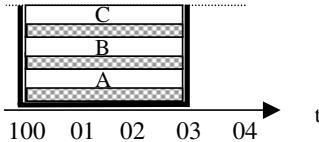
Fig. 10.8. Gantt-chart: A small bucket model (Time buckets of one time unit and two resources M1 and M2)

- loading of resource groups and
- a lead-time offset of zero time units between two successive operations.

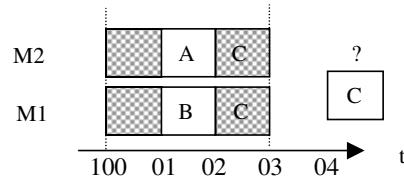
Sequence dependent setup times cannot be represented properly within a big bucket model, since the loading of a time bucket is done without sequencing. Usually a certain portion of the available capacity is reserved for setup times. However, the portion may either be too large or too small. The former leads to unnecessarily large *planned throughput* times of orders while the latter may result in an infeasible schedule. Whether the portion of setup times has been chosen correctly is not known before the disaggregation into a schedule has been performed.

Another situation where a feasible disaggregation may not exist is related to *resource groups*. As an example (see Fig. 10.9), assume that two resources have been aggregated to a resource group, the time bucket size is three time units, thus the capacity of the resource group is six time units. Each operation requires a setup of one time unit and a processing time of one time unit. Then the loading of all three operations within one big time bucket is possible. However, no feasible disaggregation exists, because a split of one operation such that it is performed on both machines requires an additional setup of one time unit exceeding the period's capacity of one machine. To overcome this dilemma one could reduce the capacity of the resource group to five time units (resulting in a slack of one time unit for the lower planning level). Then only two out of the three operations can be loaded within one time bucket. However, one should bear in mind that this usually will lengthen the planned throughput time of an order.

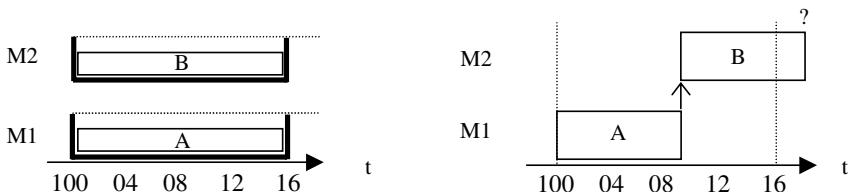
Production plan:



Corresponding detailed schedule:

**Fig. 10.9.** An example of no feasible schedule when resource groups are loaded

For a multi-stage production system with several potential bottlenecks on different production stages, a feasible schedule might not exist if an order requiring two successive operations is loaded in the same big time bucket. As an example (see Fig. 10.10) depicting a two-stage production system with operation B being the successor of operation A each with a processing time of nine hours. A production stage is equipped with one machine (M1 and M2 respectively). A time bucket size of 16 hours (one working day) has been introduced. Although the capacity of one time bucket is sufficient for each operation individually, no feasible schedule exists that allows both operations to be performed in the same time bucket (assuming that overlapping operations are prohibited).

**Fig. 10.10.** An example of no feasible schedule in multi-stage production with lead-time offset zero

A feasible disaggregation can be secured if a fixed *lead-time offset* corresponding to the length of one big time bucket is modelled. Again this may incur larger planned throughput times than necessary (32 hours instead of 18 hours in our example).

Consequently, it has to be considered carefully which of the above aggregations makes sense in a given situation. Usually, the answer will depend on the production type. Surely, an intermediate bucket oriented planning level can reduce the amount of detail and data to be handled simultaneously, but may also require some planned slack to work properly leading to larger planned throughput times than necessary.

In order to combine the advantages of both the big and the small bucket model a third approach – a big time bucket with linked lot-sizes – has been proposed in Suerie and Stadtler (2003). Here, several lots may be processed within a time bucket without considering its sequence (hence a big bucket model). However, a “last” lot within a time bucket is chosen which can be linked with a “first” lot in the next time bucket. If these two lots concern the same product a setup will be saved. While this effect may only seem to be marginally at first sight, it also allows to model the production of a lot-size extending over two or more time buckets with only one initial setup – like in a small bucket model.

Last but not least a fourth approach has to be mentioned which does not use time buckets at all, instead a continuous time axis is considered (Maravelias and Grossmann, 2003). Although this is the most exact model possible it usually will result in the greatest computational effort.

10.5.2 Limitations Due to Computational Efforts

For finding the best production schedule one has to bear in mind that there are usually many alternatives for sequencing orders on a resource (of which only a subset may be feasible). Theoretically, one has to evaluate $n!$ different sequences for n orders to be processed on one resource. While this can be accomplished for five orders quickly by complete enumeration ($5! = 120$), it takes some time for ten orders ($10! > 3.6 \cdot 10^6$) and cannot be executed within reasonable time limits for 20 orders ($20! > 2.4 \cdot 10^{18}$). Furthermore, if one has the additional choice among parallel resources, the number of possible sequences again rises sharply. Although powerful solution algorithms have been developed that reduce the number of solutions to be evaluated for finding *good* solutions (see Chaps. 28 and 29), computational efforts still increase sharply with the number of orders in the schedule.

Fortunately, there is usually no need to generate a production schedule from scratch, because a portion of the previous schedule may have been fixed (e.g. orders falling in the frozen horizon). Similarly, decomposing Production Planning and Scheduling into two planning levels reduces the number of feasible sequences to be generated at the lower planning level, due to the assignment of orders to big time buckets at the upper level.

Also, incremental planning or a reoptimization of partial sequences specified by the decision-maker will restrict computational efforts.

Further details regarding the use of Production Planning and Scheduling are presented in two case studies (Chaps. 24 and 23) as well as in Kolisch et al. (2000).

References

- Drexel, A.; Fleischmann, B.; Günther, H.-O.; Stadtler, H.; Tempelmeier, H. (1994)
Konzeptionelle Grundlagen kapazitätsorientierter PPS-Systeme, Zeitschrift für betriebswirtschaftliche Forschung, Vol. 46, 1022–1045

- Haase, K. (1994) *Lotsizing and scheduling for production planning*, Lecture Notes in Economics and Mathematical Systems, Vol. 408, Berlin et al.
- Kolisch, R.; Brandenburg, M.; Krüger, C. (2000) *Numetrix/3 Production Scheduling*, OR Spektrum, Vol. 22, No. 3, 307–312
- Maravelias, C. T.; Grossmann, I. E. (2003) *New general continuous-time state-task network formulation for short-term scheduling of multipurpose batch plants*, Ind. Eng. Chem. Res. , Vol. 42, 3056–3074
- McKay, K. N.; Wiers, V. C. S. (2004) *Production control in practice: a survival guide for planners and schedulers*, Boca Raton, Florida, USA (in print)
- Reeves, C. R. (Ed.) (1993) *Modern heuristic techniques for combinatorial problems*, Oxford
- Silver, E. A.; Pyke, D. F.; Petersen, R. (1998) *Inventory management and production planning and scheduling*, 3rd ed., New York et al.
- Suerie, C.; Stadtler, H. (2003) *The capacitated lot-sizing problem with linked lot sizes*, Management Science, Vol. 49, 1039–1054
- Tamiz, M. (1996) *Multi-objective programming and goal programming, theories and applications*, Lecture Notes in Economics and Mathematical Systems, Vol. 432, Berlin et al.
- Vollman, T. E.; Berry, W. L.; Whybark, D. C. (1997) *Manufacturing planning and control systems*, 4th ed., New York et al.

11 Purchasing and Material Requirements Planning

Hartmut Stadtler

Darmstadt University of Technology, Department of Production & Supply Chain Management, Hochschulstraße 1, 64289 Darmstadt, Germany

An indispensable part of an ERP system, Material Requirements Planning, also plays an important role in APS, because it

- generates replenishment orders (production orders) for uncritical components and parts (operations) in a multi-stage production environment (Sect. 11.1 and 11.2) and
- provides access to a transactional ERP system and thus can initiate the execution of orders.

The typical tasks of purchasing are to analyse procurement markets, to negotiate the terms of trade with potential suppliers and finally to select suppliers and to place replenishment orders. Here, we are interested in the way APS can support the selection of suppliers and the decisions on order sizes, taking into account the specific cost functions of suppliers, which often allow for quantity discounts (Sect. 11.3). This may apply to input materials for production, indirect materials and articles of merchandise.

11.1 Basics of Material Requirements Planning

Material Requirements Planning (MRP) is regarded as the core engine of an ERP system, which calculates time-phased plans of secondary demands for components and parts based on a time series of primary demands (usually finished products). Time-phased secondary demands are a prerequisite for generating production or replenishment orders so that demands for finished products can be met in time with as little work-in-process and inventory as possible.

Although most appealing, this logic suffers from the ignoring of available capacities. Consequently, production orders may result in overloaded capacities and thus infeasibilities. Experience has shown that also a two step procedure, i. e. first calculating all secondary demands and then balancing capacities by means of an ERP's capacity requirements planning (CRP) module, does not provide satisfactory solutions (for a further discussion of the drawbacks of ERP systems see Drexel et al. (1994) or Tempelmeier and Derstroff (1996)).

These drawbacks gave rise to develop APS, which do not separate the generation of secondary demands and capacity balancing. However, in order

to reduce complexity, APS concentrate on operations to be performed on potential bottlenecks, which usually are only a small subset of all operations relating to factory orders. The time needed to execute non-bottleneck operations (including transport) in between two adjacent *critical operations* is taken into account by a fixed lead-time offset. Once plans have been generated for critical operations, the timing and quantities of non-critical operations can be calculated easily by making use of the standard MRP logic. This is the topic of this section.

There are many textbooks that describe the MRP logic (e.g. Silver et al. (1998) and Vollman et al. (1997)). Thus we will only briefly describe the terms and the basic logic. More important is a discussion of issues occurring when using MRP in conjunction with an APS.

First of all, we have to decide on the series of primary demands to take as a starting point. These may be (see Fig. 11.1)

- production quantities per period for (critical) product groups calculated in Master Planning (see Chap. 8),
- production quantities per period for critical operations calculated in the Production Planning module or
- critical production orders generated in the Scheduling module (see Chap. 10).

In case that we look for the requirements of parts to be purchased from outside suppliers over a longer period of time (e.g. for negotiating contracts with suppliers or providing an outlook of expected part demands to suppliers), Master Planning will be the starting point. Note that demands for product groups have to be disaggregated into demands of respective products before starting the MRP logic.

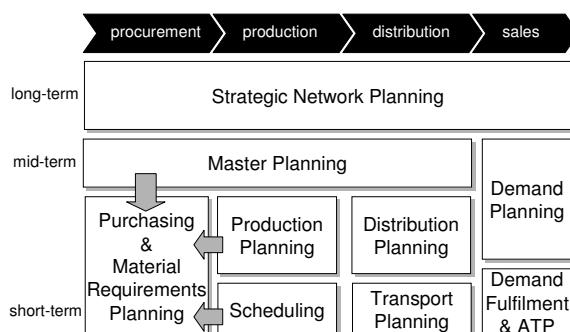


Fig. 11.1. Modules providing the input data (production quantities) for Purchasing and MRP

For placing replenishment orders or for the timing of uncritical operations (production orders), either Production Planning or Scheduling will be the

source of information. If Production Planning is chosen, demands per time bucket will result, while Scheduling will give the exact timing of the start of production orders. Hence, Scheduling best corresponds to a bucketless (continuous time axis) MRP, while the two former are best suited for a bucket oriented MRP logic. Both time axes are possible today (Vollman et al., 1997, pp. 30). In the following, we assume Production Planning to be the starting point.

As additonal data we will need:

- bill of materials, indicating for each part number, which other part numbers are required as direct inputs,
- production coefficients indicating the quantity of each direct input part needed for one unit of a given part number,
- lead-times representing a fixed interval of time needed between releasing an order for a part number and its availability,
- the inventory status, indicating for each part number, the (physical) stock at hand, scheduled receipts (i. e. outstanding orders and work-in-process), reservations, backorders and safety stock levels and
- low-level code (numbers).

A *low-level code* of a part number or operation corresponds to the longest path in the product structure starting with an end item and terminating in the respective part number. All parts visited along the path are counted yielding the level code. Due to the fact that a part number may be used in several product structures, the maximum has be taken for determining the low-level code. By definition, a low-level code “0” is attributed to end items (for an example see Fig. 11.2). Low-level codes have to be calculated preceding the bill of materials (BOM) explosion, i. e. the generation of secondary demands, to allow a pure sequential execution of calculations.

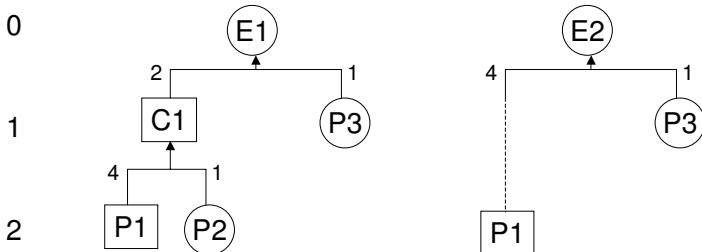
While in standard text books on MRP the level of detail for a BOM explosion is finished products, components or parts, the level of detail required in the context of APS is *operations*. Normally, several operations are required to transfer input material(s) into a specific part. Some of these operations may be critical, i. e. they have to be performed on a potential bottleneck resource, some are uncritical. Consequently, we will have to combine the BOM with the routing of operations – sometimes called the bill of capacities (BOC) (Vollman et al., 1997, p. 128).

To ease understanding we will simplify matters (without loss of generality) by assuming that there is exactly one operation to a finished product, component or part.

11.2 Generation and Timing of Uncritical Orders

The generation of uncritical orders originating from production orders scheduled on bottleneck resources will be explained now in an example. Firstly,

the data required – like the BOM – will be presented (see Fig. 11.2). Secondly, some remarks on the generation of a production plan will follow and thirdly, we will show how to derive orders for uncritical operations. Fourthly, a simplification is shown as proposed by APS vendors today.



Explanations:

- E1, E2 represent end products, C1 a component and P1, P2, P3 single parts
- single digits indicate production coefficients
- materials in circles are regarded as *critical*, materials in boxes as *uncritical*

Fig. 11.2. Bill of materials for end products E1 and E2 as well as low-level codes

E1 and E2 are completed on a highly utilized assembly line. Component C1 is produced in a manufacturing cell. Since the manufacturing cell is underutilized if only C1 is produced, surplus capacity has been sold to a partner company. The terms of the contract establish priorities for scheduling operation C1; hence, the manufacturing cell is no bottleneck. P1 is bought from an external supplier, while P2 and P3 are processed on a moulding machine, which is also often a bottleneck.

Consequently, E1, E2, P2 and P3 are regarded as critical operations for which a production plan is generated by the APS module Production Planning.

In addition to the data shown in Fig. 11.2, lead-time offsets are needed for each operation, they are always one period except for C1, which has a lead-time of two periods.

While lot-sizing plays a major role for critical operations, incurring setup times or setup costs on potential bottlenecks, this is generally negligible on non-bottlenecks. Since time is not scarce at non-bottlenecks, an hour saved by saving setup time is of no value. Hence, a lot-for-lot production, i. e. no lot-sizing, is advisable. Exceptions may only occur in case of technological reasons relating to production or transport activities requiring some minimum quantity or integer multiple of a fixed amount to work properly (e. g. production in full tub loads).

In contrast to lead-times used in an ERP system, which usually incorporate a large portion of waiting times, lead-times in the context of an APS pertaining to uncritical operations should only cater for production

and transport activities. The reason is that, by definition, utilization rates of non-bottlenecks are low and thus a production order should find the resource empty in general. However, it seems wise to include “some” safety time into the lead-time offset of uncritical operations that is the *direct predecessor* of a critical operation. This will allow for some uncertainties in processing times and will make sure that a bottleneck resource, which governs the throughput of the whole supply chain, will not run empty. Another reason why an APS can do with smaller lead-times than an ERP system (and thus smaller planned throughput times) is due to the fact that lead-times in an ERP system also cater for its inability to take into account finite capacity checks of bottleneck resources when making the BOM explosion.

From these lead-times now *cumulated lead-times* have to be calculated in between two adjacent critical operations simply by adding the single lead-times of operations along the path (in the BOM) from the upstream critical operation to the downstream critical operation – excluding the lead-time of the two critical orders. Thereby, the *starting* point (period) of the downstream critical operation is connected with the *finishing* point (period) of the upstream critical operation. Consequently, cumulated lead-times cover production times and transport activities in between two critical operations. In order to prevent overlapping operations, which might cause infeasibilities in Scheduling, another lead-time offset of one period is often added (e.g. resulting in cumulated lead-times for E1-P2, E1-P3 and E2-P3 of 3, 1 and 1 period(s), respectively).

These cumulated lead-times, as well as (cumulated) production coefficients, primary demands and the inventory status of items, parts and components, form the input to Production Planning.

Figure 11.3 shows the primary demands for finished products (critical operations) E1 and E2 and their relation to production orders to meet demands for the upcoming five periods, while taking into account a lead-time offset of one period (see solid arrows). Note, that some demands are fulfilled from initial inventory (dashed arrows).

		period	1	2	3	4	5
material							
E1	demands starting inv. order		- 30 40 10	- 20 10 30	- 30 20	- 20 30	- 30
E2	demands starting inv. order		- 20 20 10	- 20 10	- 20 20	- 10	- 30

Fig. 11.3. Primary demands and production plans for E1 and E2 (in quantities per period; inventory abbreviated by inv.)

Positive lead-times are the reason why there are no production orders for E1 and E2 in period five even though the forecast and planning horizon is five

periods. Similarly, even for materials with a low-level code greater than “0” production orders cover a smaller interval of time. Consequently, utilization rates near the planning horizon should be interpreted with caution. Furthermore, it becomes clear that a reasonable planning horizon for Production Planning should at least cover the longest path, with respect to lead-times, from a final operation (finished product) to a part with no direct predecessor in the BOM. In our example, the longest path is E1-C1-P1 or E1-C1-P2, both with an overall lead-time offset of four periods. An appropriate planning horizon should also cover a (small) *frozen horizon* and some periods for manoeuvring (e.g. for making lot-size decisions).

To keep our example small production plans for critical operations P2 and P3 are not exhibited here, because they don't cause secondary demands. Now we are in the position of calculating the time-phased order sizes of uncritical operations.

LLC	Operation		demand/order per period				
			1	2	3	4	5
0	E1	order	10	30	20	30	-
0	E2	order	10	10	20	10	-
1	C1	starting inv.	80	60	-	-	-
		gross dem.	20 (E1)	60 (E1)	40 (E1)	60 (E1)	-
		net dem.	-	-	40	60	-
		order	40	60	40	60	-
2	P1	starting inv.	200	-	-	-	-
		gross dem.	40 (E2)	40 (E2)	80 (E2)	40 (E2)	-
		net dem.	160 (C1)	240 (C1)	- (C1)	- (C1)	-
		order	280	80	80	40	-

Explanations:

- LLC: low-level code
- inv.: inventory

Fig. 11.4. BOM explosion with pegging

Here, the logic of a time-phased BOM explosion (Orlicky, 1975; Tempelmeier, 2003) has to be slightly adapted. First, finished products (i.e. final operations) are always declared “critical”. Second, all orders for critical operations *and* possessing at least one uncritical direct predecessor (i.e. upstream) operation, are labeled with low-level code “0”. Now we can start with any operation belonging to low-level code “0” and derive the associated secondary demands for all its uncritical direct predecessor operations by multiplying a period's order size (e.g. generated in Production Planning) by the production coefficient and placing it in the same time period; e.g. the order for operation C1, for 20 units, must be ready at the beginning of period 1 in order to be used for the assembly operation E1 in period 1 (see Fig. 11.4). In order to know which operation caused the secondary demand we further store its name – (see the operation's names in brackets in Fig. 11.4). This identification is called *pegging* and can be most useful in the case that operations are

not ready in time. Then, it is easy to see which orders are affected and thus specific counter actions can be initiated.

Once direct secondary demands have been calculated for all low-level code “0” operations, then secondary demands of low-level code “1” operations are complete. Next, we can calculate orders for any low-level code “1” operation and explode these into the secondary demands of its direct predecessors. This is only necessary for uncritical direct predecessors, because a production plan exists for the critical operations. (However, a BOM explosion into critical operations may also be useful in order to check the feasibility of the production plan. In case there is a mismatch of orders between the production plan and the BOM explosion, an alert should be generated automatically).

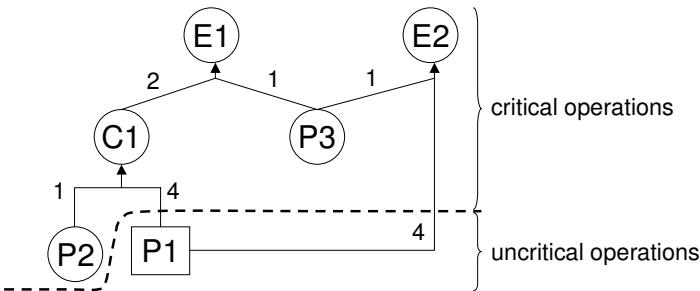
Before starting the BOM explosion, we will have to calculate net demands by netting gross demand with initial inventory. This logic may be more elaborate than shown in our example by considering safety stock requirements, outstanding orders and reservations. Given the net demands of an operation these have to be time-phased and assigned to an order period by taking into account the operation’s lead-time offset (indicated by an arrow in Fig. 11.4). These tasks are repeated until all operations have been considered.

One may ask what reasons there are for generating an alert during the BOM explosion. Obviously, if we started from an infeasible production plan, e.g. with backlogging, then the BOM explosion would also generate alerts showing that some materials are not ready in time. At this stage a popular counter measure would be *expediting*, resulting in reduced lead-times. A second reason for a mismatch of a (feasible) production plan and the result of a BOM explosion may be that lead-times used in Production Planning are independent of the amount produced, while in a BOM explosion lead-times can be calculated based on the order size. Again, any discrepancy jeopardizing efficiency or feasibility should be shown to the decision maker by an alert.

While the logic of the BOM explosion is rather simple, implementing the interface between the Production Planning module and the MRP module may be tricky. One issue is the generation and exchange of alerts between modules.

In order to reduce the complexity of an arbitrary mix of critical and uncritical operations, some APS vendors propose a distinct separation: The final operation, resulting in a finished good, is always defined as critical. Also, any upstream operation can be defined as critical. However, a critical operation may never possess a direct uncritical downstream operation. This can best be illustrated by our example (Fig. 11.2) transformed into a Gozinto graph (Fig. 11.5). Here, a separation line divides operations into the set of critical operations and the set of uncritical operations.

The advantage is that Production Planning can be executed first, followed by the BOM (or BOC) explosion for uncritical operations – and one can be sure that both plans will match. Hence, an exchange of alerts between modules is unnecessary. Also, there is no need to calculate, maintain and use



Explanations:

- E1, E2 represent end products, C1 a component and P1, P2, P3 single parts
- single digits indicate production coefficients
- materials in circles are regarded as *critical*, materials in boxes as *uncritical*
- the dashed line separates critical from uncritical operations

Fig. 11.5. Gozinto representation of the bill of materials with a separation line for the set of critical and the set of uncritical operations

cumulated lead-times or cumulated production coefficients. The disadvantage is that some formerly uncritical operations now have to be declared as critical (e.g. C1), which increases the scope and efforts of Production Planning. Especially, if the most upstream operations are processed on a bottleneck resource then (nearly) all operations in the BOC have to be defined as critical.

Referring to our example, the generation of purchase orders for P1 now starts from production orders for E2 and C1 (see Fig. 11.6). For simplification purposes, we assume here that production orders for C1, generated by Production Planning, are equal to those derived by the BOM explosion (Fig. 11.4). Now, applying the BOM explosion for P1 provides the same results as before. The only difference is that computational efforts will be smaller, while they will be larger for Production Planning (not shown here).

Given that the production plan started from is feasible and no alerts have been generated during the BOM explosion, then all production orders for critical and uncritical operations are known and can be handed over for execution (at least for the upcoming period, see Chap. 4). The only exception are purchase orders from outside suppliers, which may need further attention due to fixed ordering costs or quantity discounts – which will be dealt with next.

11.3 Quantity Discounts and Supplier Selection

Life cycle contracts are predominant today in many industries for the most important production input. Also, materials to be purchased and considered strategically important are usually procured from a supply chain partner. However, there are a number of additional materials, which are purchased

LLC	Operation	demand/order per period				
		1	2	3	4	5
0	E1	order	10	10	20	10
0	C1	order	40	60	-	-
1	P1	starting inv. gross dem. net dem. order	200 40 (E2) 160 (C1) -	40 (E2) 240 (C1) 280 80	80 (E2) - (C1) 80 40	40 (E2) - (C1) 40 -

Explanations:

- LLC: low-level code
- dem.: demand
- inv.: inventory

Fig. 11.6. BOM explosion with pegging

from outside suppliers, where it may be economical to select a supplier and to decide on the order size in the short term and to make use of quantity discounts. These materials may be commodities used as direct production input, often classified as C items, as well as materials for maintenance, repair and operations (MRO). In the case of a commodity, quality is also defined by industry standards and there are usually a number of suppliers to choose from. Also, it can be assumed that the quantity to be purchased is rather low compared to the overall market volume so that availability is no problem. Examples are standard electronic components, like a capacitor, or office equipment bought with the help of an e-catalogue.

In an abstract form the procurement decision incorporates the following features (Tempelmeier, 2002): For each item to be purchased there is a series of demands over a finite planning interval (e.g. see row “order” for item P1, Fig. 11.4). There may be one or several suppliers to choose from, each with specific costs. These costs will incur

- supplier specific fixed ordering and procurement costs (including the transport of the consignment) and
- supplier specific quantity discounts (either all-units or incremental discounts).

Figure 11.7 illustrates the two most popular forms of quantity discounts.

Here, the supplier’s fixed ordering cost is depicted as “U” on the total acquisition cost axis. The x-axis represents the order quantity. There are three purchasing intervals, each with a specific price per unit. In the *all-units discount* case, the price charged for the last unit ordered also holds for the total order quantity. In an *incremental discount* case, only those units falling within a purchasing interval are charged with the corresponding price (see lower bounds Q_1 and Q_2 of purchasing intervals 2 and 3 in Fig. 11.7). In both cases it is wise to stick to one supplier and item per period and not to split the order, because this will result in the lowest total acquisition cost. Only if the amount ordered exceeds the maximum a supplier is able to procure (Q_3) will another supplier come into play.



Fig. 11.7. Incremental discounts and all-units discount with three purchasing intervals

In general, the demand of several periods will be combined when forming purchase orders in order to make use of attractive price reductions for a large quantity. Large order quantities usually result in holding stocks for some periods; thus, holding costs counteract savings due to quantity discounts. To find a procurement plan that results in minimal costs over the planning horizon will be the task of the APS module on purchasing.

Note, that it might be difficult to specify an item's "correct" holding cost per period because a large portion of the holding cost is interest on the capital employed. Since an item's purchase price can change over time – especially if there are time-dependent, supplier-specific quantity discounts – one does not know in *advance* which items will be in inventory and at which price. One way to overcome this "problem" is to keep track of each item purchased, its purchase price, purchasing period and the period of consumption.

In a practical setting, one often has to take into account supplier-specific lead-times, delivery schedules or minimum order quantities. Also, if several items are bought from one supplier and procured by a single consignment, fixed ordering costs may be shared among these items.

A simple example is constructed to illustrate the decision situation: Let us assume that item P1 can be purchased from two suppliers ($s = 1, 2$) with one

offering all-units and the other offering incremental discounts (Table 11.1). There are three purchasing intervals ($v = 1, 2, 3$) for each supplier s with prices $p_{v,s}$.

Table 11.1. Conditions for purchasing item P1 from two suppliers

Supplier s	discount	fixed cost						
			U_s	$p_{1,s}$	$Q_{1,s}$	$p_{2,s}$	$Q_{2,s}$	$p_{3,s}$
1	all-units		100	8.00	200	7.80	400	7.60
1	incremental		50	7.90	300	7.50	500	7.20
								$+\infty$
								1000

Some additional remarks are necessary regarding the demand series generated by the BOM explosion. Namely, we require a reasonable number of period demands covering a planning interval that allows for the exploitation of quantity discounts. Also, the first replenishment decision should not be influenced by the target inventory at the planning horizon (usually set to the safety stock level). A rough rule of thumb is a planning interval and thus a demand series covering five ordering decisions (also called times between orders).

To keep our example small, we will do with five periods. Here, the demands calculated (see Fig. 11.4) suffer from the effect of the lead-time offset, i. e. there are no demands at all in period five while for periods three and four secondary demands are missing resulting from future production of item C1. Hence, it is recommended to switch to demand forecasts (see Chap. 7) for periods with incomplete secondary demands (periods two to five in our example). Still, one should check whether existing secondary demands for these periods are in line with demand forecasts. Resulting demands are shown in Table 11.2.

Table 11.2. Expected demands for item P1 resulting from BOM explosion and Demand Planning

Source of demand	Demand/order period				
	1	2	3	4	5
BOM explosion	280	80	40	—	—
Demand forecast	—	280	240	240	280
Expected demands	280	280	240	240	280

The only data missing is the interest rate to be used for capital employed within the supply chain which is 2.5 % per period.

The optimized purchasing plan (Stadtler, 2004) shows that the first order should be placed in period 1 from the second supplier with an order quantity

of 800 units while the second order is placed with the first supplier in period four with an order quantity of 520 units (Table 11.3). The total cost within the planning interval comes to 10,333.25 [MU] (monetary units). Here, holding costs sum up to 201.25 [MU] (including interest on fixed ordering costs), fixed purchasing costs are 150 [MU] and variable purchasing cost are 9,982 [MU].

Table 11.3. Purchasing plan from two suppliers

Sourcing from supplier	order quantity per period from supplier				
	1	2	3	4	5
1	—	—	—	520	—
2	800	—	—	—	—

Some APS vendors provide a separate purchasing module for exploiting quantity discounts. This may be particularly appealing for commercial enterprises and for the procurement of MRO items in general. In the case that procurement decisions incur quantity discounts and resulting costs have a strong impact on the overall cost situation of a production unit, it may be advisable to declare respective items as “critical” and to include procurement decisions into the module Production Planning (assuming that corresponding cost functions can be modelled and solved there). If procurement decisions have to cover a longer planning horizon, one might even consider including these items at the Master Planning level.

In summary, the automation of the procurement process by means of an APS module can streamline the traditional, labour intensive tasks of procurement, especially in a B2B environment. Optimized procurement decisions can further reduce holding and total acquisition costs by exploiting quantity discounts and selecting suppliers in the best way possible. However, care must be taken that the decision problem at hand is represented adequately in an APS including the specific cost functions of suppliers and obeying conditions imposed by both parties.

References

- Drexel, A., Fleischmann, B., Günther, H.-O., Stadtler, H., Tempelmeier, H. (1994) *Konzeptionelle Grundlagen kapazitätsorientierter PPS-Systeme*, Zeitschrift für betriebswirtschaftliche Forschung, Vol. 46, 1022-1045
- Orlicky, J. A. (1975) *Material Requirements Planning*, New York
- Silver, E. A., Pyke, D. F., Peterson, R. (1998) *Inventory Management and Production Planning and Scheduling*, 3rd ed., New York et al.
- Stadtler, H. (2004) *A general quantity discount and supplier selection mixed integer programming model*, Schriften zur Quantitativen Betriebswirtschaftslehre (working paper TU Darmstadt), No. 2/2004

- Tempelmeier, H., Derstroff, M. (1996) *A lagrangean heuristic for multi-item multi-level constrained lotsizing with setup times*, Management Science, Vol. 42, 738-757
- Tempelmeier, H. (2002) *A simple heuristic for dynamic order sizing and supplier selection with time-varying demand*, Production and Operations Management, Vol. 11, No. 4, 499-515
- Tempelmeier, H. (2003) *Material-Logistik – Modelle und Algorithmen für die Produktionsplanung und -steuerung und das Supply Chain Management*, 5th ed., Berlin et al.
- Vollman, T. E., Berry, W. L., Whybark, D. C. (1997) *Manufacturing planning and control systems*, 4th ed., New York et al.

12 Distribution and Transport Planning

Bernhard Fleischmann

University of Augsburg, Department of Production and Logistics,
Universitätsstraße 16, 86135 Augsburg, Germany

12.1 Planning Situations

12.1.1 Transport Systems

Transport processes are essential parts of the supply chain. They perform the flow of materials that connects an enterprise with its suppliers and with its customers. The integrated view of transport, production and inventory holding processes is characteristic of the modern SCM concept.

The appropriate structure of a transport system mainly depends on the size of the single shipments: Large shipments can go directly from the source to the destination in full transport units, e. g. trucks or containers. Small shipments have to be consolidated in a transport network, where a single shipment is transshipped once or several times and the transport is broken at *transshipment points (TPs)*. A particularly effective consolidation of small shipments is achieved by a *logistics service provider (LSP)*, who can combine the transports from many senders.

The consolidation of transport flows decreases the transport cost. As the cost of a single trip of a certain vehicle on a certain route is nearly independent of the load, a high utilization of the loading capacity is advantageous. Moreover, the relative cost per loading capacity decreases with increasing size of the vehicles. But even with a strong consolidation of shipments to full loads, e. g. by an LSP, the smaller shipments cause relatively higher cost, because the consolidation requires detours to different loading places, additional stops and transshipment (see Fleischmann (1998, pp. 65)).

The following transport processes occur in a supply chain:

- The *supply of materials* from external suppliers or from an own remote factory to a production site. Both cases are identical from the viewpoint of logistics.
- The *distribution of products* from a factory to the customers. The distribution system depends on the type of products:
 - *Investment goods*, e. g. machines or equipment for industrial customers, are shipped only once or seldom on a certain transport link.
 - *Materials for production* are also shipped to industrial customers, but regularly and frequently on the same path.
 - *Consumer goods* are shipped to wholesalers or retailers, often in very small order sizes (with an average below 100 kg in some businesses), requiring a consolidation of the transports.

Note that the transport of materials from factory to factory is part of the distribution function of the supplier as well as part of the procurement function of the receiver. Transport planning is usually the responsibility of the supplier. But there are important exceptions, e.g. in the automotive industry, where the manufacturer controls the transports from his suppliers. In this case, transport planning occurs on the procurement side as well.

An LSP may consolidate the transport flows of several "*shippers*", operating in separate supply chains, in his own network. Then he is responsible for planning how the transports are executed, i.e. by which vehicles along which routes. However, the decisions on the transport orders, i.e. the quantity, source and destination of every shipment, remain a task of the APS of the shipper. Usually, it is not practicable to include the flows of all other shippers of an LSP into the APS. However, the additional flows have an impact on the transport cost and should be taken into account implicitly by appropriate transport cost functions.

Distribution Systems

A typical *distribution system* of a consumer goods manufacturer comprises the flow of many products from several factories to a large number of customers. Products made to stock are often shipped first to central DCs on forecast. The deliveries of the customer orders may then use the following distribution paths:

Shipments may go *directly* from the factory or from a DC to the customer, with a single order. This simplest form of distribution is only efficient for large orders using up the vehicle. Smaller orders can be shipped jointly *in tours* starting from the factory or DC and calling at several customers. A stronger bundling of small shipments is achieved by a joint transport from the DC to a TP and delivery in short distance tours from there. Figure 12.1 illustrates the different distribution paths.

The transport of materials for production, as far as controlled by the supplier, is mostly done in direct shipments. A recent concept for the supply of standard materials is the *vendor managed inventory* (VMI), where the supplier decides on time and quantity of the shipments to the customer but has to keep the stock in the customer's warehouse between agreed minimum and maximum levels. In this case, the customer's warehouse has the same function as a DC, so that the planning of VMI supply is similar to the DC replenishment.

Procurement Logistics Systems

If a manufacturer controls the transports of materials from his suppliers, he can use various logistics concepts, which differ in the structure of the transportation network and in the frequency of the shipments. They may occur in parallel for different classes of materials for the same receiving factory.

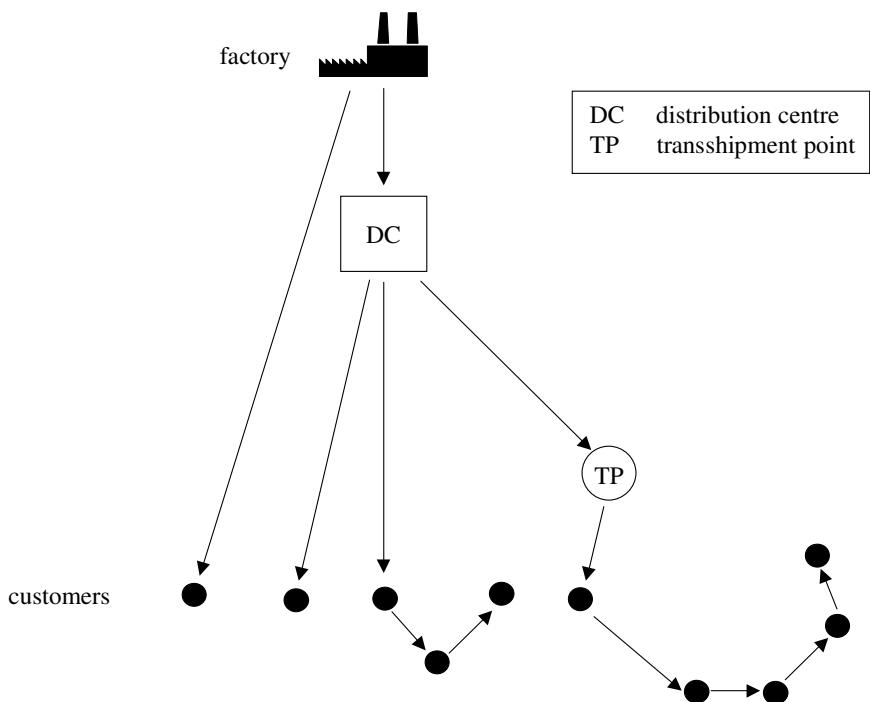


Fig. 12.1. Distribution paths

Cyclical procurement in intervals of a few days up to weeks permits to bundle the transport flow into larger shipments, but generates cycle stock at the receiving factory. *JIT procurement* with at least daily shipments avoids the inbound material passing through the warehouse. Instead, it can be put on a buffer area for a short time. If the arrivals are even *synchronized with the production sequence*, the material can be put immediately to the production line where it is consumed. The latter case is called *synchronized procurement* in the following.

The following transport concepts exist for procurement:

- *Direct transports* from the supplier are suitable for cyclical supply and, if the demand is sufficiently large, also for daily supply. Only if the distance is very short, direct transports may be used for synchronized procurement.
- A *regional LSP* collects the materials in tours from all suppliers in his defined area, consolidates them at a TP and ships them in full trucks to the receiving factory. This concept permits frequent supply, up to daily, even from remote suppliers with low volume. The trunk haulage can also be carried out by rail, if there are suitable connections.

- An *LSP warehouse* close to the receiving factory suits for synchronized procurement: The LSP is responsible for satisfying the short-term calls from the receiver by synchronized shipments. The suppliers have to keep the stock in the warehouse between agreed minimum and maximum levels by appropriate shipments, like in the VMI concept.

12.1.2 Interfaces to Other APS Modules

The common expression “Distribution and Transport Planning” denotes a set of various functions which overlap with other APS modules. In the proper sense, “Transport Planning” is the generic term, and it may occur on the procurement side as well as on the distribution side, as explained in the previous section. Moreover, it extends from the mid-term aggregate planning of transport processes, which is part of Master Planning, down to the shortest-term planning level: Planning deliveries of known customer orders is the last step of Demand Fulfilment and the release of orders for delivery from stock is part of the ATP function (see Chap. 9).

Distribution and Transport Planning is linked to the other modules by the following data flows:

Strategic Network Planning (see Chap. 6) provides the structure of the transport network, i.e.

- the locations of factories, suppliers, DCs and TPs,
- the transport modes and potential paths,
- the allocation of suppliers and customers to areas and of areas to factories, DCs, TPs and
- the use of LSPs.

Master Planning (see Chap. 8) determines

- aggregate quantities to be shipped on every transport link and
- the increase and decrease of seasonal stocks at the factory warehouses and the DCs,

where the first point can also be considered as part of mid-term transportation planning. The aggregate transport quantities should not serve as strict instructions to the short-term transport planning in order to keep the latter flexible. The main purpose of that quantity calculation is to provide appropriate resources and capacities and to take the duration of the various transport links into account. However, in case of multiple sources – e.g. if a material can be ordered from several suppliers or if a product is produced in several factories or if a customer can be supplied from several DCs – the aggregate quantities reflect the global view of Master Planning. Then they represent important guidelines for short-term transportation which could be used, for instance, as fractions of the demand sourced from different locations.

Also *Demand Planning* (see Chap. 7) provides essential data for transport planning:

- customer orders to be delivered,
- forecast of demand at the DCs and
- safety stocks at the DCs.

The relationship with *Production Scheduling* is twofold: On the one hand, Transport Planning may determine

- net requirements, timed at the planned departure of shipments from the factory, as input to Production Scheduling,

on the other hand, the latter module provides

- planned and released production orders as input to Transport Planning for the very short term decisions on the release of shipments.

12.1.3 Planning Tasks

As mentioned before, Distribution and Transport Planning comprises mid-term and short-term decisions, which are explained in the following.

Mid-Term Planning Tasks

The *frequency* of regular transports on the same relation is a key cost factor. It is a mid-term decision variable for the DC replenishment on the distribution side and for the supply of materials on the procurement side. The objective is to optimize the trade-off between transport cost and inventory (see Sect. 12.2.1). The resulting frequencies set target values for the short-term decisions on shipment quantities. Moreover, they determine the necessary transport lot-sizing inventory (see Sect. 2.4), which should be a component of the minimum stock level in Master Planning as well as in Production Planning and Scheduling.

The selection of *distribution paths* for the delivery of customer orders usually follows general *rules* fixed by mid-term decisions. They are mostly based on limits for the order size, e.g. orders up to 30 kg by a parcel service, up to 1000 kg from DC via a TP, up to 3000 kg directly from DC and larger orders directly from factory.

On the procurement side, the *assignment of material items to the supply concepts* - direct, via regional TP or via LSP warehouse - also has to be fixed on a mid-term basis. As explained in the previous section, these decisions are closely related with the supply frequencies.

The determination of *aggregate transport quantities* on every transport link in the supply chain is an essential mid-term planning task. As far as the distribution side is concerned, this is the “Distribution Planning” part of “Distribution and Transport Planning”. But this task should be integrated in the Master Planning in order to guarantee a close coordination of the production and transportation flows in the supply chain.

Short-Term Planning Tasks

Short term transport planning is usually carried out daily with a horizon of one day or a few days. This task, also called *Deployment*, consists of the following decisions:

The *quantities to be shipped* on the current day have to be determined, in the distribution system for the *replenishment* of every DC and VMI customer by every product, in the procurement system for the supply of every material. The shipment quantities can be influenced by the mid-term decisions on shipment frequencies and aggregate quantities.

The task of *vehicle loading* is to adjust the sum of the shipment quantities of the various items on the same transport link to a full vehicle load or a multiple thereof. It is relevant for DC replenishment and supply of materials, if the vehicle, as usual in these cases, is used exclusively for the concerned supply chain.

For the *deliveries to customers*, the quantity is fixed by the customer order, but there may be several *sources* from where to deliver and several *distribution paths*. These choices normally follow the guidelines set by the Master Planning quantities and by the general rules on the distribution paths, as explained above.

The deployment function for products made to stock is closely related to the *ATP function* (see Chap. 9): Customers expect orders to be delivered from stock within a short agreed lead-time, mostly between 24 and 72 hours, necessary for order picking, loading and transportation. If the incoming orders of the current day in total exceed the available stock of a certain item, the orders cannot be released according to the standard rules. Instead, some of the following measures have to be decided on:

- shipping some orders from an alternative source,
- substituting the item by an available product, if the customer accepts it,
- reducing the quantities for DC replenishment which are in competition with the customer orders to be shipped from factory and
- reducing some customer orders in size, delaying or canceling them: This most undesired decision is usually not completely avoidable. Even if it is only necessary for a very small percentage of all orders, the concerned orders must be selected carefully.

Vehicle scheduling comprises two different tasks:

- scheduling the short distance tours for delivering small orders from a TP in smaller vans and
- scheduling the trunk haulage from the factory to the DCs, from DCs to TPs and the direct delivery tours from a factory or a DC to customers.

These decisions can again be prestructured by longer term planning of fixed areas for the short distance tours and a regular line schedule for the circulation of trucks between factories, DCs and TPs. However, except for the case

where the vehicles are used exclusively for the supply chain under consideration, vehicle scheduling is typically the task of the LSPs in charge. As far as an LSP uses the vehicles for clients outside the considered supply chain – and this is a source of efficiency of the transport processes – vehicle scheduling cannot be integrated into advanced planning.

12.2 Models

12.2.1 Transport and Inventory

Transport planning has a strong impact on the inventory in the supply chain. It directly creates transport lot-sizing stock and transit stock (see Sect. 2.4) and influences the necessary safety stock. The lot-sizing stock results from the decision on the transport frequencies. Unfortunately, the present APS do not (yet) support the optimization of mid-term transport planning with regard to inventory. Nevertheless, this section presents some generic planning models, since the resulting frequencies and inventories are also important data for other APS modules. When setting these data, the following relationships should be taken into account. A review of combined transportation and inventory planning is given by Bertazzi and Speranza (1999).

Single Link, Single Product

The simplest case is a transportation process linking a production process of a certain product at location A with a consumption process at location B. Both production and demand are continuous with a steady rate. In this case, the optimal transportation scheme consists in regular shipments of the same quantity. Figure 12.2 depicts the cumulative curves of production, departure

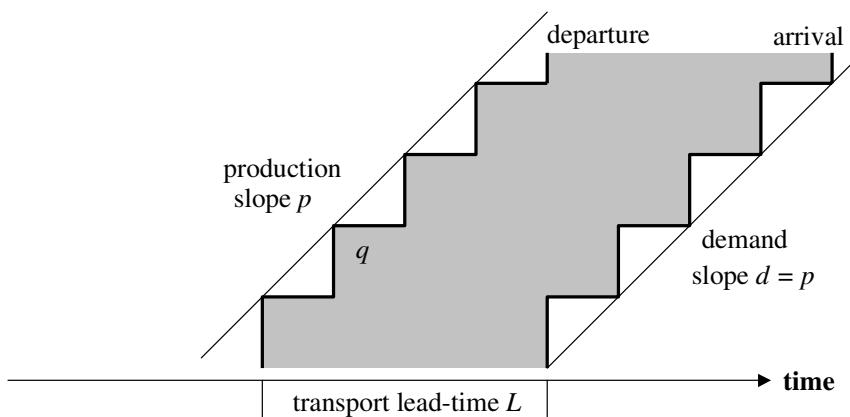


Fig. 12.2. Cumulative production, departure, arrival and demand

from A, arrival in B and consumption. The vertical distances between these curves represent the development of the stock in A, in transit and in B. With the notations

p	production rate (units per day)
$d = p$	demand rate
Q	maximum load per shipment
L	transport lead-time
t	cycle time
$q = d \cdot t$	shipment quantity
h	inventory holding cost (per unit and day)
$T(q)$	cost of a shipment of quantity $q \leq Q$

the following relationships are obvious: The average transit stock is $L \cdot d$. As it does not depend on the transport schedule, it can be neglected in transport planning, as long as the transportation time is fixed. Therefore, $L = 0$ can be assumed in the following. The total cost per day due to transportation is

$$hq + T(q)d/q. \quad (12.1)$$

As the transport cost usually shows economies of scale, i. e. $T(q)/q$ is decreasing with increasing q , there is a tradeoff between inventory and transport cost which can be optimized by the choice of q . If $T(q) = F$ is fixed for $0 < q \leq Q$, i. e. the shipment is exclusive for the quantity q , the optimal q is obtained from the usual EOQ formula (see Silver et al. (1998, Chap. 5.2)) with two modifications: The factor $\frac{1}{2}$ of the holding cost h is missing and q must not exceed Q , i. e.

$$q^* = \min(Q, \sqrt{Fd/h}). \quad (12.2)$$

However, in most cases, the transportation costs are dominant, so that the transport in full loads $q^* = Q$ is optimal.

It follows from Fig. 12.2 that Production Planning must consider the demand in B shifted by the time $L+q^*/d$ or, equivalently, guarantee a minimum stock of $Ld + q^*$.

Single Link, Several Products

Now, several products i are produced in A and consumed in B, each with a steady rate d_i and holding cost h_i . If the transport cost F per shipment is fixed again, it is optimal to ship always all products together, i. e. with a common cycle time t and quantities $q_i = d_i t$ (see Fleischmann (1999)). The optimal cycle time is

$$t^* = \min\left(Q / \sum_i d_i, \sqrt{F / \sum_i h_i d_i}\right). \quad (12.3)$$

Even if demand fluctuates, it is optimal, at a certain shipment, to ship all products with positive net demand in the following cycle. Rules for determining shipment quantities in this case are discussed in Sect. 12.2.2.

General Case

The above assumption of steady demand may be realistic in case of consumer goods, whereas the consumption of materials in production and the output from production mostly take place in lots. Blumenfeld et al. (1991) and Hall (1996) investigate the influence of production and transport scheduling on inventories in various supply networks and underline the difference between *independent* and *synchronized schedules*. Synchronization of transports and the consumption of materials is the basic idea of JIT procurement. Synchronization of production and distribution is the rule in a make-to-order or assemble-to-order situation. Production to stock is by its nature not synchronized with the shipments of customer orders.

But shipments from a factory to remote DCs or to VMI customers can be synchronized with production to stock. However, in case of many items produced cyclically on common lines and distributed to several destinations, the synchronization may become very difficult or impractical. Figure 12.3 depicts the cumulative production, transportation and demand curves for a

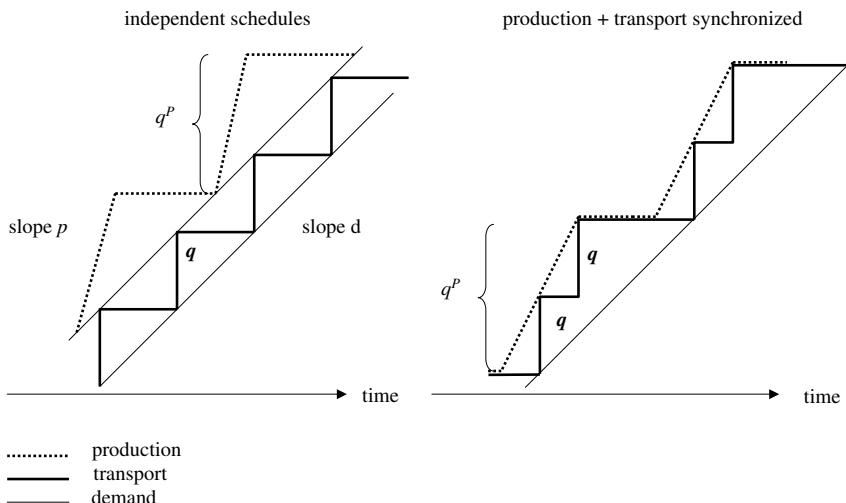


Fig. 12.3. Independent and synchronized schedules

single product and a single destination in case of independent and of synchronized schedules. In the latter case the production lot-size q^P is an integer multiple of the shipment quantity q . Note that the production rate p is now

greater than the demand rate d , because the production line has to produce other items in the intervals between the depicted lots. Obviously, synchronization reduces the average stock level which is in total (at the factory and at the DC)

$$I = \frac{1}{2}q^p(1 - d/p) + q \quad \text{for independent schedules} \quad (12.4)$$

$$I^S = \frac{1}{2}q^p(1 - d/p) + qd/p \quad \text{for synchronized schedules} \quad (12.5)$$

(see Blumenfeld et al. (1991)). But the difference is less than the shipment size q which is often small compared with the production lot-size q^p . Note that in the case of independent schedules production and transportation scheduling are decomposed by a demand line, which is left-shifted from the true demand line by the transport cycle time (plus the transit time which is not shown in Fig. 12.3). Production has to satisfy this demand line, whereas transportation planning assumes this line as continuous supply as in the single link cases considered above.

Transportation and Safety Stocks

In a distribution system for products made to stock, the safety stocks that are necessary for guaranteeing a certain service level, depend on the strategy of the transports between the factory and the DCs (see Silver et al. (1998, Chap. 12.4)): In a strong *push system* any production lot is distributed immediately to the DCs. A modification consists in retaining some central safety stock at the factory warehouse which is distributed in case of imminent stock-out at some DC. In a *pull system*, transports are triggered by the local stock at every DC, when it reaches a defined reorder point. In a push system, global information on the demand and stock situation at every DC is required for the central control. But also in a pull system, global information can improve the central allocation of stock in case of a bottleneck. In an APS, such global information should be available for the whole supply chain.

The push system corresponds to the case of synchronized production and distribution and thus requires less cycle stock, but in general higher total safety stock or more cross-shipments between the DCs. The local safety stock at a DC has to cover the local demand uncertainty during the transport lead-time, the total system safety stock has to cover the total demand uncertainty during the production lead-time and cycle time. In a consumer goods distribution system, the transport cycle time is usually very short, as a DC is usually replenished daily, but the production cycle time may last weeks to months, if many products share a production line. Therefore, the system safety stock calculation should be based on a periodic review model with the review period equal to the production cycle.

12.2.2 Deployment

The general task of deployment is to match the short-term demand with the available and expected stock for the next day or few days. As the *source locations* (factories, suppliers), where stock is available, are in general different from the demand locations (DCs, customers), it has to be decided how much to ship from which source location to which demand location.

A Network Flow Model

This task can be formulated as a network flow problem with the data

source locations S_i with available stock a_i ($i = 1, \dots, m$),

demand locations D_j with demand d_j ($j = 1, \dots, n$),

transport cost c_{ij} per unit from S_i to D_j ,

and the decision variables

shipment quantities x_{ij} from S_i to D_j

as follows:

$$\text{minimize } \sum_{i,j} c_{ij} \cdot x_{ij}, \text{ subject to}$$

$$\sum_j x_{ij} \leq a_i \quad \text{for every source location } S_i$$

$$\sum_i x_{ij} = d_j \quad \text{for every demand location } D_j$$

$$x_{ij} \geq 0 \quad \text{for all } i, j.$$

This is a special LP problem which can be extended to the case of several products and restricted transport capacity. It is in fact an extract from the Master Planning LP for the entire supply chain (see Chap. 8), restricted to transport processes and to a shorter horizon. It is therefore easy to integrate into an APS as it is offered by most APS suppliers. In the following we consider the more detailed release of single shipments. It can be supported by the above model in certain cases.

Delivering Known Customer Orders

In a make-to-order situation, the completion of the orders in due time is the responsibility of production planning and scheduling. Deployment can only deal with completed orders ready for delivery, and the shipment size is fixed by the customer order.

In a make-to-stock situation, many customer orders may compete for the same stock. If the stock at every source is sufficient for the normal allocation of orders, again, all order quantities can be released for delivery.

Otherwise, ATP decisions about measures against shortage have to be taken as explained in Sect. 12.1.3. If there are several sources with sufficient stock in total, reallocations can be made, either by transshipments from source to source or by directly reallocating certain customer orders from their normal source to an exceptional one. The latter measure is both faster and cheaper, in particular if customers are selected near the border between the delivery areas of the concerned sources. While this is difficult in conventional distribution systems with local control within the areas, it is no problem in an APS with global information and central control of deployment.

The optimal combination of the measures against shortage for all customers competing for the stock of a certain product can be determined with the above network flow model, with the following interpretation:

- Every customer j is modeled as a demand location.
- Besides real locations with available stock, the source “locations” i include other potential measures, in particular a “source” with unlimited availability that stands for reducing or canceling orders.
- The cost c_{ij} includes penalties for delaying, reducing or canceling a customer order, depending on the priority of the customer.

Replenishment of DCs and Procurement

Shipment quantities for replenishment and procurement are not determined by customer orders but have to be derived from Demand Planning. Moreover, the calculation requires the prior specification of a certain *transport cycle time* (or of the transport frequency) for every relation, as explained in Sect. 12.2.1. The *net demand* for a shipment is then

$$\begin{aligned} d^N = & \text{ demand forecast at the destination} \\ & \text{ during the following transport cycle and the transport lead-time} \\ & + \text{ safety stock for the destination} \\ & ./. \text{ available stock at the destination.} \end{aligned}$$

In a *pull system* the shipment quantity is set equal to d^N , if there is sufficient stock at the source *for all destinations*. The quantities may be modified by a vehicle loading procedure, as explained below. If the stock at the source is not sufficient, it is allocated to the destinations using a “*Fair Shares*” rule which takes into account the demand and stock situation of every destination and therefore requires global information and central control (see Silver et al. (1998, Chap. 12.4.3)). The basic idea of fair shares is to balance the stock at various demand locations so that the expected service level until the arrival of a new supply at the source (e. g. by a production lot) is equal at all locations. If the local stocks are included into the allocation procedure, it may result that, for some destination, the allocation is lower than the available stock, indicating that stock has to be transferred by lateral transshipments.

Distribution Requirements Planning (DRP) (see Silver et al. (1998, Chapter 15.6)) can be used to propagate the net demand upstream in a network, if

every node is supplied by a fixed single source. It is an extension of the MRP demand calculation to the distribution network and permits, like MRP, to consider time-phased dynamic demands and lead times from node to node.

In a *push distribution*, every supply arriving in the source is immediately distributed to the destinations according to fair shares. In case of short transport lead-times and long supply cycles for the source, it is advantageous to retain some central safety stock at the source which is distributed later according to updated fair shares.

In the case of shortage, the determination of the DC replenishment quantities can also be integrated in the network flow model, together with the deliveries of customer orders, where a DC appears as demand location with the above net demand.

Vehicle Loading

The previous calculations of shipment quantities are carried out separately for every product. They do not consider joint shipments of many products in appropriate transport units (e.g. whole pallets). This is the task of vehicle loading which starts from those shipment quantities and fits them to the vehicle capacity. As far as the quantities represent net demand, they can only be increased, but in general, the demand calculation can specify minimum quantities below the proposed quantities. An upper bound is given by the stock which is ready for shipment. Vehicle loading comprises the following steps:

- round up or down the shipment quantity of every product to whole transport units (e.g. pallets),
- adjust the size of the joint shipment, i.e. the sum of the single product quantities, to a full vehicle capacity, where the vehicle is eventually selected from a given fleet.

Both steps have to consider the minimum quantities and the available stock, the second step should try, within these bounds, to balance the percentages of increase (or decrease) over the products.

Vehicle Scheduling

As explained in the previous section, vehicle scheduling has only a limited importance for advanced planning. Therefore, and in view of the huge body of literature, models and algorithms for vehicle scheduling, this subject is not dealt with here. Instead, the reader is referred to the following review articles. Most literature concerns scheduling round trips of vehicles starting and ending at a single depot. This case is relevant for delivering small orders to customers from a TP and for collecting small orders for materials from suppliers by a regional LSP. A recent survey is Laporte (1998).

Vehicle scheduling for trunk haulage, as it occurs on the relations factory – DC, DC – TP, for direct deliveries to customers and in procurement transports, has been investigated only recently (see Stumpf (1998)).

12.2.3 APS Modules

There is no standard structure of the APS modules for Distribution and Transport Planning. In any APS, these tasks are covered by several modules or by multi-functional modules, but with different allocations within the SCP-matrix. In the following, essential features of these modules are explained regarding the planning tasks of Sect. 12.1.3. This Section is based on information from i2 Technologies (2004), J.D. Edwards (2001) and Knolmayer et al. (2002).

Mid-term Planning

The optimization of *transport frequencies* (see Sect. 12.2.1) w.r.t. transport and inventory cost is not supported. The same is true for establishing *rules on the use of distribution paths* and for *assigning materials to supply concepts*. However, the effect of such tactical decisions can be studied by means of analytical modules like the i2 Transportation Modeler.

The integration of *Distribution Planning* in the Master Planning function is standard in all the above APS. Thus, using the LP solver or heuristic algorithms of Master Planning, aggregate quantities can be determined for every transport link in the supply network.

Short-term Planning

For the short-term *deployment*, the APS provide the same modules as for Distribution Planning, used with a shorter horizon and more detailed demand information. Alternatively, there are special heuristics for calculating deployment quantities following a push or pull strategy, but restricted to the case, where every order has a specified single source. They work in two steps: First, a DRP calculation is performed upstream, starting from the net demand at the demand locations. If the available stock is not sufficient at some location, then fair share rules are applied downstream in a second step. The fair share rules are rather simple, e.g. the inventory is distributed proportionally to the demand or such that the same proportion of the target stock level at every location results (SAP APO). They do not consider service levels. The DRP calculation may differentiate several types of demand: customer orders, forecasts, safety stock replenishment and pre-built stock. Then, the allocation of tight inventory proceeds in this order and fair shares are only applied within one type of demand.

The Production & Distribution Planning Module of the Peoplesoft APS (former J.D. Edwards) incorporates a special submodule called “Connect

Algorithm” for the allocation of insufficient inventory. It runs after the normal deployment algorithm (LP or heuristics) and allocates the resulting inventory to the customer orders and forecast. It considers multiple sources and tracks the effects of reallocations along the supply network.

Vehicle loading is supported in all the above APS by particular modules or submodules running after the deployment. For every shipment they perform

- rounding procedures to multiples of transport units for single items, and
- building vehicle loads containing several items.

At least the first step considers the effects of quantity changes on the planned inventories. In the Peoplesoft APS (former J.D. Edwards) the two functions are split: The first one runs as “Rounding Engine” after the Deployment and before the reallocation, the second one is done by a separate Vehicle Loading module.

The modules Transportation Planning and Vehicle Scheduling of SAP APO as well as Transportation Modeler and Transportation Manager of the i2 APS perform a detailed planning of the single shipments and aim at an efficient consolidation of the shipments and an optimal use of the vehicles. They adopt primarily the view of an LSP:

- Input data are shipments (customer orders) with given quantity, origin and destination.
- The paths of shipments through consolidation points (TPs and hubs) and the routes of the vehicles are planned.
- The use of various carriers is controlled.
- Various transport tariffs of the carriers and for billing the customers can be considered.

In contrast, a manufacturer deals with customer orders that specify only quantity and destination, but leave the source location open. For DC replenishment shipments, even the quantity is open. If the transports are outsourced to one or several carriers, then it is usually an LSP who is responsible for the above tasks. However, for the collaboration between an LSP, who works in several supply chains, and the manufacturers in these supply chains, these modules are certainly useful.

References

- Bertazzi, L.; Speranza, M.G. (1999) *Models and algorithms for the minimization of inventory and transportation costs: A survey*, in: Speranza, M.G.; Stähly, P. (Eds.) New trends in distribution logistics, Berlin et al., 137–157
- Blumenfeld, D.E.; Burns, L.D.; Daganzo, C.F. (1991) *Synchronizing production and transportation schedules*, in: Transportation Research B, Vol. 25B, No. 1, 23–37
- Fleischmann, B. (1998) *Design of freight traffic networks*, in: Fleischmann, B.; Nunen, J.A.E.E. van; Speranza, M.G.; Stähly, P. (Eds.) Advances in distribution logistics, Berlin et al., 55–81

- Fleischmann, B. (1999) *Transport and inventory planning with discrete shipment times*, in: Speranza, M.G.; Stähly, P. (Eds.) New trends in distribution logistics, Berlin et al., 159–178
- Hall, R. (1996) *On the integration of production and distribution: economic order and production quantity implications*, in: Transportation Research B, Vol. 30, No. 5, 87–403
- i2 Technologies Inc. (2004) *Homepage*, URL: <http://www.i2.com>, Date: July 19, 2004
- J. D. Edwards (2001) *Production & Distribution Planning 4.0, User's Guide*, Denver
- Knolmayer, G.; Mertens, P.; Zeier, A. (2002) *Supply chain management based on SAP systems*, New York et al.
- Laporte, G. (1998) *Recent advances in routing algorithms*, in: Labbé, M.; Laporte, G.; Tanczos, K.; Toint, P. (Eds.) Operations research and decision aid methodologies in traffic and transportation management, Berlin et al.
- Silver, E.A.; Pyke, D.F.; Peterson, R. (1998) *Inventory management and production planning and scheduling*, 3rd ed., New York et al.
- Stumpf, P. (1998) *Vehicle routing and scheduling for trunk haulage*, in: Fleischmann, B.; Nunen, J.A.E.E. van; Speranza, M.G.; Stähly, P. (Eds.) Advances in distribution logistics, Berlin et al., 341–371

13 Coordination and Integration

Jens Rohde

Darmstadt University of Technology, Department of Production & Supply Chain Management, Hochschulstraße 1, 64289 Darmstadt, Germany

A strong *coordination* (i. e. the configuration of data flows and the division of planning tasks to modules) of APS modules is a prerequisite to achieve consistent plans for the different planning levels and for each entity of the supply chain. The same data should be used for each de-centralized planning task and decision. APS can be seen as “add-ons” to existing ERP systems with the focus on planning tasks and not on transactional tasks. In most cases an ERP system will be a kind of “leading system” where the main transactional data are kept and maintained. The data basis of APS is incrementally updated and major changes on master data are made in the ERP system. This task will be called *integration* of APS with ERP systems.

The coordination between the different planning modules described in Part II of this book is very important to derive dovetailed detailed plans for each supply chain entity. Section 13.1 will show which guidelines are given, which data are shared and how feedback is organized. Furthermore, one can see which modules are normally used centrally and de-centrally, respectively.

As we have already seen in Chap. 5, some decisions and tasks are left to the ERP system. These tasks and data which are used by APS but are kept in ERP systems are described in Sect. 13.2. The definition of the interface between ERP and APS has to determine which ERP data are used in APS and which data are returned. Moreover, *Data Warehouses* which keep important historical data and are mainly used by Demand Planning build interfaces to APS (see also Chap. 7).

A detailed knowledge of the status of supply chain operations and the occurrence of events within the supply chain is getting more and more important. Thus, the concept of Supply Chain Event Management to effectively manage the different categories of events occurring in a supply chain is discussed in Sect. 13.3.

Modules of APS that support collaboration of supply chain entities as well as external customers and suppliers are part of Chap. 14 and will not be discussed in this chapter.

13.1 Coordination of APS Modules

A general structure for coordination of the different modules cannot be suggested. There are several architectures that range between individual planning modules, which can be used as stand alone systems, and fully integrated

planning systems. A fully integrated system regularly has the advantage of an identical look-and-feel for all modules and accessibility to all modules by a single user interface. Furthermore, a single database provides data needed by every module and avoids redundancies and inconsistencies in the planning data caused by multiple databases. Different modules can interact via sending messages and exchanging data directly. In contrast, individual modules mostly do not have an identical look-and-feel and regularly no common data basis. An advantage of this architecture is that modules can easily be combined and chosen (if not all modules are needed) for a specific line of business. Most APS providers with such architectures provide special integration modules that enable controlled data and information exchange within the system (see also Chap. 18). Furthermore, an Alert Monitor is often responsible for the handling of alert situations from different APS modules in one central module.

The following paragraphs describe which guidelines are given and how feedback is organized to generate the different plans for a supply chain as a whole. Figure 13.1 gives a general view of the main interactions. The data flows are exemplified, as they can be different from one supply chain to another (see Chaps. 3 and 4). The main feedback is derived by periodic updates of plans while considering current data. Chapters 6–12 illustrate the interactions between APS modules in more detail.

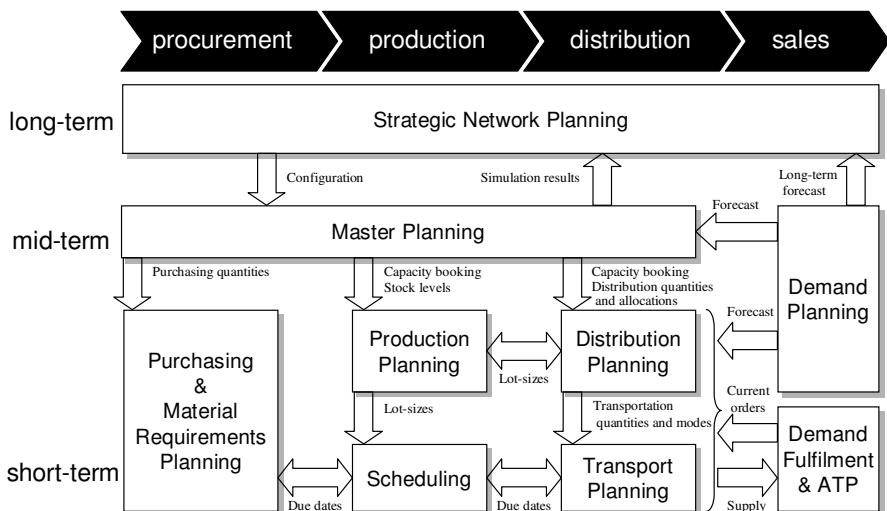


Fig. 13.1. Coordination and data flows of APS modules

Strategic Network Planning Strategic Network Planning determines the configuration of the supply chain. This configuration consists mainly of loca-

tions for each supply chain entity and possible distribution channels. Long-term Demand Planning gives input about trends in future demand. Simulated master plans can provide useful hints for capacity enhancements. However, the strategic goals of a supply chain (i.e. market position, expanding into new regions and markets, etc.) specify the framework for this module.

Demand Planning Demand Planning provides demand data for mid-term Master Planning as well as for short-term Production and Distribution Planning. The forecast for end products of a supply chain is input for Master Planning. The short-term planning modules use current, more accurate short-term forecasts from Demand Planning. Furthermore, de-centralized Demand Planning modules provide demand data for products not planned in Master Planning (e.g. non-critical components).

Master Planning Master Planning determines a production, distribution and purchasing plan for the supply chain as a whole with given demand from Demand Planning on an aggregated level. Therefore, this task should be done centrally. The results provide purchasing guidelines for the de-central Purchasing and Material Requirements Planning – like purchasing quantities from external suppliers, production guidelines for the de-central Production Planning and Scheduling – like capacity booking for potential bottlenecks and stock levels at the end of each period and distribution guidelines – like distribution channel chosen, and distribution quantities for de-central Distribution and Transport Planning. Feedback from short-term modules is derived by current stock levels, updated forecasts and current capacity usage. The average realization of given guidelines from short-term modules should be used for model adjustment in Master Planning.

Demand Fulfilment and ATP For Demand Fulfilment and ATP demand data from Demand Planning, production quantities for disaggregated products and intermediates, due dates from Production Planning and Scheduling, distribution plans and detailed vehicle routes from Distribution and Transport Planning and purchasing due dates as well as selected suppliers from Purchasing and Material Requirements Planning are used. Furthermore, current inventory levels at each production and distribution stage are needed as input. To be able to influence production and distribution plans, unused capacity bookings have to be known, too.

Production Planning and Scheduling The main guidelines from Master Planning are capacity bookings and stock levels for each period for the de-central units. Production Planning and Scheduling requires detailed, disaggregated information. Furthermore, current (short-term) forecasts and availability of production resources update the guidelines from Master Planning.

Lot-sizes and due dates from this module are exchanged with Distribution and Transport Planning to coordinate production and transportation lot-sizes as well as with Purchasing and Material Requirements Planning to coordinate purchasing lot-sizes and due dates in a more detailed way than it is done by Master Planning.

Purchasing and Material Requirements Planning Purchasing quantities derived from Master Planning provide valuable input for mid-term supplier contracts and supplier selection. Based on these quantities discounts can be negotiated. Considering short-term production due-dates and lot-sizes as well as mid-term contracts for critical components feasible purchasing plans are obtained. Purchasing plans have to be aligned with production schedules to secure an adequate and timely supply of materials.

Distribution and Transport Planning The coordination of Distribution and Transport Planning is similar to Production Planning and Scheduling. The short-term coordination by lot-sizes and due dates enables accurate production-distribution plans. The actual production quantities provide main input for the transport plans. Furthermore, time windows from customer orders are additional constraints for building and routing vehicle loads.

Alert Monitor The alert monitor depicts the concept of *management-by-exception*. Management-by-exception is a technique to control guidelines. It differentiates between *normal cases* and *exceptional cases*. Here, the decision whether a situation is an exceptional case or not is delegated to the APS. The prerequisites for this concept are detailed information about tolerances for normal cases, exact definitions for reporting and delegation of decisions (along the lines of e.g. Silver et al. (1998) and Ulrich and Fluri (1995)).

The APS raises alerts if problems or infeasibilities occur (see Fig. 13.2). To pass the right alerts to the right organizational units within a supply chain, it is necessary to filter these alerts first. Afterwards, filtered alerts are sent to the responsible organizational unit of a supply chain entity. Specifying these responsibilities is part of an implementation project (see Chap. 17). Finally, these alerts have to be sent physically, e.g. by e-mail or an Internet based application.

The responsible units react on alerts by generating new plans, moving orders, using reserved teams, etc. The new or adjusted plans are then sent back to the APS. The APS has to process the changes made and propagate them to each unit affected by the changes.

13.2 Integration of APS

To use an APS effectively, it has to be integrated in an existing IT infrastructure (see Fig. 13.3). The main interactions exist between APS and *online*

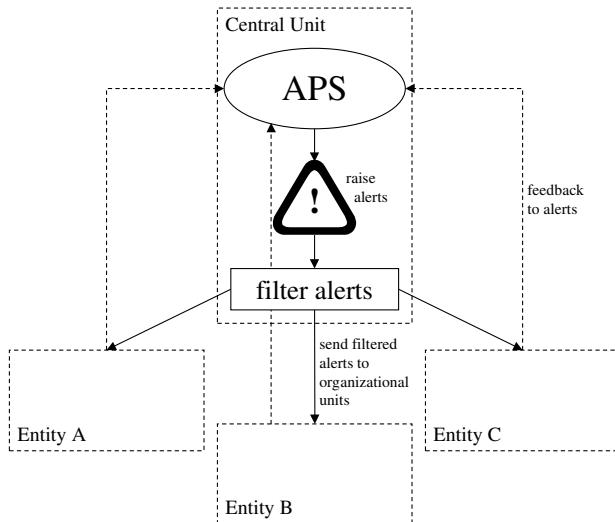


Fig. 13.2. Alert monitor

transaction processing (OLTP) systems, e.g. ERP and legacy systems. Another important system – especially for the demand planning task – is a *Data Warehouse* (DW). This “warehouse” stores major historical data of a business and supply chain, respectively. The next subsection will illustrate the integration of APS with OLTP systems. Afterwards, the integration with Data Warehouses is described. The integration of OLTP and Data Warehouses will not be subject of this book.

New middleware technology, subsumed as *Enterprise Application Integration* systems, provides a platform for integration of various tools and databases. The last subsection of this chapter will give a brief insight.

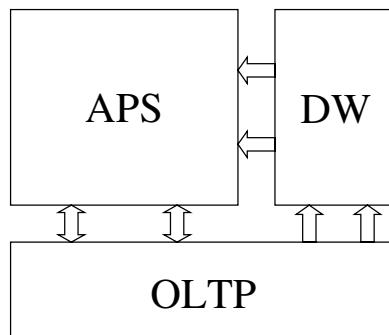


Fig. 13.3. APS integration

13.2.1 Integration with OLTP

An APS does not replace an existing ERP or legacy system. On the contrary, APS extend them by additional planning functionality. Transactional data are kept and maintained in the OLTP system. An APS is only responsible for its specific data, more exactly, all data that are needed but are not part of the OLTP system's data basis.

As Fig. 13.4 shows, an APS regularly communicates with several OLTP systems of different supply chain entities. Furthermore, planning tasks like BOM explosion for non-critical materials and ordering of materials are mostly left to ERP systems (see Chaps. 4 and 5). The *integration model* defines which objects are exchanged, where they come from and which planning tasks are performed on which system. The *data exchange model* specifies how the flow of data and information between the systems is organized.

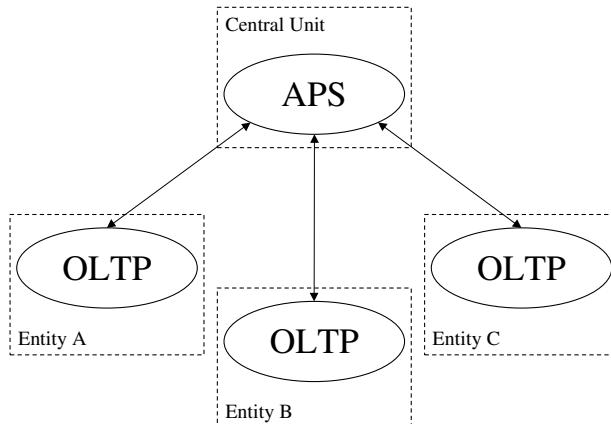


Fig. 13.4. Integration of several OLTP systems

Most APS provide a *macro-language* to define these models and enable an automatic exchange of data. While OLTP systems are regularly older systems, the adjustment has to be done by APS. That is, an APS has to be able to match data items from the OLTP system to its implicit structure and to handle different import and export formats, because it is mostly not possible to do all adjustments needed on the OLTP system. Also, it must be possible to maintain specific data like penalty costs and aggregation rules (see Chap. 8) within the APS.

Integration Model

Within the integration model, objects which are exchanged between OLTP and APS are defined. If, like it is done in most cases, not all products are

planned in APS, the integration model has to define which products and materials are critical. Also, the potential bottleneck resources have to be defined. These objects are e.g.:

- BOMs,
- routings,
- inventory levels and
- customer orders.

Furthermore, the integration model has to define which data are exchanged with which OLTP system. A supply chain consists of several entities with local systems. The right data have to be exchanged with the right system. This assignment can be done by modelling different sites with their flows of materials in the Master Planning module (see Chap. 8).

The integration model also defines which results are returned to the OLTP systems and the planning tasks done by an APS or ERP system (e.g. performing BOM explosion of non-critical components in local ERP systems). By defining several integration models it is possible to simulate different alternatives of a division of labour between APS and ERP system.

Data Exchange Model

Data *which* have to be exchanged between OLTP and APS are mainly defined by the integration model. The data exchange model defines *how* these data are exchanged. The data transfer between OLTP and APS is executed in two steps (see Fig. 13.5).

The first step is the *initial data transfer*. During this step data needed for building the Master Planning, Production Planning and Scheduling and Distribution and Transport Planning models are transferred from OLTP to APS (e.g. the BOM and routings of critical products, properties of potential bottleneck resources, regular capacities, etc.). After the models are generated “automatically”, it is necessary to maintain the APS specific data like optimizer profiles, penalty costs and aggregation rules.

In the second step, data are transferred *incrementally* between the systems. The OLTP system should only transfer changes that have been made on the data to the APS and vice versa (*netchange*). The data exchanged are divided into *master data* and *transactional data*. Changes on master data require a model adjustment in the APS. For example, this could be the purchase of a new production resource or the long-term introduction of a second or third shift. Transactional data are transferred to and from APS as a result of planning tasks. For example, the following transactional data are sent incrementally to an APS:

- current inventories,
- current orders,

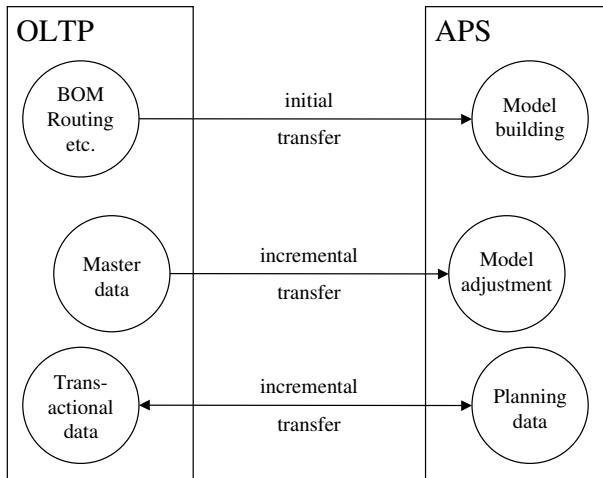


Fig. 13.5. Transferring data between OLTP and APS

- availability of resources and
- planned production quantities and stock levels, respectively.

Current inventories are needed for every APS module. For Master Planning they can be regarded as a feedback in an incremental planning process, the ATP module uses this data to perform online-promises, while in Distribution Planning it is necessary to calculate the actual distribution quantities, etc. Current orders are used to match planned orders. Those planned orders are a result of the Production Planning and Scheduling module if this planning task is performed on forecasts. All short-term planning modules have to consider the availability of resources like machines, vehicles, etc. to generate feasible plans. Planned production quantities and stock levels are for instance needed to perform BOM explosions for non-critical components in local ERP systems.

However, the separated data basis for APS modules poses problems of redundancies and inconsistencies. These problems have to be controlled by the data exchange model. Even though redundancy of data enables the APS to simulate different plans without affecting OLTP systems, it is very difficult to ensure that all systems have the correct data. Changes in each system have to be propagated to avoid an inconsistent data basis. That is, every modification has to be recorded and sent to the relevant systems. If too many changes are made, too many data are transferred between systems and data updates paralyze the APS. The trade-off between 100%-consistency and paralyzation of the APS has to be considered during the implementation process (see Chap. 17). That is to say, it is possible to perform updates in predefined time intervals to avoid trashing by data transfer but with a reduction in consistency.

13.2.2 Integration with Data Warehouses

While OLTP systems depict the current state of a supply chain entity a *Data Warehouse* is its “memory”. Nearly all data are available – but not information. The goal of a Data Warehouse is to provide *the right information at the right time*. The Data Warehouse has to bring together disparate data from throughout an organization or supply chain for decision support purposes (Berry and Linoff, 1997, p. 360).

The terms *knowledge discovery in databases* (KDD) and *data mining* keep arising in combination with Data Warehouses. The term KDD is proposed to be employed to describe the whole process of extraction of knowledge from data. Data mining should be used exclusively in the discovery stage of this process (Adriaans and Zantinge, 1996, p. 5).

The interaction between APS and the Data Warehouse is a read-only process – the Data Warehouse is updated incrementally by transactional data from OLTP systems. The main use of Data Warehouses is in Demand Planning which uses historical data regularly to find patterns in demand and sales data and to analyze those time-series with statistical tools (see Chaps. 7 and 26). KDD, especially data mining, provides important input for every step in model building for all modules of an APS. While data mining tools have the focus in *finding patterns* in data, in contrast, *online analytical processing* (OLAP) tools are fast and powerful tools for *reporting* on data. OLAP tools provide a fast way for APS to access data of the Data Warehouse. The conventional way by queries (esp. SQL) is also possible to access data (see Fig. 13.6).

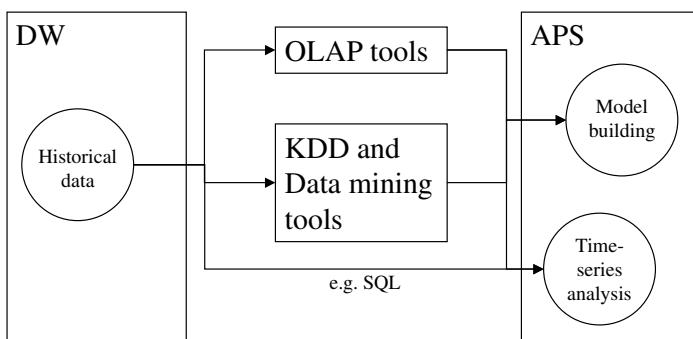


Fig. 13.6. Integration of Data Warehouse

SCM is a new challenge for the design of Data Warehouses. Not only have the data of a single company to be collected, but also supply chain wide transactional data have to be brought together in a consistent way to enable decision support for the supply chain as a whole.

13.2.3 Enterprise Application Integration

The growing number of different systems within a single organization makes point-to-point integration no longer applicable. Integrating the various systems of an organization or even the entire supply chain, called *Enterprise Application Integration* (EA), is a challenging task that could not be performed without powerful middleware systems. According to the task these systems have to perform they are called EAI systems. The goal of those systems is the decoupling of applications. Figure 13.7 visualizes the difference between point-to-point integration and decoupled applications.

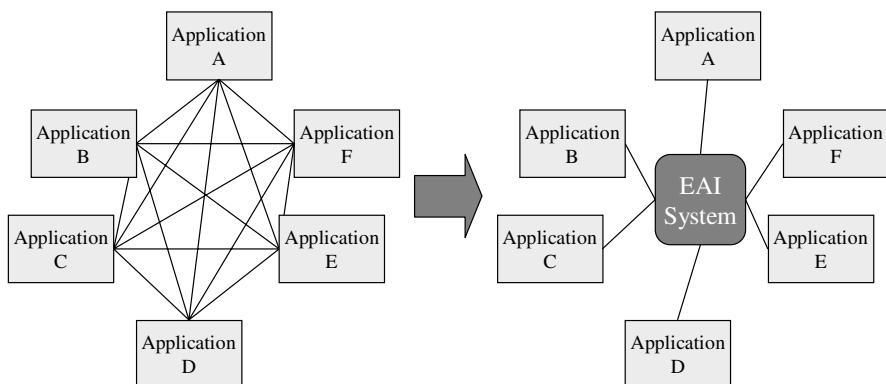


Fig. 13.7. Point-to-point integrated vs. decoupled applications

Independent of the underlying software component architecture like *Enterprise Java Beans* (EJB), *CORBA* or Microsoft *DCOM*, an architecture for EAI has to be identified. Such an identification provides essential input for software selection and implementation. Lutz (2000) distinguishes the following five EAI architecture patterns:

Integration adapter Via this architecture an existing (server) application interface is converted into the desired interface of one or more clients. The client application will then exclusively invoke services of the server application through the interface of the integration adapter. An interface is dedicated to a single server application, but multiple clients can access the server application by using a common integration adapter. Changes in the server application no longer reflect each (accessing) client; only the adapter has to be adjusted. The integration adapter does not provide any logic. Solely, a mapping of the server API (application programming interface) to the API provided by the adapter is performed. Usually, the integration adapter does not know of the existence of clients and the server application does not know about the existence of the integration adapter unless the server application needs adjustment to be able to provide desired services.

Integration messenger Communication dependencies between applications are minimized by this approach. Here, the application interaction logic is not decoupled. The integration messenger delivers messages between applications and provides location transparency services, i. e. distributed applications do not have to know about the actual location of each other to be able to communicate. The design of the integration logic is still left to the applications. One example for an integration logic is *remote method invocation* where an application is able to perform “direct” method calls on another one. In this case both applications have to provide the required services and public interfaces for remote method invocation. A change on the integration logic in one application still affects the other ones while application location changes only concern the integration messenger.

Integration facade The interface to several server applications is simplified by providing an integration facade. Server functionality is abstracted to make the back-end applications easier to use. The integration facade has to perform the mapping of its own interface to the interfaces of the server application. Furthermore, internal logic has to be provided to enable the abstraction. For example, an ATP request (see Chap. 9) invokes services on various systems to get information about product availability. The different system are e. g. inventory management systems of different distribution centres. Without an integration facade the client has to invoke these services on each server application. Whereas the integration facade is aware of the different systems. It provides a thin interface to the client(s) and takes over the invocation of the services on different systems as well as processing the information. Server applications are unaware of the existence of the integration facade while the integration facade itself does not know of the existence of clients.

Integration mediator This architecture pattern encapsulates interaction logic of applications and decouples this logic from the participating applications. In contrast to the integration facade the participating applications are aware of the existence of the integration mediator. No direct communication between these applications is permitted. Each interaction has to invoke the integration mediator. Via this pattern dependencies between applications are minimized and maintenance is simplified owing to the centralized interaction logic. The interaction logic encapsulated by the integration mediator includes message content transformation (e. g. mapping of different product IDs) and controlling message destination(s). In stateless scenarios this logic is only dependent on the current content of a message. In contrast, stateful scenarios are additionally dependent on previous application interactions (e. g. accumulation of events). Stateful integration mediators are much more complex to handle as they usually need state management and state persistence to span shutdown situations, for example.

Process automator The goal of this architecture is to minimize dependencies between process automation logic and applications. It automates sequencing of activities in a process. The process automator pattern consists of a process controller, activity services and applications providing desired services. The sequencing activity logic of a process is implemented by the process controller. The activity services abstract from the applications and provide request-based services to the process controller, i.e. all system interactions are hidden. The activity service is a specialty of the integration facade pattern that abstracts interactions to the level of an activity. By providing such a simplified and uniform interface to the service applications, the process controller is decoupled from the special APIs of the services. The application integration logic is encapsulated.

The different architecture patterns can be combined. For example, integration adapter and integration messenger can be combined in such a way that the integration adapter provides the interfaces that the integration messenger is expecting. This architecture decouples server APIs and the application itself from the API of an integration messenger.

13.3 Supply Chain Event Management

The task of *Supply Chain Event Management* is to *manage* planned and unplanned events in a supply chain. The effectiveness of the supply chain is to be improved while reducing costs by handling events. Managing events does not only mean to react to events, but also to affect or even prevent their occurrence. Following Otto (2004), SCEM can be characterized as a management concept, a software solution and a software component. Here, we will focus on SCEM as a management concept.

Table 13.1. SCEM Definitions (Alvarenga and Schoenthaler, 2003)

Supply chain event	Any individual outcome (or nonoutcome) of a supply chain cycle, (sub) process, activity, or task.
Supply Chain Event Management	The application of statistical, process, and technology identification and control solutions to standard and nonstandard supply chain events.
Event category	A logical grouping of supply chain events.
Event probability index (EPI)	The statistical measure, on a scale of 0 to 1, of the tendency of an event to occur within the supply chain [...] over a given time interval.
Standard event	An event that tends to occur within the supply chain; that is, has an EPI of .50x or higher.
Nonstandard event	An event that tends not to occur within the supply chain; that is, has an EPI of .49x or lower.
Event management plan (EMP)	A documented process that outlines the steps taken to control or react to an event.
Planned event	An event for which a documented EMP exists.
Unplanned event	An event for which a documented EMP does not exist.

To understand SCEM Alvarenga and Schoenthaler (2003) define the terminology given in Table 13.1. A supply chain event can occur on all levels of detail in the supply chain (from broader cycles to detailed tasks). To be able to manage these events efficiently a reasonable grouping into event categories is inevitable. By giving the probability of an event to occur events can be classified into standard and nonstandard events, assuming that nonstandard events are generally more costly to manage. A documented reaction to supply chain events avoids ad hoc decisions under pressure of time. Planned events are generally less costly to manage.

A reduction in cost by SCEM can be achieved in two ways. Firstly, the reaction to an event can be accomplished more efficiently. Secondly, more costly events can be shifted to less costly events by defining EMPs for so far unplanned events or by eliminating events following the idea of continuous improvement.

When changing the category of events from unplanned to planned the trade-off between costs for defining an EMP and the probability of occurrence and the resulting ad hoc decision should be taken into account. Furthermore, it should be inspected whether external influences like material shortages can be affected, e. g. by using the concept of vendor managed inventory. Thus, probably shifting an event from a standard to a nonstandard event makes a single occurrence more costly, but in sum less costly due to less occurrences.

A software based approach to SCEM has to provide online access to pre-defined supply chain events. Furthermore, events have to be categorized and management of event shifts has to be provided. By making categorized events and actions to be taken in case of occurrence accessible throughout the complete supply chain a valuable *supply chain visibility* is created. Here, an Alert Monitor offers an important platform for identification of events and notification of their respective owners.

References

- Adriaans, P.; Zantinge, D. (1996) *Data mining*, Harlow et al.
- Alvarenga, C. A.; Schoenthaler, R. C. (2003) *A new take on Supply Chain Event Management*, Supply Chain Management Review, March/April 2003, 28–35
- Berry, M. J. A.; Linoff, G. (1997) *Data mining techniques: For marketing, sales, and customer support*, New York et al.
- Lutz, J. C. (2000) *EAI architecture patterns*, EAI Journal, March 2000, 64–73
- Otto, A. (2004) *Supply Chain Event Management: Three perspectives*, International Journal of Logistics Management, Vol. 14, No. 2, 1–13
- Silver, E. A.; Pyke, D. F.; Peterson, R. (1998) *Inventory management and production planning and scheduling*, 3rd ed., New York et al.
- Ulrich, P.; Fluri, E. (1995) *Management*, 7th ed., Bern

14 Collaborative Planning

Christoph Kilger¹ and Boris Reuter²

¹ j & m Management Consulting AG, Kaiserringforum, Willy-Brandt-Platz 5, 68161 Mannheim, Germany

² axentiv AG, Uhlandstraße 9, 64297 Darmstadt, Germany

The preceding chapters deal with planning processes within one *planning domain*, e.g. an enterprise (demand planning, master planning) or a factory (production planning). The term *planning domain* constitutes a part of the supply chain and the related planning processes that are under the control and in the responsibility of one planning organization. However, the quality of a plan and the quality of the decision-making process that is based on that plan can often be improved by considering additional information that is beyond the scope of the individual planning domain.

In this chapter, we describe so-called *collaborative planning processes*, which span multiple planning domains. The idea is to directly connect planning processes that are local to their planning domain in order to exchange the relevant data between the planning domains. The planning domains *collaborate* in order to create a common and mutually agreed upon plan. Thus, input data is updated faster and planning results become more accurate. Figure 14.1 shows the Supply Chain Planning Matrices of two planning domains that are connected by a *collaboration*.

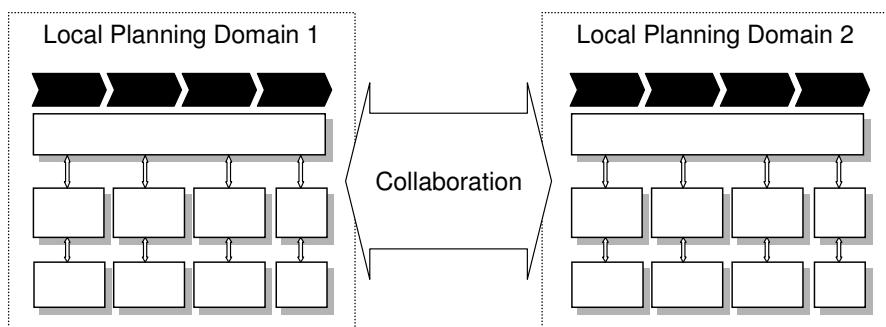


Fig. 14.1. A Collaboration connects the planning processes of planning domains

Collaborative planning concepts can be applied to the planning processes that interface with customers (e.g. sales planning) and to those that interface with suppliers (e.g. procurement planning). Further, collaborations can be distinguished by the objects that are exchanged and collaboratively planned, such as supply capability of suppliers or material demand of customers.

There also exist some approaches that cope with mutual reconciliation of activities, such as *Collaborative Planning, Forecasting and Replenishment* (CPFR) developed in the consumer goods industry (see VICS (1999)) and *Collaborative Development Chain Management* (CDCM), which follow the ideas of simultaneous engineering and focus on joint development of products by several partners with the use of web-based computer systems (see Becker (2001)).

Section 14.1 gives an introduction to collaborative planning and the type of planned objects, and Sect. 14.2 copes with types of collaborations. In Sect. 14.3 we describe a generic collaborative planning process. Section 14.4 gives an overview of APS-technology that supports collaborative planning processes.

14.1 Introduction

The following example illustrates the use of collaborative planning processes. Consider a manufacturer of headlight modules for the automotive industry. The manufacturer supplies headlight modules to two car manufacturers. A subcontractor can be employed to cover peak demand situations, providing additional assembly capacity. Headlight modules consist of a body and a glass cover. The body is produced by the manufacturer itself. The headlight glass covers are supplied by an external supplier in a make-to-order process. Bulbs are provided by a second supplier from stock.

Figure 14.2 illustrates the supply chain and gives some examples of collaborations:

- The car manufacturers are interested in getting a reliable supply of headlights. Therefore, they provide their demand forecast of headlight modules to the headlight manufacturer. From the headlight manufacturer they request a commitment to fulfil the demand forecast and information about the maximum supply capabilities. The latter is needed in case the actual demand exceeds the demand forecast, and in order to catch up fulfilment of the demand in case of a supply shortage (for instance if a die was broken). The headlight manufacturer is interested in getting a minimum demand commitment from the car manufacturer and has to plan its business and provide the appropriate capacity.
- Between the headlight manufacturer and the supplier of the glass covers a collaborative planning process helps to plan the demand and the supply capabilities in advance.
- Compared to this make-to-order business, bulbs are standard products that are made-to-stock. Both the supplier of bulbs and the headlight manufacturer keep a specific safety stock against demand and supply variabilities. Collaboration on inventory and on demand forecast helps to optimize bulb inventory and replenishment of bulbs.

- The subcontractor matches forecasted production demand with his actual production capacity. Thus, the manufacturer and the subcontractor collaborate on the use of the capacity at the subcontractor's site.

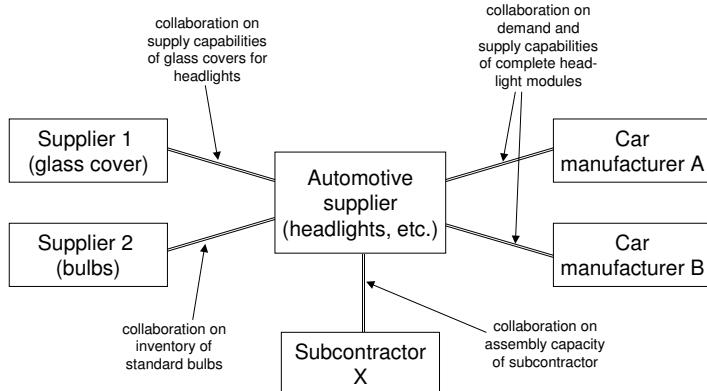


Fig. 14.2. Supply chain structure of a collaboration scenario

The example illustrates the main aspects of collaborative planning. A *collaboration* represents a business relationship between a supplier of an item and a customer of that item. The customer represents the *item demand* of the collaboration, the supplier represents the *supply capability* (see Fig. 14.3).

In business-to-business scenarios cooperations are described by some constitutive attributes (Jansen, 2001)

- Coordination and joint fulfillment of tasks.
- Legally and – in general – economically independent companies.
- Potentially higher success through joint activities.
- Voluntariness.
- Admissible.
- Agreement on common or at least compatible goals.

The term “collaborative” means – according to Webster’s Dictionary – “to work jointly with others or together especially in an intellectual endeavor” (Webster, 2001). In a more precise manner, Bruner (1991, p. 6) defines collaboration in the context of educational environments: “Collaboration is a process to reach goals that cannot be achieved acting singly (or, at a minimum, cannot be reached efficiently). [...] Collaboration includes all of the following elements: jointly developing and agreeing to a set of common goals and directions; sharing responsibility for obtaining those goals; and working together to achieve those goals, using the expertise of each collaborator.” Anderson and Narus (1998, p. 24) distinguish *transactional* and *collaborative* relationships. The former focus on the “timely exchange of basic products for

highly competitive prices [. . . The latter is characterized by . . .] strong and extensive social, economic, service, and technical ties over time, with the intent of lowering costs and/or increasing value, thereby achieving mutual benefit.” The relationship of supplier and customer might range between those two extremes. Collaborative planning requires a collaborative relationship with the intent of establishing a mid-term relationship to enable planning activities and the exchange of expertise-based partner information to create additional value. Thus, the collaboration is not a “one-transaction-relationship” with spot-market character.

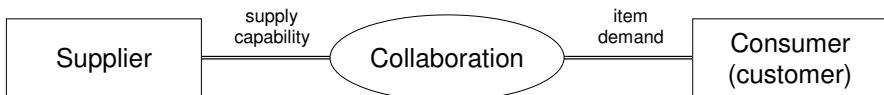


Fig. 14.3. Collaboration as a demand-supply relationship

A collaboration is related to a specific item, that – as illustrated by the example – may consist of sub-items (described by a bill-of-materials). In collaborative planning an important step is to manage the individual (item-related) collaborations. In the automotive industry, the car manufacturers define new items such as new headlights, as part of the design process of a new car. A decision workflow is initiated for each new item that ends with the definition of a new collaboration for this item (or the decision not to collaborate on that item).

Items may be organized in a hierarchy. For instance, the light system of a specific car model may be a top level item, and the headlight and backlight modules may be second level items. Demand and supply capabilities can be attached to any level. In the example, the demand may be attached to the top level item, expressing the total demand for light systems. This is broken down to the demand for headlight and backlight modules. These modules may be supplied by different suppliers. Thus, the individual supply capabilities of the headlight supplier and backlight supplier are attached to the second level in the item hierarchy.

14.2 Types of Collaborations

Collaborations can be classified according to multiple dimensions: Leadership, objects, and structure of the collaboration network.

Usually, one of the partners participating in the collaboration has a *leading role*, while the other (or others) are *followers*. Taking an example from the computer industry: computer manufacturers like Dell and Fujitsu Siemens Computers are in a leading role towards their suppliers of disk drives, memory modules, controllers, etc., but they are in a follower role towards Intel. The leading partner initiates and drives the collaborative planning process,

whereas a follower supports the process. Collaborations can be classified according to the leadership in *supplier-driven collaborations* (supplier has the lead) and *consumer-driven collaborations* (consumer has the lead). This classification corresponds to the notion of leadership in supply chains (as described in Chap. 1).

Collaborations are related to specific items that are provided by the supplier and are used by the consumer. Supplier and consumer exchange information about the demand and supply of those items. This information may be about the item itself (*material-related collaboration*) or it may be about capacity or services that are required to make the item, to install it, to transport it, etc. Collaborations of this type are called *service-related collaborations*.

A collection of collaborations forms a network structure. The nodes of the network represent the suppliers and the consumers, the directed edges represent the item-relationships connecting suppliers with consumers. If the maximum length of any path in the network (following the item-relationships downstream) is two, the network is a *two-tier collaboration network*. In a two-tier collaboration network there are no inner nodes, i.e. each node is either a supplier or a consumer (there is no node that is both, supplier and consumer). If there is a path in the network consisting of three or more nodes, the network is a *multi-tier collaboration network*. Note that in a multi-tier collaboration network the inner nodes are acting as supplier *and* as consumer, as they are connected to at least two partners, one being a consumer and one being a supplier. Dudek (2004) describes the properties of two-tier and multi-tier collaboration networks in detail.

In the following we describe typical collaborative planning processes.

14.2.1 Demand Collaboration

In a typical supply chain some of the supply processes are driven by the forecast. The interface between order-driven and forecast-driven processes in a supply chain is called *decoupling point* (see Chap. 9, p. 183). Multiple departments are involved in the creation of the forecast, e.g. sales, marketing, product management and central planning. Sales enters the forecast for specific customers or regions, and plans the influence of market trends. Marketing adds the influence of marketing activities and promotions to the forecast, and product management provides information about the phase in/phase out of products. Central planning consolidates the plan from an overall perspective. The consolidation of the inputs from the departments is called *consensus-based forecasting*. The statistical forecast serves as a reference and starting point for the human planners (in sales, marketing, etc.). Procedures for the integration of statistical forecasting and the structured judgement of human planners is described in Chap. 7.

The customers – whose demand is being planned – may add valuable input to the forecasting processes. For instance, customers can provide the mid-term material requirements based on their master plan as an input for

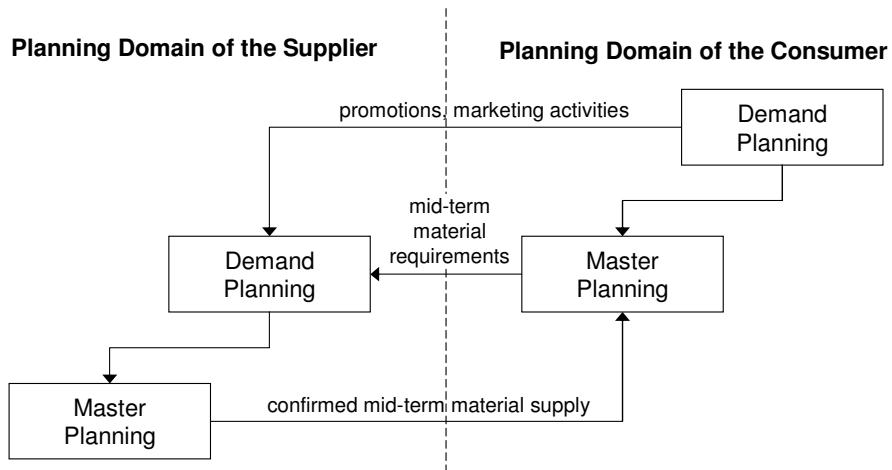


Fig. 14.4. Relationships between the planning domains in a demand collaboration

the demand planning of their suppliers. The suppliers use this information as input to the consensus forecasting processes as described above. Further, the consolidated and approved forecast can be confirmed to the consumers (customers) for consideration in their master planning (represented as confirmed mid-term supply). In this case a *demand collaboration* network is formed between the suppliers and the consumers, connecting the forecasting processes of the local planning domains. A demand collaboration is usually driven by the supplier – as he is interested in getting accurate information about the future demand. Fig. 14.4 shows the connection between the local planning domains in a demand collaboration.

The scenario depicted in Figure 14.2 (p. 261) provides an example for a demand collaboration: The headlight manufacturer might setup a demand collaboration with both car manufacturers to create a collaborative forecast reflecting the actual demand for headlights in the supply chain. Another example of a demand collaboration is a joint collaborative promotion planning: the consumer provides detailed information about planned promotions or other marketing activities, the supplier considers this information as input to the demand planning process. Smaros (2003) gives further examples of collaborative demand planning.

Prerequisites to enter a demand collaboration are harmonized master and transactional data. Every local planning domain is able to analyze the planning process by custom views. Deviations are reported by alerts and should be discussed and adjusted in cyclic planning meetings. The result of the planning meetings is an agreed demand plan. The quality of past decisions and planned forecasts have to be analyzed ex-post based on the historic sales figures.

The demand of a consumer (customer) participating in a demand collaboration must be treated differently from the demand of other customers not participating in a demand collaboration: Partners in a collaboration are more open and provide better and more reliable input to the collaborative planning processes than other customers. To avoid shortage gaming (see Chap. 1, p. 30), the demand of partners participating in a collaboration has to be fulfilled with higher priority compared to the demand of other customers.

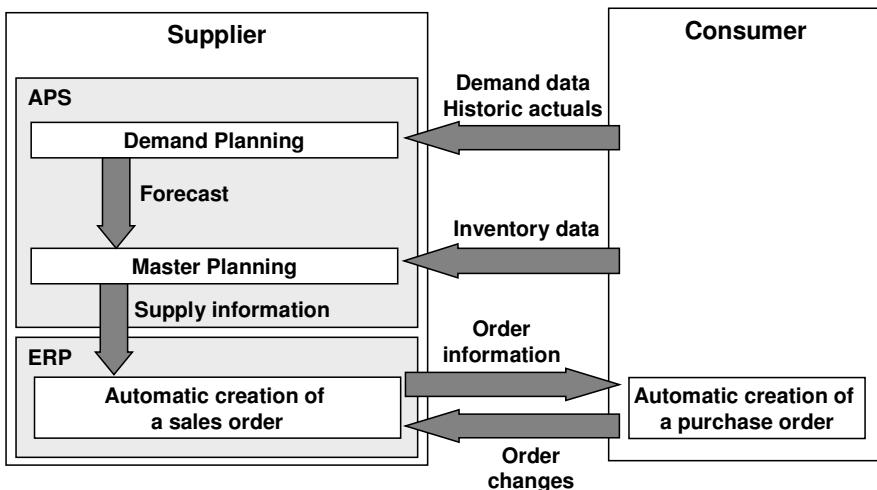


Fig. 14.5. Relationships between supplier and consumer in an inventory collaboration

14.2.2 Inventory Collaboration

Inventory collaboration is a special application of demand collaboration. The consumer provides information about his future demand and about the current inventory to the supplier. The supplier uses this information to create the requirements of his own products at the sites of the consumer (e.g. factory sites, warehouses). Consequently, the consumer needs no longer to create and send replenishment orders to the supplier. The replenishment of the inventory is automatically planned by the supplier; time-lags due to the replenishment planning and ordering processes of the consumer do no longer occur. The replenishment decisions are driven by pre-defined service level agreements between supplier and consumer (e.g. expressed as minimum coverage time of the stock level). Inventory collaboration is a service that is usually requested by the consumer (or is at least tolerated by the consumer). The process itself is driven by the supplier. For inventory collaborations also the term *vendor managed inventory (VMI)* is used.

To control his own inventory and his customer's inventory simultaneously the supplier has to be able to access the consumer's major inventory levels and forecasts and has to plan with respect to a system-wide inventory. This could be done by using the so-called base-stock-system (Tempelmeier, 1999). Modern EDI-techniques support the electronic exchange of the necessary information (e.g. inventory data, demand data, planned replenishments). Usually the supplier automatically generates a sales order in his ERP system based on the replenishment plan. The order information is sent to the ERP system of the consumer and a purchase order matching the sales order is created automatically. Fig. 14.5 summarizes the relationships between supplier and consumer in an inventory collaboration.

14.2.3 Procurement Collaboration

A procurement collaboration – also called supply-side collaboration (Fu and Piplani, 2004) – is similar to a demand collaboration. The consumer and the supplier exchange demand and supply information as shown in Fig. 14.4. The main difference is that a procurement collaboration is driven by the consumer, whereas a demand collaboration is driven by the supplier.

Mid-term procurement collaborations provide information about material constraints to the master planning (Chap. 8). Short-term procurement collaborations interface with the purchasing and material requirements planning (Chap. 11), and the short-term material supply information is used to update short-term plans, for example production schedules.

14.2.4 Capacity Collaboration

The collaborations discussed so far – demand, inventory and procurement collaboration – are related to the exchange of demand and supply information of material (items). A capacity collaboration is an example for a service-related collaboration: Supplier and consumer exchange information about demand and availability of production services. For instance, a manufacturer (i.e. the consumer) collaborates with a subcontractor (i.e. the supplier) on the usage of the subcontractor's production facility based on the manufacturer's master plan. The manufacturer wants to ensure that he gets a reservation for a specific amount of capacity, without knowing for which production order the capacity actually will be used and what product actually will be produced. Similar to procurement collaborations capacity collaborations are usually driven by the consumer.

Besides the forecasted capacity, a minimum and maximum capacity level is often negotiated between the two parties:

- The subcontractor (supplier) is interested in defining a minimum capacity to ensure the load of his production facilities.

- The manufacturer (consumer) of the capacity is interested in knowing the maximum capacity that is provided by the subcontractor.

The difference between the forecasted capacity and the maximum capacity is called the *upside flexibility* of the subcontractor. However, if the subcontractor has multiple consumers using the same capacity this upside flexibility range might be announced to more than one consumer. In this case multiple manufacturers are sharing the flexibility range.

The typical goal of a capacity collaboration is to provide additional upside flexibility for the manufacturer (the consumer), in case his own capacity is fully loaded. However, in practice it often occurs that the manufacturer has to load the capacity of the subcontractor first before loading its own capacity in order to create the contracted minimum load of the subcontractor's production facilities and to avoid penalties.

14.2.5 Transport Collaboration

Transport planning and vehicle scheduling is one of the operational tasks of purchasing and distribution (Chap. 12). Often, several logistic service providers are involved in the main purchasing and distribution process of an enterprise or a certain part of the supply chain. The transportation services (for inbound and for outbound transportation) are nowadays provided by external transportation and logistics providers.

A transport collaboration is similar to a capacity collaboration: both are service-related collaborations driven by the consumer. While the capacity collaboration is related to production services, the transport collaboration is related to transportation services. In a transport collaboration the consumer is typically a manufacturer or retailer, and the supplier is a transportation and logistics provider.

For example, a transport planner of a manufacturer (i. e the consumer) uses a planning tool to assign transport requests to a provider either by hand or through an optimization run. The requests are sent to the provider, e. g. by e-mail, containing a hyperlink to a website or an XML-document. The provider checks the request and accepts or modifies the conditions, e. g. route, pick-up points and delivery dates, or rejects it in a predefined time window. Alerts are generated, if for example the requested transport capacity exceeds the agreed quantities, response is belated or the request has been changed or rejected. For the latter cases the transport planner can accept the change or choose a different provider. With the acceptance of a request a predefined order fulfilment workflow starts.

14.2.6 Material- and Service-Related Collaborations

Demand, inventory and procurement collaborations are material-related collaborations, capacity and transport collaborations are service-related. Besides

these “pure” material- or service related collaborations there exist combined material- and service-related collaborations. These collaborations are formed mainly in industries where materials and services have to be synchronized in order to efficiently and reliably fulfill the customer demand.

As an example for a material- and service related collaboration we consider the procurement of computer equipment. Large organizations such as banks, assurance companies, public administration etc. procure large quantities of computer equipment, including computers, servers, printers, network components etc. The procurement process is often organized as a “rollout project” that is managed by a specialized service company (see Fig. 14.6). Hardware suppliers provide products (materials), service providers provide transport, customization and computer installation services. For instance, the computer equipment that is to be installed in one floor of an office building is collected at the customizing centre. If it is complete, all servers, workstations, printers and further equipment are customized, software is installed, network addresses are assigned etc. After customization is complete the transportation and logistics provider forwards the equipment to the installation site. Technicians are arriving at the same time on site and install and replace the old equipment by the new one.

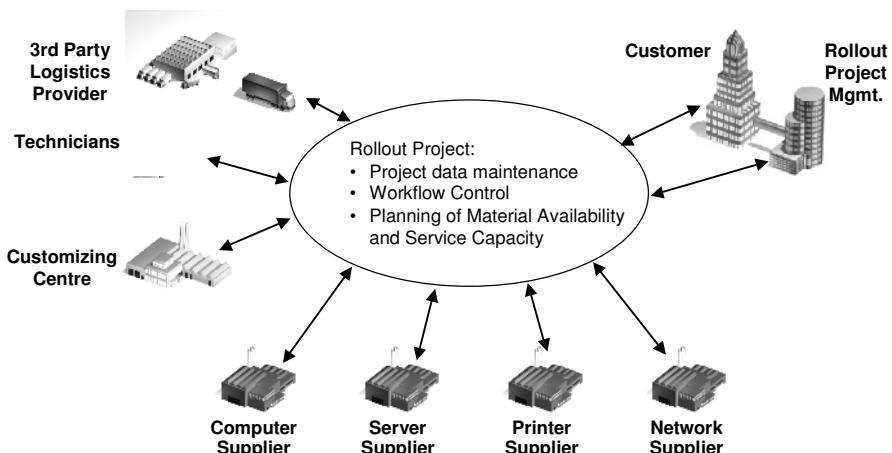


Fig. 14.6. Structure of a rollout project

Traditionally, the availability of materials and service capacity is controlled manually, leading to an insufficient information flow and slow reaction in case of changes. As a consequence, the delivery performance is low and large inventories are stocked at all involved sites to buffer against shortages. In order to better coordinate all involved parties by faster information exchange a material- and service related collaboration is formed. This collaboration is usually driven by the service company coordinating the project.

The customer receiving the customized computer equipment has the role of the consumer, the hardware suppliers and service providers act as suppliers.

The material- and service-related collaboration may be supported by collaboration modules of APS, typically in combination of an Internet portal consolidating all information flows and providing role-specific views for all involved parties. As an example, consider a material shortage at the server supplier. Having an APS-based collaboration process installed, the server supplier updates the availability information for the servers that are needed for the rollout project. The remaining hardware suppliers use this information to adjust their production and distribution plans. The service providers update their plans accordingly, and may for instance re-allocate available capacity to other customers or may even reduce capacity in case they employ subcontracted workforces.

The synchronization of services and materials by APS-based collaboration processes gets more and more important as services gain a broader share in many industries. Other examples of industries with a high fraction of services that have to be synchronized with material availability are telecommunications, building and construction industry, and medical technology industry. For further details on collaboration on services and materials refer to Kilger and Holtkamp (2001) and Keinert and Ötschmann (2001).

14.2.7 Multi-Tier Collaboration Networks

The collaboration types described so far are *single-tier* collaborations, connecting a customer with its direct suppliers. If the supply chain extends over multiple supplier tiers, e.g. in the automotive industry, this results in a chain of individual single-tier collaborations (see Fig. 14.7). Each collaboration connects one supplier-customer pair. Information about a changed demand-supply situation must be propagated along all collaborations before the entire supply chain works according to the new situation. If for instance the customer and all three suppliers have a weekly planning cycle, it takes three weeks until a changed or new demand signal reaches the tier 3 supplier.



Fig. 14.7. Chain of single-tier collaborations

In order to speed up the information exchange in the supply chain a *multi-tier* collaboration can be established, directly connecting the customer with the tier 1, tier 2, and tier 3 suppliers. Figure 14.8 visualizes such a multi-tier collaboration. All members of the supply chain work according to the same “beat”; information about demand or supply changes are propagated

within one planning cycle to all supply chain members (Kilger and Stahuber, 2002). One example of a multi-tier collaboration has been implemented by DaimlerChrysler, connecting the planning processes of DaimlerChrysler and tier 1 to tier 7 suppliers for parts of door modules (see Graf and Putzlocher (2002)).

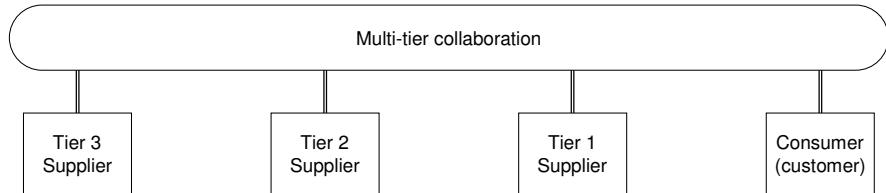


Fig. 14.8. Multi-tier collaboration

A successful multi-tier collaboration requires one distinguished supply chain member who is driving the collaborative planning processes and defines the rules and standards of the collaboration. In the automotive industry, this role is usually taken over by the automotive manufacturer, as he is controlling the supplier network.

Common issues of a multi-tier collaboration are different batching rules and inventory policies of the suppliers. For instance, assume the tier 2 supplier shown in Fig. 14.8 has a batching rule telling him to order multiple quantities of 100 from the tier 3 suppliers. The multi-tier collaboration has to know this batching rule, because otherwise the demand signal reaching the tier 3 supplier from the collaboration will be different from the actual demand of the tier 2 supplier.

14.3 A Generic Collaboration Process

A typical generic collaboration process is as follows (see Fig. 14.9):

1. Definition
2. Local domain planning
3. Plan exchange
4. Negotiation and exception handling
5. Execution
6. Performance measurement

Definition The definition of a collaborative relationship of business partners incorporates the goal of working together in some mutually defined ways by a formal agreement. Four main issues have to be considered: gives & gets, the collaboration items including planning horizons, the time horizon of the collaboration and an agreed dispute resolution mechanism in case of conflicts (along the lines of (Anderson and Narus, 1998, p. 25)):

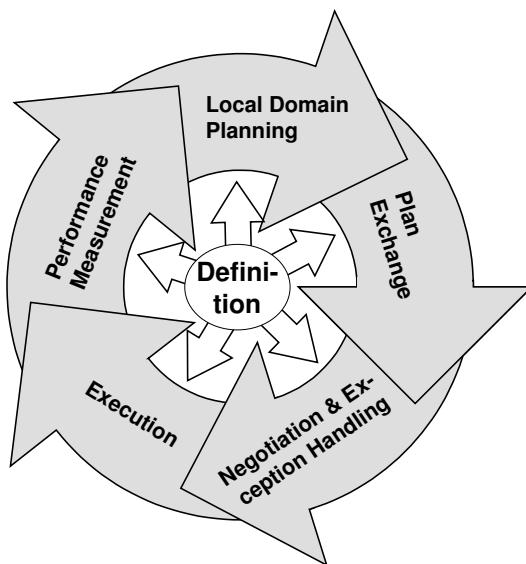


Fig. 14.9. Collaboration cycle illustrating the generic collaboration process

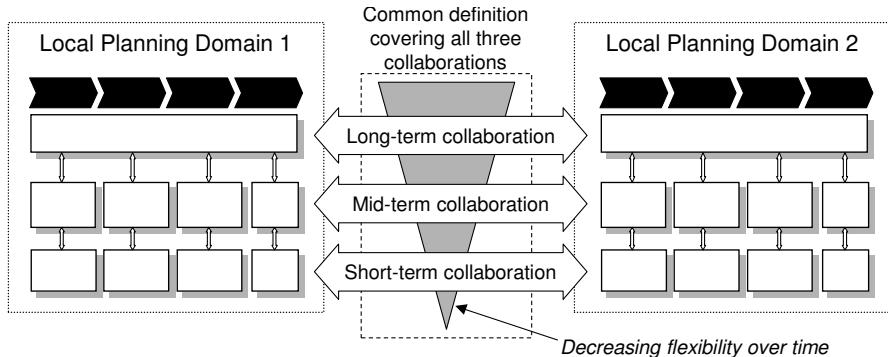


Fig. 14.10. Long-term, mid-term and short-term collaborations

- “Gives” address the contribution of each partner to the collaboration, e.g. personnel, fixed assets, money, knowledge, commonly used software, whereas “gets” are the specific gains of each partner participating in the collaboration, e.g. greater expertise, broader market access and additional earnings. Conflicts often occur if one partner’s perception of gives compared to gets received is not balanced. Monitoring gives & gets by success metrics, e.g. KPIs, helps to avoid discrepancies, supports compensations and fosters a continuous improvement process.
- The collaboration items are products and/or services to which the collaboration is related. By focusing on main material flows in the supply chain,

important products such as bottleneck raw materials, products with long lead-times or high-value are included in a collaboration process. Related to the items are parameters such as minimum demand quantities, exception rules as well as classification of importance for several partners. Furthermore, the reconciliation of plans on different planning horizons has to be defined.

- The time horizon determines the duration of the collaboration. It also contains milestones for common aims and review points to analyze the relationship. At the end of the time horizon the partners have to decide whether to continue, expand or curtail the relationship.
- Close relationships include potential disagreements and conflict situations. Thus, an agreed dispute resolution mechanism has to be established. Depending on the severity of the conflict, different mechanisms might be taken into account, e.g. negotiation processes to rearrange agreements, mediation to focus on objective conflict issues by external moderation, or arbitration to accept a third parties' decision as final and binding.

Within the definition of the collaboration items, the collaboration is applied to the three *planning horizons* of the Supply Chain Planning Matrix (see Chap. 4): long-term, mid-term and short-term. Thus, a supplier and a customer can adjust their long-term, mid-term and short-term collaborative planning process by connecting their corresponding local domain planning processes. The collaborative plans are then locally disaggregated (see Fig. 14.10).

As mentioned above, a collaboration covers a specific *time horizon*. The time horizon of a collaboration is usually structured into multiple time phases, each representing a specific degree of decision flexibility (see also Fig. 14.11). The *history phase* represents actuals of a collaboration, e.g. ordered and supplied quantities, actual inventory levels etc. The actuals are used as a foundation on which the future development of the collaboration is planned. The *frozen phase* covers the next time buckets, e.g. the next four weeks. In that time horizon the plan is fixed for execution. During the *commit phase* the plan is being reviewed in detail and approved to become fixed for execution. The length of the commit phase indicates the (maximum) duration of the commitment process. The *forecast phase* covers the remaining time buckets of the time horizon, reflecting the current plan of the collaboration.

After the definition of a collaboration, the framework for operational business is established.

Local Domain Planning A planner organizes his future activities in a local domain plan, that takes into account a certain local planning situation, his individual objective function, current detailed internal information, know-how about process restrictions and assumptions about the environmental development. In particular, assumptions about planned activities of suppliers and

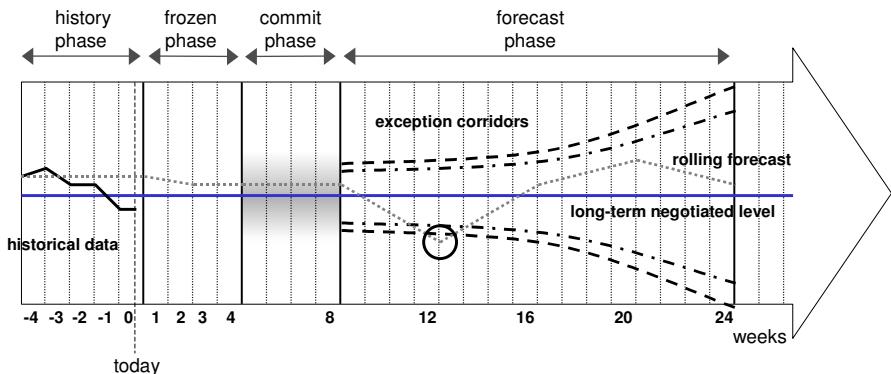


Fig. 14.11. Time phases of a collaboration (example)

customers are uncertain without collaboration. In a decision making process several plans are created, evaluated and ranked by an objective function to identify the best one. Plans having similar objective function values may have very different structures. Thus, alternative plans should not be discarded, but stored in separate versions. This enables the planner to react to changes in the planning environment such as changes in restrictions. In a collaboration the locally created plan will be the basis for communication with partners.

Plan Exchange The plan exchange process is highly sensitive. The partners intent to augment planning quality by exchanging information. In the definition phase of a collaboration objects are defined such as products on which data might be exchanged. Depending on the content, e. g. inventory of a certain delivered product or inventory of all similar products, the accuracy and the use of data lead to more or less valuable information. The sources of data might be transactional data of suppliers and customers, that are maintained in ERP-systems, or their local domain plans such as forecasted demand, replenishment orders or supply commitments retrieved from an APS.

Negotiation and Exception Handling The partners exchange planning information under the terms defined in the collaboration process. This enables partners to gain an overview of the planning situation and identify whether the predefined goals are achieved.

Dudek (2004) describes a negotiation-based collaboration process that is based on an iterative improvement process between consumer and supplier. The process is depicted in Fig. 14.12. Both – consumer and supplier – exchange the following information items:

- compromise order/supply patterns,
- local cost effect of a compromise pattern,

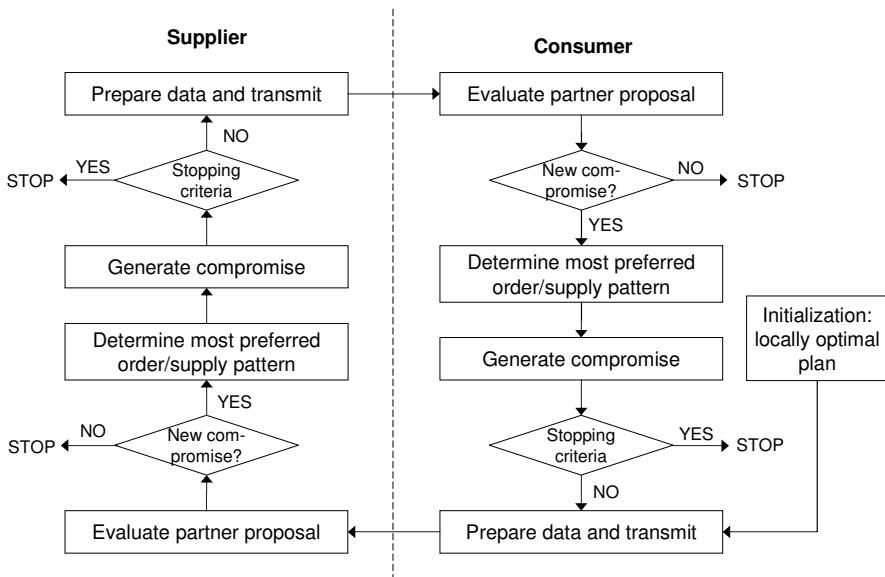


Fig. 14.12. Negotiation-based collaborative planning process (adapted from Dudek (2004))

- cost increase due to partner's previous proposal,
- total cost effect of previous solution,
- total cost effect of best solution, and
- identifier of best solution.

The negotiation process is initialized by the locally optimized plan of the consumer. This plan is transmitted to the supplier, who evaluates this proposal. The evaluation of the partner plan is accomplished by intra-domain planning, enhanced by additional constraints. The role of the stopping criterion is to determine whether or not to continue the iterative process based on the current, previous and best cost outcome detected so far. Dudek suggests to use a Simulated Annealing meta-heuristic to control the decision process whether to go on with the improvement or to stop the iterations (Dudek, 2004). The purpose of the determination of preferred order/supply patterns is to find all modifications to the received order/supply pattern which improve the local cost situation. In the next step a compromise proposal is generated based on the locally optimized order/supply pattern. Dudek describes MIP models for the generation of locally optimized patterns and the generation of a compromise proposal including the anticipation of partner costs. For the case that the selected compromise proposal is equivalent to the compromise of an earlier iteration, methods are described to compare the current proposal to the former ones and to generate additional compromise proposals based on a model-based approach.

After completion of the iteration steps of the supplier the same sequence of steps is executed by the consumer. The procedure terminates if one of the stopping criterions is fulfilled, or if no new compromise proposals can be found. Dudek generalizes the procedure to general two-tier collaborations (multiple suppliers connected with multiple consumers) and to multi-tier collaborations.

Stadtler (2003) introduces a multi-agent system that adjusts local domain plans between two partners in a supply chain, based on the exchange of purchase orders, supply proposals, latest due dates and compensations. A compensation is requested by the consumer if the supplier deviates from the plan submitted by the consumer. The procedure leads to near optimal solutions for the simultaneous integration of the two optimal local domain plans. A comprehensive survey of electronic negotiation processes gives Rebstock (2001).

The common goals and conditions of the partners are measured by KPIs. Planning results, both for the local domain and collaborative plans, are compared with the real-world data based on the KPIs. Analysis of plan deviations helps in identifying ways in which future plans may be improved. Various data views and aggregation levels of the data to be compared support this analysis. Reactions to deviations from the plan are closely associated with the plan review.

If the partners have decided on a particular threshold value for a given KPI exceeding this value should trigger a process which either pushes the KPI back within its allowed range or allows an exception to occur. The first case strictly disciplines unauthorized actions by partners, initiates a negotiation process to mutually align plans between partners or is used to achieve a desired supply chain behaviour (such as less planning “nervousness”). The second case comes into play where structural changes or other exceptional situations occur. Causes for exceptions might be internal, e.g. planning faults or insufficient decision support, or external such as changed economical or competitive situations. Exception handling is triggered by alerts indicating specific planning problems, for example:

- mismatch between the forecasted demand and the supply capability,
- violation of a minimum demand level,
- a missing response from the supplier to match a forecasted demand,
- an item demand planned by a customer that is not yet released for collaboration by the supplier.

Execution An adjusted plan leads to replenishment-, production- and purchasing orders to fulfil the planned goals and is then executed.

Performance Measurement After plan execution, “as-is/to-be” analysis is used to measure the effects of the implementation of the global plan. Planning results are more easily accepted by everyone in the case of a win-win

situation throughout the entire supply chain. More difficult is the case in which some members lose. Compensation approaches must be developed in this case which lead to reimbursement of the members who agree to “lose” for the good of the supply chain as a whole. The deviation from a local domain plan can be used as a measurement.

14.4 Software Support

Planning and coordination within an enterprise are difficult tasks for today’s software, and the addition of collaborations increases this complexity. Challenges of collaboration planning tools include master data integration, user-specific secure data access and the mutual decision-making process. Systems that enable collaborative planning must support partners during each step of the process.

Definition The definition-step establishes a framework of collaboration and consists of a management agreement to confidential cooperation as well as the definition of common goals, objects of collaboration, success metrics and incentives/penalties. The selection process of appropriate items or partners is supported by reporting systems based on Data Warehouses. As the selection process is qualitative and thus not supported by APS, the results of a collaborative planning agreement must be customized in the APS. For example, SAP APO allows the authorization of specific users, the specification of the type of collaboration as well as the definition of exchanged data such as master data and – in case of forecast data – the time series of dedicated key figures. This issue is of critical importance, because an incorrect mapping of master data or time series granularity causes severe planning problems.

Local Domain Planning The step of local domain planning to generate individual plans by each partner is the main focus of APS. As planning becomes more complex with respect to the consideration of partners’ plans, several “good” plans with different structures or containing changed data-scenarios might be stored in so-called versions. Thus, changes of considered planning restrictions are anticipated and responsiveness is improved.

Plan Exchange The step of exchanging plans with a partner is related to the data implemented by the customizing of the collaboration. Furthermore, the way in which data are exchanged is defined by workflows. These include entering data using a web-based interface, or exchanging information such as orders or time series in one of several formats: XML-documents, RosettaNet, EDI, Excel spreadsheet or flat file.

SAP ICH (Inventory Collaboration Hub) is a typical example for a web-based exchange platform for demand and supply plans. SAP ICH supports

two basic collaborations methods on the procurement side. The first one – *Supplier Managed Inventory (SMI)* – supports a traditional min/max-based supplier-driven replenishment and inventory monitoring process. The second method is *Release Processing*. It represents support of buyer-driven replenishment via SAP R/3 delivery schedules, typically derived from buyer-side internal MRP runs. In addition, SAP ICH supports the alert monitoring, the master data maintenance and the processing of advanced shipment notes (ASN).

One of the main problems facing todays APS is the lack of considering interdependencies between multiple exchanged items and their concurrent use of capacities. Thus, bottlenecks cannot be avoided by planning items one at a time, and the resulting reconciliation efforts might increase with the number of items and the complexity of planning restrictions.

Negotiation and Exception Handling To support the step of negotiation and exception handling, rules that trigger information flows indicating specific planning problems have to be defined. The rules are related to planning objects such as resources indicating capacity overload, materials indicating shortage situations or lateness of an order. Most APS contain some predefined rules (e.g. SAP APO Alert Monitor profiles) or have a programming interface to trigger alerts by deviations from calculated key figures such as exception corridors shown in Fig. 14.11 (e.g. i2 – violation of funnel agreement; SAP APO – MacroBuilder to define user specific alerts). Depending on the severity of deviations from the agreed limit and the ability to influence the plan either negotiation processes are started by defined workflows to align the plan (e.g. splitting an order) or an exception is allowed (e.g. sourcing from a partner's competitor).

Execution The execution step contains the fulfilment of an aligned plan between the partners. It leads to activities in transactional systems (e.g. SAP R/3) such as entering production or replenishment orders. Shop floor control systems support “track and trace” of orders and material flows, resource loads and staff assignments.

Performance Measurement In order to identify the quality of a fulfilled plan or a sequence of plans, results must be measured. “Plan vs. as-is” data are analyzed using reporting tools. For example, input of transactional data such as sales are compared to sales forecast data to identify gaps and opportunities for improvement. Inside a single APS it is customary to define KPIs and KPI schemes in order to measure supply chain performance. Special tools (e.g. SAP APO Plan Monitor) are then used to keep track of the KPI values. KPI schemes throughout the entire collaboration have to be customized in each of the partners' APS. That is, in order to measure collaboration performance, the KPI schemes must be agreed upon by each collaboration partner.

References

- Anderson, J. C.; Narus, J. A. (1998) *Business market management: Understanding, creating, and delivering value*, New Jersey
- Becker, T. (2001) *Collaborative Development Chain Management*, Supply Chain Management, No. 1, 27–36
- Bruner, C. (1991) *Thinking collaboratively: Ten questions and answers to help policy makers improve children's services*, Washington DC, Education and Human Services Consortium, Cited from North Central Educational Laboratory Homepage, URL: <http://www.ncrel.org/sdrs/areas/issues/envrnmnt/drugfree/sa1lk26.htm>, Date: July 19, 2004
- Dudek, G. (2004) *Collaborative planning in supply chains – A negotiation based approach*, Berlin et al.
- Fu, Y.; Piplani, R. (2002) *Supply-side collaboration and its value in supply chains*, in: European Journal of Logistics, 152, pp. 281-288
- Graf, H.; Putzlocher, S. (2002) *DaimlerChrysler: Integrierte Beschaffungsnetzwerke*, in: Corsten, D.; Gabriel, C. (Eds.) *Supply Chain Management erfolgreich umsetzen*, Heidelberg, 47–61
- Huang, G. Q.; Lau, J. S.; Mak, K. L. (2003) *The impacts of sharing production information on supply dynamics: a review of literature*, in: International Journal of Production Research, Vol. 41, 7, 1483–1517
- Jansen, S. (2001) *Mergers and Acquisitions – Unternehmensacquisitionen und -kooperationen*, Wiesbaden
- Keinert, W.; Ötschmann, K. (2001) *Vom SCM zum e-Service-Chain Management*, Computer@Production, No. 3, 68–69
- Kilger, C.; Holtkamp, R. (2001) *Koordinierung von Dienstleistungen und Material mit eBusiness- und Supply-Chain-Management-Technologien*, in: Dangelmaier, W. et al. (Eds.) *Die Supply Chain im Zeitalter von E-Business und Global Sourcing*, Paderborn, 373–388
- Kilger, C.; Stahuber, A (2002) *Integrierte Logistiknetzwerke in der High Tech-Industrie*, in: Baumgarten, H.; Stabenau, H.; Weber, J. (Eds.) *Management integrierter logistischer Netzwerke*, Bern
- Rebstock, M. (2001) *Elektronische Untersttzung und Automatisierung von Verhandlungen*, in: Wirtschaftsinformatik, Vol. 43, 6, 609–617
- Smaros, J. (2003) *Collaborative forecasting: A selection of practical approaches*, in: International Journal of Logistics, Vol. 6, 4, 245–258
- Stadtler, H. (2003) *Advanced planning systems – basics and beyond*, Schriften zur Quantitativen Betriebswirtschaftslehre (working paper TU Darmstadt), No. 3/2003, 1–16
- Tempelmeier, H. (1999) *Material-Logistik: Modelle und Algorithmen für die Produktionsplanung und -steuerung und das Supply Chain Management*, 4th ed., Berlin et al.
- VICS (1999), *CPFR guidelines*, URL: <http://www.vics.org>, Date: July 19, 2004
- Webster (2001) *Homepage*, URL: <http://www.webster.com>, Date: July 19, 2004

Part III

Implementing Advanced Planning Systems

15 The Definition of a Supply Chain Project

Christoph Kilger

j & m Management Consulting AG, Kaiserringforum, Willy-Brandt-Platz 5,
68161 Mannheim, Germany

Supply Chain Management aims at improving competitiveness of the supply chain as a whole, by integrating organizational units along the supply chain and by coordinating material, information and financial flows in order to fulfill (ultimate) customer demands (Sect. 1.1). Supply Chain Management projects range from functional improvements on the IT level to large-scale change programmes. Functional improvements might be the introduction of a new forecasting method or the adjustment of the master planning optimization profile. Examples for larger SCM projects are the optimization of the supply chain network, the redesign of the planning processes or the adjustment of the business strategy based on SCM concepts. In either case, the goal of SCM projects is to improve competitiveness of the supply chain.

In recent years since the peak of the e-business hype Supply Chain Management and especially Advanced Planning Systems were viewed more and more critically by industry firms, as many SCM projects failed or did not realize the promised business value. There are three reasons for that.

The first reason for the failure of SCM projects is the perception “that the more you spend on IT (e.g. APS) the more value you will get from it” (Willcocks et al., 2002). This attitude leads to technical capabilities searching for business problems to be solved. The role of IT (and APS) was clearly over-estimated in the past as the single source of business value. In order to get “value” out of an APS implementation the SCM concept must be formulated prior to the APS implementation. The SCM concept defines the needs and priorities for the APS implementation; the APS supplies advanced planning functionality to be utilized by the SCM concept. For example, the SCM concept describes the network organization, the planning processes and the supply chain performance targets and indicators.

A second reason for the failure of SCM projects in the past is an inadequate alignment of the SCM concept with the supply chain strategy. Many decisions that have to be taken as part of SCM projects have a direct impact on the supply chain strategy (i.e. support the strategy or not) and must therefore be aligned with the supply chain strategy. Examples are the integration with supply chain partners, coordination and leadership in the supply chain and the results from strategic network planning. The supply chain strategy sets the objectives and the direction for the SCM concept; the SCM concept must support the supply chain strategy. The supply chain strategy itself must be derived from the overall business strategy and may influence the business

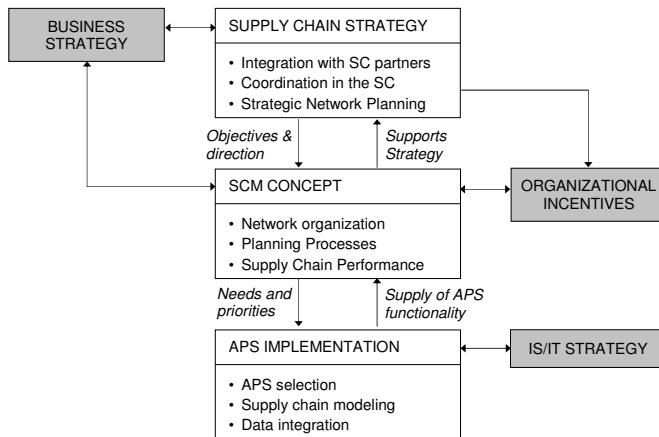


Fig. 15.1. Strategic SCM framework (adapted from Ward and Peppard (2002))

strategy. Further, decisions about organizational incentives (e.g. profit sharing, management bonus) for the supply chain participants must be aligned with the supply chain strategy and concept. Note that the APS implementation layer must be aligned with the information systems and information technology strategy (IS/IT strategy). In particular, the selection of an APS vendor is governed by the IS/IT strategy (see also Chap. 16). The relationship between business/supply chain strategy, SCM concept, organizational incentives, APS implementation and IS/IT strategy is depicted in Fig. 15.1.

A third reason for the success or failure of SCM projects is found in the organizational and managerial culture of industry firms. Based on our experience from many SCM projects six management practices are important prerequisites to make SCM projects successful. These practices were also described in a study by Collins (2001) of 11 large companies that consistently and massively outperformed their competitors over a decade or more:

- a mature and strong leadership,
- a focus on people and their strengths and skills,
- an ability to confront the brutal facts without losing faith in the end goal,
- a clear and well-formulated business strategy, backed by a viable financial model, passion and ability to be world class in delivering the idea,
- a culture of discipline and
- seeing technology as an accelerator of business performance rather than a single cause of momentum and breakthrough.

The initial phase of a SCM project must deliver a thorough understanding of the current situation, potential improvement areas and associated risks. Sect. 15.1 describes the phases of a supply chain evaluation process. Further, it describes the functional areas of a supply chain that have to be examined in order to make an initial assessment of the current structure and performance

of the supply chain. The supply chain evaluation answers the question *Where are we today?*

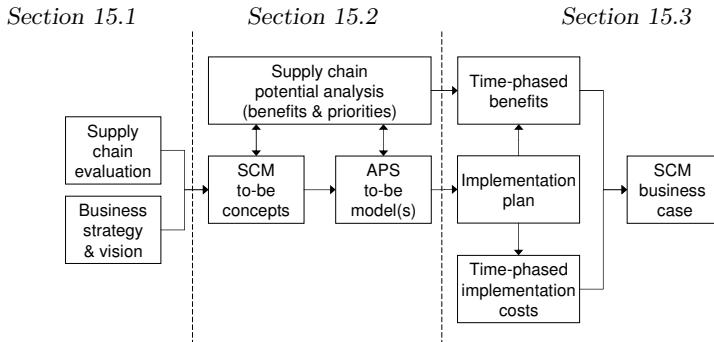


Fig. 15.2. Structuring the phases of a SCM project definition

Based on the business strategy and the results from the supply chain evaluation the improvement areas in the supply chain are identified. These are mapped to SCM concepts capable of improving the supply chain performance. To-be APS models are designed supporting the SCM to-be concept, and the related benefits are described and quantified – with the help of logistical and financial *key performance indicators* (KPIs). In particular, the impact of *external factors* influencing the performance of the supply chain can be attenuated by SCM concepts, as closer integration and coordination enables quicker and optimized reactions to external changes. This phase – Supply Chain Potential Analysis – is described in Sect. 15.2, and answers the question *Where do we want to go?*

In the last step of the definition of a SCM project the total scope of the project is broken down into smaller sub-projects, each of those having a specific business objective. The sub-projects are time-phased according to a high-level implementation plan. Benefits and implementation costs are time-phased based on the implementation plan, resulting in a business case and a return on investment calculation for the SCM project. Sect. 15.3 describes the procedure to create a project roadmap, answering the question *How do we get there?* All three phases together – supply chain evaluation, potential analysis, and roadmap – constitute a systematic approach to defining a supply chain project (see Fig.15.2).

15.1 Supply Chain Evaluation

The supply chain evaluation is structured according to the functional areas of the supply chain organization, including executive management, the IT function, suppliers, customers and competitors. The following paragraphs

discuss topics that have to be clarified with the various functional entities in an organization before entering an APS implementation project. Figure 15.3 shows the structure of a supply chain evaluation.

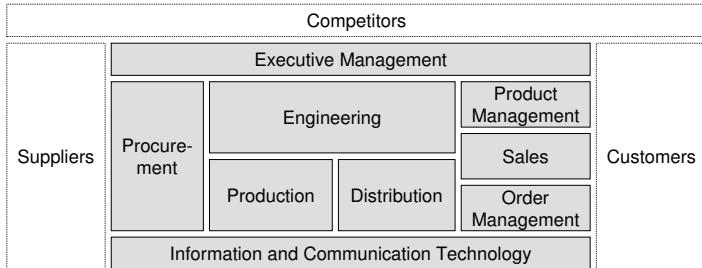


Fig. 15.3. Participants of a Supply Chain Evaluation

To get an initial overview of the supply chain, the SCOR methodology can be applied as part of the supply chain evaluation. With SCOR, the logistical structure of the supply chain can be visualized using Supply Chain Threads, and the processes of the supply chain can be documented by mapping them to SCOR process categories on level 2. For details on the SCOR methodology and further techniques for analyzing supply chains refer to Chap. 2.

Another tool that has proven to be helpful in the early stage of SCM projects is the lead time analysis of supply chain decisions shown in Fig. 15.4. On the right hand side of the figure the typical distribution of customer orders on hand – based on their due date and the number of orders or the order quantities – is shown over time. On the left hand side the lead time of decisions that have to be taken prior to order fulfillment is depicted. For example, if the procurement lead time of some material is 4 weeks, assembly lead time is 1 week and distribution and transportation lead time is also 1 week, the procurement decision must be taken 6 weeks in advance before customer orders using that material can be fulfilled. This analysis helps to understand the position of the decoupling point in the supply chain which is an important indicator for many decisions related to Supply Chain Management, in particular the need for planning processes. The lead time analysis is conducted with the help of interviews with order management, distribution, production and procurement – its purpose is to give a rough overview of the demand and supply lead times in the particular supply chain.

Fig. 15.4 shows an example of a lead time analysis. In this example procurement of materials with a short lead time and assembly operations can be executed order-driven, whereas the remaining supply decisions have longer lead times and – therefore – need to be executed driven by the forecast.

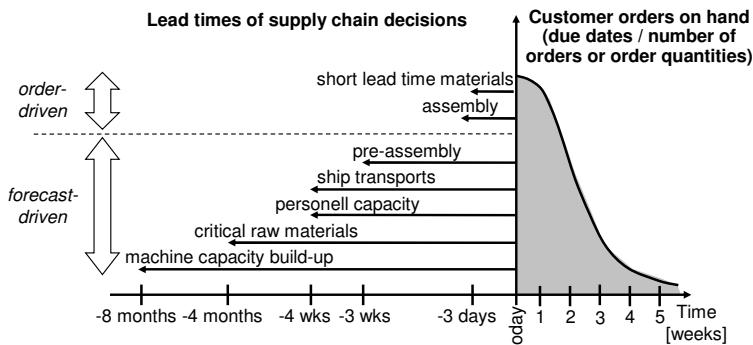


Fig. 15.4. Lead Time Analysis of Supply Chain Decisions

15.1.1 Executive Management

Executive management is an important information source related to strategic issues, cross-functional change programmes and the right target levels of the supply chain.

An example for a strategic decision related to Supply Chain Management is the creation of a new e-business sales channel in parallel to the existing sales channels (Kilger, 2000). This decision will clearly influence the existing sales channels – direct sales via sales representatives as well as indirect sales via channel partners – and hence, has to be managed from the top. Another example is the question whether to enter an electronic marketplace on either side of the supply chain – suppliers or customers. In particular, if some marketplace is open to other companies, it is possible that competitors will participate as well in the marketplace. Thus, this decision must be aligned with the strategic differentiation from competitors and the targeted collaboration level with suppliers and customers, respectively.

Strategic decisions are long lasting decisions and have a big impact on all aspects of the business. They should be taken before entering a change programme, as they influence the general direction and the initial scope of the SCM project. Furthermore, strategic decisions in a supply chain should be centered around a common vision of the supply chain participants (see Chap. 1 and Poirier (1999)).

The second role of executive management is to enable the change of procedures across multiple functional areas and departments. Most planning processes stretch across multiple functional areas. It must be investigated how these planning processes are done currently and how management is enforcing the collaboration between the departments. Examples for cross-functional planning processes are demand planning and master planning. Executive management must break barriers between the departments, bridge gaps in understanding and enforce collaboration between the functional areas.

The third role of executive management is to set the target levels and organizational incentives for the functional areas and the departments. In many cases, the target definitions of the departments are not consistent or even contradictory. For example, procurement is responsible for keeping inventory levels low, order management has to guarantee a high due date performance and production must ensure a high utilization rate of the production equipment. All three have to collaborate in the master planning process and thus have to agree on uniform goals and targets for the master planning process. Executive management must approve the common master planning targets and make sure that all three departments work towards these targets. Conflicting targets are even more likely to occur if the APS project includes multiple partners in the supply chain, e. g. suppliers and customers.

15.1.2 Sales

The sales force is closest to the customers and thus can give input about the behaviour of the customers, the market segments, the demand patterns per segment etc. Sales should be the owner of the demand planning process, resulting in an unconstrained demand plan representing future market demand. Important aspects to investigate are the planning frequency, the planning cycle time, the planning accuracy, the structure of the forecast (along the three forecasting dimensions product, geography and time) and issues like seasonal demand patterns, product cannibalism etc. (for details on the demand planning process see Chap. 7). Also the basic questions *What is being forecasted?* and *Who is giving input to the forecast* must be clarified. Often there is no common understanding of the definition of the forecast – which is a prerequisite for measuring the accuracy of the forecast against actuals and to setup a collaborative forecasting process, with input from sales, product management, production, procurement, management and customers. It must also be clarified who is committing to the approved forecast.

One common trap when talking to sales must be observed. In most companies, there is a central sales organization in the headquarter, and there are multiple sales organizations, e. g. regional sales offices. The sales representatives create the demand plan for their sales region. The regional plans are merged to one uniform demand plan by headquarter sales. When talking to sales it is important to distinguish between these two sales functions. As SCM projects are started by the headquarter, it is obvious that project activities focus on headquarter sales. However, it is very important also to involve the sales regions in the project from the beginning on, as the sales regions are closer to customers, and by that, have better knowledge about the customers' behaviour and expectations. Further, changes in the forecasting procedures directly affect central sales operations and the sales regions.

15.1.3 Product Management and Engineering

Product management is responsible for the definition and the positioning of the product lines on the markets. Engineering is in charge of designing new products defined by product management. One important aspect of product management and engineering is the product life cycle management. Especially in industries where the product life cycles are short as in the high tech industry, product life cycle management directly impacts the performance of the supply chain. At the beginning and at the end of a product life cycle, supply and demand are difficult to predict, often leading to excess inventory and/or unmet demand. As an example consider the assembly of computers. Disk drives being one of the major components of a computer have an average product life cycle of four months – there are three product generations per year. Product management/engineering must align the product life cycles of their own products with those of the disk drives and make sure that all planning processes that are dependent on the supply of disk drives – e. g. demand planning, master planning, order promising and production scheduling – observe the product life cycles.

Secondly, product management and engineering gives input to postponement strategies (see also Chap. 1). Postponement helps to reduce marketing risk. Every differentiation which makes a product more suitable for a specified segment of the market makes it less suitable for other segments (Alderson, 1957). Thus, differentiation has to be applied as late as possible in order to be able to react to demand from a large variety of market segments.

A third aspect of product management are marketing activities, e. g. marketing events, the definition of special product bundles or special product offers etc. Marketing activities may influence the demand plan by creating additional demand. All planning processes related to procurement, production and order fulfilment must be aligned with the marketing activities.

15.1.4 Procurement

All in-bound supply processes are executed by procurement. In many industries the supply lead-time (from procurement to shipment) is greater than the order lead-time. As a consequence raw materials must be ordered based on the forecast. In order to give the suppliers a forecast of what will be procured in the future, procurement forecasts the future purchasing decisions as part of master planning. In practice the following issues related to the supplier forecast are often apparent:

- *Gap between sales forecast and supplier forecast:* Theoretically, the sales forecast is the direct input to the supplier forecast. However, in many companies there are gaps between demand planning and master planning, the latter creating the supplier forecast. These gaps may be due to disconnected information systems or due to communication barriers

within the company's organisation (i. e. sales department and procurement department).

- *No feedback to sales about feasibility of the forecast:* In material constrained industries it is very important to get an early feedback from the suppliers whether they can fulfill the forecasted quantities. Especially in an allocation situation, where material is short in the market, the master plan will be constrained by the supply. In this case, sales should receive information about the supply they can expect (see also Chap. 14 about collaborative planning).
- *No clear representation of supplier flexibility:* In order to represent the supply capabilities of a supplier, a “flexibility funnel” can be defined, specifying – per time bucket – the lower bound of the quantities that have to be purchased and the upper bound of the quantities to which the supplier is bound by the terms and conditions of the supplier contract.
- *Accuracy of supplier forecast not being measured:* The supplier forecast accuracy measures the forecasted quantities against the actually procured quantities. The supplier forecast accuracy is a KPI for the procurement processes, as it steers the production of the procured materials at the suppliers' sites.

Besides the supplier forecast, the number of inventory turns (or average days of supply), the distribution of the inventory age and the on time delivery of suppliers are important KPIs indicating the performance level of the procurement processes.

15.1.5 Order Management

The management of customer orders gets more important as markets get more competitive. The responsibility of order management is to manage and control customer orders throughout the order life cycle, i. e. from the first customer inquiry to the delivery of an order. Order management is responsible for the creation of an initial order promise. Together with sales, order management defines specific allocation policies, allocating the feasible production supply to customer segments (allocated ATP – see Chap. 9).

If the supply, the capacity or the demand situation changes, orders have to be rescheduled in order to get a new feasible promise. In many organizations, orders are not rescheduled even if the situation has changed – leading to unrealistic order promises.

As products on global markets get more and more interchangeable due to comparable quality and features, the reliability as a supplier becomes more important – being measured by the customer service level. The customer service level is measured by three KPIs, the on time delivery, the order fill rate and the order lead-time (see Chap. 2). Besides the customer service level, the order volume, the average number of orders per day and the peak order entry rate are important measures of the order management processes.

15.1.6 Production

In industries with a complex production process and significant production lead-time, one of the most important performance criteria is the work-in-process (WIP) inventory level. Low WIP inventory levels have a positive impact on many related processes and performance indicators (Goldratt and Fox, 1985):

- *Low WIP reduces production lead-time and increases on time delivery:* The production lead-time directly depends on the WIP level. The more material sits in the queue in front of a workstation the longer is the average queuing time, leading to a longer production lead-time. The variability of the production lead-time is increased if the queue in front of a workstation grows. This directly reduces the on time delivery, as it is more difficult to predict the exact production time and to confirm orders accordingly.
- *Low WIP improves quality of products:* In most industries production failures leading to quality problems occur in the early production steps, but are detected at later production stages (usually the testing operations). In order to improve the quality of the products, the quality of the whole process must be improved. If the WIP level is high the average lead-time is also high (see previous item). A long lead-time (induced by a high WIP level) may result in a long time lag between the actual producing operation and the final test operations. Thus, the test operation is reached a long time after the operation being responsible for the failure has terminated. Potentially the whole production process changed in the meantime, and the root cause of the quality problems cannot be determined – preventing an improvement of the process. Thus, the lower the WIP, the easier is the detection of quality problems in a complex production environment.
- *Low WIP speeds up time-to-market of new products:* As product life cycles get shorter, the importance of the time-to-market of new products grows. If the WIP level is high the production lead-time is also high – leading to a longer time to market. Furthermore, the old products that are still in production can often be sold only for a lower price. Thus, lower WIP enables a business to bring new products more quickly to market and to get a higher margin for their products.
- *Low WIP improves forecast accuracy:* The accuracy of the sales forecast depends on the input sales get from customers. In many industries a specific time window exists, and customers give their suppliers visibility of their demand within that window. This “window of visibility” is often derived from the average production lead-time of this industry – and thus depends on the WIP level. If actual production lead-time is below the average forecast accuracy will be high. If production lead-time is above the average forecast accuracy will be low as sales does not get an accurate demand signal outside the window of visibility. This increases

the risk that purchasing will procure the wrong materials, production will start the wrong production orders, WIP levels increase even further.

Besides the WIP level, manufacturing lead-times, excess capacity, bottlenecks in production and sourcing decisions are further potential improvement areas.

15.1.7 Distribution

Distribution can give information about the distribution strategy, distribution and transportation planning processes, merge in transit operations, physical material flows and inventory levels at distribution centres (see also Chap. 12). It is important that these processes are synchronized with the demand (i. e. the customer orders) and with the production supply. One of the main issues found in distribution is the synchronization of the supply feeding multiple order line items that have to be shipped together. If the supply is not synchronized unnecessary inventory is build up, and the delivery of the complete order in time is jeopardized.

15.1.8 Coordination and Integration Technology

One root cause for disconnected planning processes is the extensive use of spreadsheets to support the planning processes:

- Spreadsheets maintain data locally; they do not enforce data consistency and data integrity. Thus, it is highly probable that planners use different data sets, leading to inconsistent planning results.
- Spreadsheets are highly flexible; they can easily be adapted to the needs of the individual planners. However, this flexibility leads to a continuous change of the spreadsheets, making it difficult for others to understand the planning process and the planning results.
- Spreadsheets are stored as individual files, limiting the integration with transaction systems (for loading historic sales, orders on-hand, etc.) and restricting the capabilities to exploit historic data as input to planning.
- Disconnected, spreadsheet based planning processes normally do not consider constraints, leading to planning results without checking feasibility.

Due to the sequential execution of the planning processes based on spreadsheets and the insufficient decision support functionality of spreadsheets planning cycles tend to be long, decreasing the quality of the planning results.

The second important aspect of integration technology is the availability of data (Kilger and Müller, 2002). APS require highly accurate data, including data elements that are normally not maintained within spreadsheets. Even ERP systems like SAP R/3 and Peoplesoft do not maintain data at a level of detail as required for an APS. For example, the detailed product structure and geographic structure as needed by an APS to support forecasting is normally not maintained in spreadsheets or ERP systems. But

also “standard” data like routings and BOMs are often not maintained in a quality requested by an APS – especially if no planning functionality has been employed that would need this data. The precise review of the available data and the data maintenance processes in place are important input to the supply chain evaluation.

15.1.9 Graphical Visualization of the Supply Chain

In order to make the communication with the supply chain experts in the organization more effective, graphical visualization techniques should be employed. Especially the visualization of the material and the information flows of the supply chain helps in the discussions with the various departments and is a good starting point for identifying constraints and/or improvement areas in the supply chain. Additional information to the operation and the material buffer representations of a supply chain flow model can be attached representing specific characteristics like vendor managed inventory, multi-plant sourcing, security stock levels, batch sizes, lead-times etc. If already possible in this step, all constraints in the supply chain should be identified in the model, as well as locations of inventory.

The next step in a supply chain evaluation would be to get an overview of the planning processes, e. g. *sales forecasting, master planning, production planning, distribution planning, detailed scheduling*. A simple process flow notation can be employed, showing sequential relationships between the individual planning processes, the IT systems (decision support systems, transactional systems, ERP systems) supporting the planning processes and the data flows between the IT systems. Chapt. 4 gives an introduction to the various planning processes. The most important item to be checked is the integration of the planning processes. In many organizations, planning processes are performed sequentially and disconnected. Planning results of a former process step are not or only partially used as input to the subsequent steps. This leads to non-synchronized process chains and sub-optimal planning decisions.

15.2 Supply Chain Potential Analysis

Based on the results from the supply chain evaluation and the analysis of the business strategy potential improvement areas are identified and the initial scope of the project is defined. To achieve the improvements and related benefits specific SCM concepts are applied and to-be models for an APS implementation are designed. To-be SCM concepts include

- *processes*: planning, execution, performance measurement,
- *organizational models*: intra-organizational and inter-organizational models (e.g. collaboration mode with supply chain partners),

- *structure of the supply chain:* physical structure of the production and distribution network, and
- *IT support:* support from APS and other IT systems to support the intended to-be models.

The design of the to-be SCM concepts and the required APS-functionality to support these concepts must – even in this early phase – be mapped against the capability of the organizations participating in the supply chain and the project. As Willcocks et al. (2002) observe in the context of e-business initiatives, “people are at the heart of strategic transformation. [...] An essential part of the planning process is a detailed analysis of the current capabilities of the available resources. An assessment of the skills and competencies necessary to deliver and implement in a world where change is continuous and where the contribution of the IT department is measured as much by its intellectual capital as by the reliability of its systems.” The capability of people must be assessed on two levels: On the project level and the operational level. On the project level, the question must be answered: *Do we have the right people and skills to improve business by applying SCM concepts?* On the operational level, the question is *Are our employees capable of operating the new system and work according to the new processes in their daily work?* Both questions must be answered positively before advancing with the project.

The SCM concepts that can be applied to improve business performance are described in detail in Part I; the APS modules to support these concepts are described in Part II of this book. In this chapter we focus on the *benefits* that can be created from SCM in an industrial organization.

15.2.1 Financial Performance Indicators

Following Goldratt and Fox (1985), the goal of an industrial organization (or supply chain) is to be profitable and to improve earnings (defined as revenue minus cost of sales, operating expenses and taxes). Financial benefits can be measured in three ways. *Net profit* is an absolute measurement of making money. However, if we know that a company earns 20 million \$ a year, we cannot tell whether this is a good or a bad performance – as the performance of a company depends on the money that has been invested in the business.

In the business environment at the beginning of the third millennium the performance of a business relative to the invested capital is in the focus of managers and shareholders. The term *shareholder value* is ubiquitous. The *return on capital employed (ROCE)* is “a measure of the returns that a company is realizing from its capital. ROCE is calculated as profit before interest and tax divided by the difference between total assets and current liabilities. The resulting ratio represents the efficiency with which capital is being utilized to generate revenue” (InvestorWords, 2004). The invested capital consists of multiple components, e. g. cash, receivables, inventories, property, buildings, equipment and liabilities. SCM concepts mainly affect the assets,

not the financial components of the invested capital like debts and equity. That is why from a Supply Chain Management perspective the *return on assets (ROA)* is often used as relative business performance measure instead of the ROCE.

The third measurement of the financial performance of a business is the *cash flow*. “Cash flow equals cash receipts minus cash payments over a given period of time; or equivalently, net profit plus amounts charged off for depreciation, depletion and amortization” (InvestorWords, 2004). Cash Flow is rather a short-term measure of a company’s financial health than a long term performance indicator.

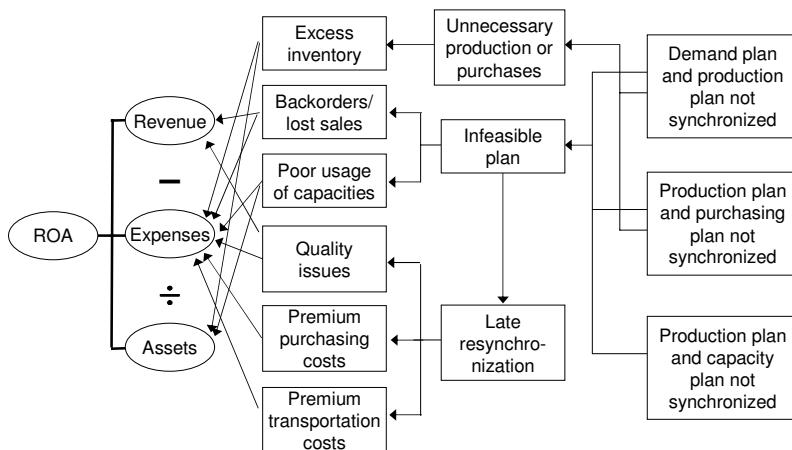


Fig. 15.5. Impact of SCM planning on the ROA

15.2.2 Return On Assets

In the following, we focus on the return on assets as the bottom line performance measure. A common definition for the ROA is as follows (InvestorWords, 2004):

$$\text{ROA} = \frac{\text{Earnings}}{\text{Assets}} \quad (15.1)$$

$$= \frac{\text{Revenue} - \text{Cost of Sales} - \text{Operating Expenses} - \text{Taxes}}{\text{Assets}}$$

Revenue is all the money the customers pay for the offered products and services. Cost of Sales – also called Cost of Goods Sold (COGS) – equals the cost of purchasing raw materials and manufacturing finished products. Operating Expenses are expenses arising in the normal course of running

a business. Assets include all equipment and material that is involved in turning inventory into sales. On a balance sheet, assets are equal to the sum of liabilities, common stock, preferred stock and retained earnings.

In order to evaluate the benefits from SCM, we have to analyze how revenue, costs/expenses and assets can be improved by SCM concepts. Figure 15.5 gives examples of how SCM planning capabilities impact the ROA.

15.2.3 External Variability

Let us illustrate the impact of poor planning capabilities on the ROA by means of an example (adapted from Kilger (1998)). Figure 15.6 shows a supply chain with suppliers, factories, national and regional distribution centres and customers. For example, let us assume that in one factory a machine goes down and due to the service required the machine will be up again in one week. An ERP system would adapt the schedule of that machine accordingly and move out all manufacturing orders that are impacted by the change. But what is the impact of the downtime of that machine on the ROA of the complete supply chain? In order to answer that question the new situation has to be propagated upstream and downstream:

- *Upstream:* Due to the machine downtime of one week manufacturing orders have been moved out and the required raw material will be consumed later. The supplier can peg this material potentially to other customers (factories) and thus make revenue elsewhere.
- *Downstream:* The national distribution centre receiving the finished goods one week later may run into a stockout situation, if the supply that is now delayed is required to fulfil all customer orders on time. This would reduce revenue and increase inventory (assets) and expenses.
- *Planning scenario:* In order to assess whether the plan at the national distribution centre can be re-optimized a planning scenario is created to check whether the national distribution centre may receive the short material from an alternate factory. This potentially can help to ship the orders in time, by that securing revenue.

This example indicates that the performance of a supply chain is to a large extent influenced by external disturbances and external variability. Thus, in order to assess the potential benefits of SCM concepts and an APS implementation, focus should be laid on the impact of external factors on the ROA components – revenue, expenses and assets. It is interesting to note that transaction systems like Enterprise Resource Planning (ERP) systems (SAP R/3, Peoplesoft, etc.) focus rather on the internal processes than on the external factors influencing the ROA. For example, the MRP and production planning modules of an ERP system help to create an initial production schedule (that is often not feasible) and support the tracking of the material flow on the shop floor – but they do not provide simulation capabilities and problem resolution functions to quickly react to external changes.

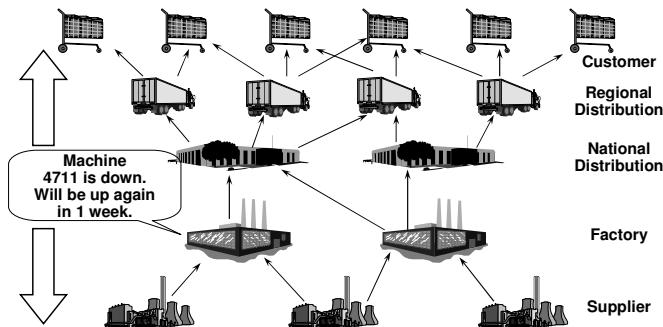


Fig. 15.6. Propagation of changes in a supply chain

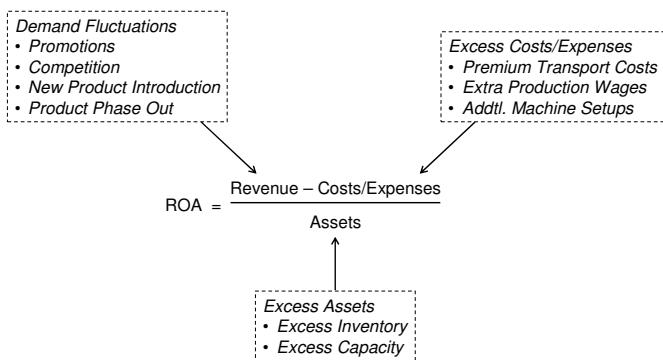


Fig. 15.7. Impact of external variability on the ROA

In the following, we discuss examples illustrating the impact of external factors on demand, expenses/costs and assets (see also Fig. 15.7).¹

15.2.4 Demand Fluctuations

The impact of external factors on the revenue is obvious as sales are generated by external customers. However, there are some specific situations in which it is particularly difficult to predict sales quantities:

- Promotional actions (special product offerings, reduced promotional prices, etc.) may create higher additional demand than planned for;
- competition may attack in certain markets or in specific product areas, leading to an unexpected drop in demand;
- the introduction of new products may be more successful than expected and/or may cannibalize the demand for other products;
- the phase out of products may result in a demand drop of other products, as customers were used to buying both products together.

¹ The examples are partially based on discussions with Sanjiv Sidhu (1999).

Fluctuations of demand directly impact revenue. In addition, they may have side effects on expenses and assets. For example, introducing a product in a new regional market costs money due to marketing activities. If the resulting demand is not properly planned or if it is not properly supplied (as other product/market combinations are more profitable or are prioritized for other reasons), the additional demand will not be transformed into additional revenue – but the additional expenses have been spent.

15.2.5 Excess Expenses/Costs

Expenses are partially controlled internally and partially externally. In every industry there is a cost segment that is determined by the type of production. For example, computers are assembled in a similar way all over the world: get all components you need and assemble them into the housing, test the device, pack it and ship it. The costs implied by this process are comparable across the computer industry and – more important – there is only a marginal impact of Supply Chain Management techniques on the internal production costs. However, there is a big impact of Supply Chain Management on the expenses determined by external factors, being for example:

- Pay premium air freight fare to get material because of late or short supply by the standard suppliers;
- pay extra production wages for subcontractor manufacturing in order to account for peak load situations due to additional demand or delayed production;
- reschedule the production plan (including the need to pay overtime rates for the workers) because of short term additional demand or delayed supplies of material;
- buy critical components on the spot market (e.g. processors, memory) for a higher price compared to the preferred supplier.

15.2.6 Excess Assets

The total value of assets in a supply chain can be split into “base assets” that are required for the production operations and “excess assets” that are being used to shield the supply chain from external variability. For example, excess inventories may exist for raw material, work in progress and finished goods. Excess inventories are used to buffer production from the demand variability of the market, and may lead to increased material costs due to

- price reductions (e.g. the price of electronic components reduces by 2% per week in the average),
- obsolescences (i.e. components that cannot be sold on the market any longer because a better successor has been introduced to the market for the same price),

- stock keeping costs,
- internal capital costs and
- material handling.

By better controlling the volatility in the supply chain, excess inventory may be reduced and business performance will increase. However, please note that some safety stocks will in most cases be required due to uncertainties that are “inherent” to the supply chain and may not be controlled (Chap. 7 gives an overview of safety stock policies). Further, the reduction of excess inventories may lead to an increased risk of capacity shortages and – as a consequence – to excess expenses as described in the previous section.

Excess capacity is built up in order to have sufficient capacity to cover peak load situations. Due to the interdependencies in a production system – a resource can start a production operation only if all preceding operations have been completed and all required raw material is available – the load variability of a resource increases with the number of preceding operations. Thus, load variability is higher the more downstream the resource is located, and – because of that – excess resources are often built up at the end – downstream – of the supply chain (this is typically the test area or the distribution network).

15.3 Project Roadmap

In the preceding section we have shown that external factors and external variability influence the financial performance of a supply chain. Especially demand fluctuations, excess expenses and excess assets have a negative impact on the ROA. SCM concepts supported by APS functionality enable a supply chain to quickly react to external changes and by that help to improve the ROA.

However, having identified improvement potentials related to revenue, expenses or assets does not necessarily tell how to realize these potentials. Which levers exist to create additional demand and transform this into additional revenue? What is the root cause of excess expenses or assets? Which SCM concepts can help? What module of an APS do I need? In case there are multiple things I can do to improve the ROA, how shall I prioritize?

15.3.1 Enabler-KPI-Value Network

In order to answer these questions and to create a project roadmap, the targeted financial improvements have to be related to concrete project activities. The bridge between these financial criteria and the project activities is formed by logistical KPIs. The following steps describe the way to define a SCM project roadmap by a value driven approach:

1. identify improvement potentials based on financial performance indicators (i.e. conduct a supply chain potential analysis as described in Sect. 15.2),

2. transform the targeted improvements of the financial indicators into targeted improvements of logistical KPIs, and
3. map the targeted improvements of the KPIs to SCM concepts and/or APS modules enabling the improvement.

As there are many SCM concepts and a broad range of APS functionalities – as described in detail in Parts I and II of this book – it is very important to start the definition of an APS implementation project from the value perspective, as indicated by steps 1-3 listed above. Starting the definition of an SCM project from the functional perspective bears the risk that the system gets overengineered, i.e. the system would contain many functions that do not necessarily help to improve the business performance.

For example, in the computer industry, order promising and production planning is normally constrained by the material supply and not by capacity. Thus, exploiting the finite capacity planning abilities of APS in order to improve the production scheduling process would not lead to a big business improvement. Following the three steps listed above, one could define the project scope as follows:

1. The main target of the project is to generate additional sales and to increase revenue.
2. Additional sales can be generated by improving the on time delivery. (This value proposition can be backed by industry benchmarks and interviews with the customers; refer to Chap. 2 for additional details.)
3. On time delivery can be improved by an APS by
 - synchronization of purchasing decisions and order promising based on forecast/ATP,
 - creation of feasible master plans considering all constraints and
 - simulation of receipts according to different scenarios / rescheduling of orders in order to improve the short term production plan.

Thus, the focus of the project should be laid on forecasting, master planning and order promising, instead of production scheduling.

In general, the relationships between financial performance indicators, logistical key performance indicators and APS enablers form a complex network. The structure of the network strongly depends on the particular situation of the supply chain, its improvement potentials and the initial scope of the SCM project. Figure 15.8 shows an Enabler–KPI–Value network based on the example given above, connecting APS enablers with logistical KPIs and their relation to financial performance criteria. The arrows in the boxes indicate whether the value of the KPI or financial performance indicator will be increased (arrow upwards) or decreased (arrow downwards). A detailed description of logistical KPIs can be found in Chap. 2. Setting up one or – in complex situations – multiple Enabler–KPI–Value networks defines the framework for improvements, linking SCM and APS enablers with financial performance indicators by logistical KPIs. Each path through that network

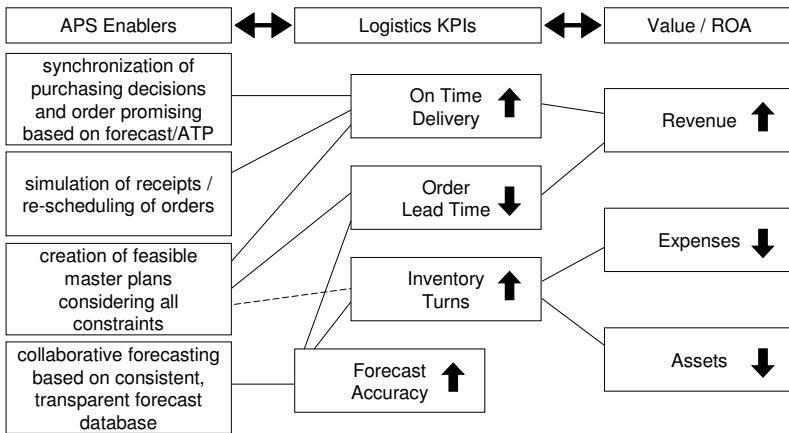


Fig. 15.8. Enabler–KPI–Value Network

represents a logical relationship of an enabler, a KPI and a financial performance indicator. Usually, the enabler has a *positive* impact on the logistical KPI (e.g. reduction of order lead time or increase of inventory turns). In some cases an enabler may also have a *negative* impact on a KPI. For example, the creation of an optimized master plan may result in higher inventory (reducing the KPI inventory turns) if additional inventory is needed to buffer the supply chain against large demand peaks and to achieve a high service level. Negative implication of an enabler to a KPI is indicated by a dashed line (Fig. 15.8). Note that a negative implication of an enabler should be compensated by a greater positive implication on some value component via other paths in the network – in order to assess this enabler as beneficial.

15.3.2 KPI-Driven Improvement Processes

Having identified a collection of logistical KPIs that shall be improved the next step is to further detail the improvement. This is done by setting up a *KPI profile* for each of the KPIs. A KPI profile consists of the following constituents (Kilger and Brockmann, 2002):

1. The first step is to determine the current as-is value of the KPI. This fixes the starting point of the targeted improvement activities and is the base for measuring the success of the project.
2. Then, the targeted to-be value of the KPI has to be set. This gives us the goal we want to reach by the project.
3. In the next step, the time horizon to reach the to-be KPI value is estimated. This can only be a rough estimate as a detailed project plan has not yet been created.
4. From the targeted improvements the enablers of APS that can help to reach the targets have to be determined, as well as additional influencing

factors like process restructuring, reshaping the organizational structures, analyzing high-level data requirements etc. Especially process changes are required in most cases to realize the full benefit as expressed by the to-be KPI value.

5. Based on the as-is value, the to-be value, the estimated time horizon and the considered APS enablers, actual project activities are setup and implemented. It is important to note that each of these sub-projects have to generate business value in a given time period, by applying predefined APS enablers.
6. In order to enter a continuous improvement process, one can go back to step 1 and start the cycle again from a higher performance level.

Note that at this point in time, we are still in the definition phase of the project. The KPI profiles help breaking down the complete project scope into a sequence of sub-projects, each having a clear objective and a well-defined scope. By that we make sure that the definition of the sub-projects is value driven and not driven by the “nice functional features” of an APS – helping to prevent the system to get overengineered. The result of the project roadmap definition phase is a high level project plan, consisting of the identified sub-projects, including preliminary milestones and a first estimate of project schedule, resources and implementation costs. The targeted financial benefits and the implementation costs can be structured along the milestones and the project schedule, resulting into a cash flow series and an initial ROI calculation and business case for the project.

From the APS enablers that are used in the KPI profiles a requirements list for the selection of an APS can be derived. In the next chapter we focus on the selection process of Advanced Planning Systems – with the requirements list being one major input to the APS selection. However, despite the fact that APS are providing advanced planning capabilities that may help to improve business, it is important to realize that additional measures have to be taken to achieve the full business objectives as documented by the KPI profiles. Especially process changes and the provision of additional data for the APS are required in most cases and should be roughly planned already at this stage.

References

- Alderson, W. (1957) *Marketing behavior and executive action*, Homewood, Illinois
- Collins, J (2001) *Good to Great*, New York
- Goldratt, E. M.; Fox, R. E. (1985) *The Race*, Great Barrington
- InvestorWords (2004) www.investorwords.com, July 19, 2004
- Kilger, C. (1998) *Optimierung der Supply Chain durch Advanced Planning Systems*, in: Information Management & Consulting, Vol. 13, No. 3, 49–55
- Kilger, C. (2000) *Supply-based eBusiness: Integration von eBusiness und Supply Chain Management*, in: Dangelmaier, W.; Felser, W. (Eds.) Das reagible Unternehmen, Paderborn, 74–86

- Kilger, C.; Müller, A. (2002) *Integration von Advanced Planning Systemen in die innerbetriebliche DV-Landschaft*, in: Busch, A.; Dangelmaier, W. (Eds.) Integriertes Supply Chain Management, Wiesbaden, 213-234
- Kilger, C.; Brockmann, K. (2002) *Supply Chain Detectives in der High Tech Industrie*, in: Dangelmaier, W.; Emmrich, A.; Kaschula, D. (Eds.) Modelle im E-Business, Paderborn, 139-154
- Poirier, C. (1999) *Advanced supply chain management: How to build a sustained competitive advantage*, San Francisco
- Sidhu, S. (1999) *private conversation*
- Ward, J.; Peppard, J. (2002) *Strategic Planning for Information Systems*, Chichester
- Willcocks, L. P.; Petherbridge, P.; Olson, N. (2002) *Making IT Count*, Oxford

16 The Selection Process

Christoph Kilger and Ulrich Wetterauer

j & m Management Consulting AG, Kaiserringforum, Willy-Brandt-Platz 5,
68161 Mannheim, Germany

Advanced Planning Systems are a relatively new software technology. One of the first Advanced Planning Systems was OPT that was implemented end of the Eighties (Schragenheim and Ronen, 1990; Silver et al., 1998). OPT is based on the *Theory of Constraints* (Goldratt, 1990), stipulating that the constraints of a production system have to be represented in detail in a planning system in order to exploit and to control its performance.

Since these early days, the SCM market has grown tremendously and many new APS vendors appeared on the market, others were acquired by larger software companies. A good overview of the APS software market is given by two recent studies that appeared in 2003 (Laakmann et al. (2003) and Busch et al. (2003)). Further, analysts like Gartner and AMR publish regularly reports on SCM and APS. Kortmann and Lessing (2000) give a detailed overview of the APS market as of mid 1999, including a classification scheme for APS and information on the selection of APS.

Beginning of the year 2004, the APS market is still growing rapidly and the systems get more and more mature. The potential user has a large variety of systems to choose from, and in many cases, a clear indication which system to buy and implement is not at hand. Thus, a systematic approach for the selection of an APS is required. The following four steps provide a guideline and a proven methodology for the selection of an APS:

1. Create a *short list of APS* based on parameters such as supported planning processes, industry specifics, information on the APS vendor companies, license fees and typical implementation time and effort for the APS (Sect. 16.1).
2. Assess the APS on the short list based on the *requirements* that have been collected in the definition phase of the APS project (see Chap. 15). Remove APS from the short list that do not fulfill the major requirements (Sect. 16.2).
3. Setup a detailed implementation plan including a refined estimate of the effort and the timelines for the *implementation and integration* of the APS (Sect. 16.3).
4. Compare the APS vendors based on their *post implementation effort and support model* (availability and costs for user support, service fees, release migration, etc.; Sect. 16.4).

The results from the requirements analysis, implementation and integration planning and the support model are consolidated, resulting in a ranking of the

APS vendors. In the following sections we detail the selection methodology for the selection of an APS.

16.1 Creation of a Short List

In the early phase of the selection process the “strategic fit” of the APS with the targeted planning processes, the industry within which the supply chain is operating (as far as industry solutions are concerned), the budget targeted for the APS implementation project and the planned implementation time may be equally important as the features and functions. The assessment of the APS by these criteria cuts down the number of APS that have to be considered in the subsequent detailed analysis. By that, time and effort for the selection process can be reduced.

16.1.1 Planning Processes

The planning tasks that may be supported by an APS are summarized by the Supply Chain Planning Matrix (see Fig. 4.3 on page 87). During the definition phase of an SCM project (Chap. 15) the planning processes that are to be supported by the APS are specified.

Tab. 16.1 lists a range of APS vendors. The columns represent APS software modules covering the SCP matrix (the software modules are also shown in Fig. 5.1 on page 109). In addition to the modules shown in Fig. 5.1 we included two further processes: Alert Management (Chap. 13) and Collaborative Planning (Chap. 14). The APS vendors included in Tab. 16.1 (and in the other tables of this chapter) were aggregated from two studies about the APS software market:

- The first study is a market survey of Supply Chain Management software conducted 2003 by the Supply Chain Management Competence & Transfer Center (Laakmann et al., 2003). Laakmann et al. (2003) compare 23 APS vendors and give detailed information about the individual modules of the 23 APS, grouped by vendor.
- The second study was originally published by Felser et al. (1999) and Kortmann and Lessing (2000). Parts of this study were updated end of 2001 for the 2nd edition of this book. Based on this study an updated study was conducted 2003 (Busch et al., 2003). Busch et al. (2003) compare 14 APS vendors, structured by nine categories.

From the APS vendors included in the two studies we have selected a range of APS suites that do best cover the modules of the SCP matrix (as listed in Tab. 16.1).¹

¹ Besides the APS described in this chapter Laakmann et al. (2003) include Ablay & Fodi, Logistik World, Retek and T-Systems in their study; Busch et al. (2003) include Demand Solutions and Descartes.

Table 16.1. Planning processes covered by APS modules

	Strategic Network Planning								
	Demand Planning								
	Master Planning								
	Demand Fulfillment / ATP								
	Production Planning & Scheduling								
	Distribution & Transport Planning								
	Collaborative Planning								
	Alert Management								
Adexa ^a	●	●	●	●	●		●	●	
Agilisys ^e	●	●	●	●	●	○	●	●	
Aspen Tech	●	●	●	●	●	●	●	●	
Axxom	●	●	●	●	●	●	●	●	
Baan	●	●	●	●	●	●	●	●	
DynaSys		●	●	○	●	○			
Flexis			●	●	●			●	
GEAC		○	○	○	●			●	
i2 Technologies	●	●	●	●	●	●	●	●	
ICON		●	●	●	●	○	●	●	
Intentia	●	○	●	●	●	○	●	○	
Manugistics	●	●	●	●	●	●	●	●	
Mapics ^b	●	●	●	●	●	●	●	●	
Oracle		●	●	●	●		●	●	
Peoplesoft ^c	●	●	●	●	●	●	●	●	
SAP	○	●	●	●	●	●	●	●	
TXT e-solutions		●	●	●	●	○		●	
Viewlocity ^d	●		○	●	●	●	●	●	
Wassermann	○	○	●	●	●		●	○	

- Functionalities are fully covered
- Functionalities are partially covered

^aformerly Paragon

^bformerly Frontstep and Symix Systems

^cformerly J. D. Edwards and Numetrix

^dformerly SynQuest

All information based on Laakmann et al. (2003), except

^ebased on Busch et al. (2003).

The APS module Strategic Network Planning is covered by many APS suites as indicated in Tab. 16.1. However, in many projects specialized software tools are used to support Strategic Network Planning. The reasons are (1) that Strategic Network Planning is only loosely coupled with the other planning processes of the SCP matrix and (2) that proprietary heuristics for specific planning tasks like investments and transportation cost optimization are applied which cannot be used in other software modules (representation of many binary variables, non-linear cost functions for transportation and non-linear effects of grouping of material flows). The following vendors are

mentioned here as representatives for providers of specific tools for Strategic Network Planning: 4Flow (www.4flow.de), LogicTools (www.logic-tools.com), logsolut (www.logsolut.de), and Prologos (www.prologos.de).

16.1.2 Industry Focus and Experience

The supported industry sectors are an important selection criteria for APS, as most vendors have expertise in specific industries, supporting the planning processes of these industries better than those in other industries (Enslow, 1998). The manufacturing processes, the used terminology, the business rules, the planning processes, the optimization procedures and the reporting requirements strongly differ across industries (Felser et al., 1999). For a number of reasons, APS vendors often focus on one or two specific industries, for example:

- The engineers that are responsible for the design and the implementation of the system already had experience in these industries.
- The first successful implementations were installed in these industries.
- For strategic reasons the APS vendor is focusing on these industries.
- Specific planning features are a prerequisite for specific industries. Unless there are potential clients no effort is spent to include these features in the APS.

Some of the APS vendors launch implementation initiatives for specific industries, trying to extent the scope of their expertise to a new area.

The main improvement areas of an APS implementation strongly depend on the type of industry and the type of the supply chain according to the supply chain typology, respectively (refer to Chap. 3). See Chap. 4 for a describtion of the dependency between industry specific planning tasks and the supporting planning concepts and methods. In distribution intensive industries the main potentials are in the optimization of the distribution and transportation operations, including the deployment of supply and the reduction of inventory. In asset intensive industries major improvements are possible by the optimization of the throughput, the detailed scheduling of the capacity bottlenecks and the reduction of change over time. In material intensive industries forecasting and procurement decisions influence business performance and should be optimized by the APS. Tab. 16.2 gives an overview of the industries supported by the APS vendors.

A remark has to be made related to the number of installations. The procedure to measure the number of installations strongly depends on the APS vendor. Some vendors take only the number of production sites that are supported by their APS, others count all installations of individual APS modules separately, leading to a larger number of installations. Furthermore, some vendors consider any installation, whether productive or in an early implementation stage, whereas others consider only installations where the

Table 16.2. Industry focus and experiences of APS vendors

	Automotive									
	Clothing / Apparel					Retail				
						Semiconductor				
						Electronics / High Tech				
						Consumer Packaged Goods				
						Food & Beverage				
						Logistics Service Providers				
						Aerospace & Defense				
						Machinery				
						Pharma / Chemicals				
						Traffic & Tourism				
Adexa ^a	○	●		●	●	●	○		○	●
Agilisys ^e						●	●			●
Aspen Tech	○	○		○	○	○	●			●
Axxom	○		●			●	●	○		○
Baan	●	●	●	●	●	●	●	●	●	●
DynaSys	○	○	●		○	●	●			●
Flexis	●			○	○		○		○	●
GEAC	●	●	○	○	○	●	●	○	●	●
i2 Technologies	●	●	●	●	●	●	●	●	●	●
ICON	●				●					
Intentia	●	●	●		○	●	●	○	○	●
Manugistics	●	●	●	●	●	●	●	●	●	●
Mapics ^b	●	●	●	●	●	●	○		●	●
Oracle	●	●	●	●	●	●	●	○	●	●
Peoplesoft ^c	●	○			●	●	●	●	●	●
SAP	●	●	●	●	●	●	●	●	●	●
TXT e-solutions	●	●	●	○	○	●	●	○	○	○
Viewlocity ^d	●	●	●	●	●	●	○	●	○	●
Wassermann	●	○			●	●		○	●	●

- Numerous references available for this industry sector
- Limited number of references available for this industry sector

^aformerly Paragon

^bformerly Frontstep and Symix Systems

^cformerly J. D. Edwards and Numetrix

^dformerly SynQuest

All information based on Laakmann et al. (2003), except

^ebased on Busch et al. (2003).

customer has announced that the system is being used productively. Thus, it should be defined precisely by the vendor how the number of installations is being measured.

16.1.3 Information on the APS Vendor Companies

Besides the supported planning processes and the industry focus, information on the APS vendor companies are important for the selection process. Tab. 16.3 lists the following information:

- The year the company entered the APS market,
- number of employees in the SCM field (if this information is not available, the total number of employees is given (in parentheses) as an indication of the size of the company),
- financial information for the years 2001 and 2002,
- information on additional software packages that are provided by the company, i. e. Enterprise Resource Planning (ERP) and Customer Relationship Management (CRM) packages,
- information whether the APS may be hosted by the vendor, i. e. provision of an ASP model (application service provider).

Note that many APS vendor companies went through a series of mergers and acquisitions, making the historic information difficult to interpret and compare.

16.1.4 License Fees

Typically, the size of the customers of an APS vendor also relates to the license fees. Whereas APS vendors with larger customers tend to be in the upper price segment, the APS vendors with small and medium sized customers are more often found in the lower price segment. In many cases, the license fees are determined based on the number of users and the expected business benefits created by the implementation of the APS – measured by KPI improvements as described in Chap. 15. The license fees should match the expectations and the targeted budget of the APS implementation project. However, it is often difficult to get information about the pricing model applied by the APS vendors without entering actual contract negotiations.

16.1.5 Implementation Time and Costs

Besides the supported planning processes, the industry focus and the general vendor information and license fees, the typical implementation time and costs should be considered. From this, an estimate of the required use of internal resources as well as external consultants and experts from the APS vendor may be derived. The best information source to estimate the time and effort are reference projects in the same industry – or in related industries, as direct competitors most probably will not talk about their experiences. The APS vendors should provide a list with references where projects with a similar scope had been completed and set productive. A visit at one or several reference sites is strongly recommended at an early stage of the selection process in order to learn from the experiences that have been made with the APS vendor and its systems.

Table 16.3. Information on APS vendor companies

	Year of APS market entry		Number of employees in SCM field		Revenue 2001			Revenue 2002 ^h								
Adexa ^a	1994	270	44 M €	—	○	○	●	—	—	—						
Agilisys ^g	1969	250	—	—	—	—	—	—	—	—						
Aspen Tech	1981	350	307 M €	306 M €	○	○	●	—	—	—						
Axxom	1996	55	40 M €	—	—	—	○	—	—	○						
Baan	1990	150	—	134 M €	●	●	●	—	—	—						
DynaSys	1986	38	3,6 M €	4 M €	○	○	○	—	—	—						
Flexis	1997	35	—	—	○	○	○	—	—	—						
GEAC	1997	(3400)	—	687 M €	●	●	●	—	—	—						
i2 Technologies	1988	3355	986 M €	908 M €	—	—	●	—	—	●						
ICON	1994	45	3,2 M €	—	○	○	●	—	—	—						
Intentia	1998	(3135)	437 M €	398 M €	●	●	●	—	—	—						
Manugistics	1986	1300	295 M €	304 M €	●	—	●	—	—	●						
Mapics ^b	1996	>300	118 M €	191 M €	●	●	●	—	—	—						
Oracle	1990	(42000)	12,1 B €	10,2 B €	●	●	●	—	—	—						
Peoplesoft ^c	1977	250	804 M € ^e	959 M € ^e	●	●	○	—	—	—						
SAP	1998	(28900)	7,3 B €	7,4 B €	●	●	●	—	—	—						
TXT e-solutions	1994	292	30 M €	42 M €	—	—	—	—	—	—						
Viewlocity ^d	1995	138	36 M € ^f	7,1 M €	—	—	—	—	—	—						
Wassermann	1995	80	621 M €	—	●	○	●	—	—	—						

(n) total number

^aformerly Paragon

of employees

^bformerly Frontstep and Symix Systems

– no information

^cformerly J. D. Edwards and Numetrix

● is available

^dformerly SynQuest

○ not available

^efinancial results for J. D. Edwards^ffinancial results 2001 for SynQuest

All information based on Laakmann et al. (2003), except

^gbased on Busch et al. (2003) and ^hbased on own research.

16.2 APS Requirements

The main result of the definition phase of the APS project has been the detailed requirements list (see Chap. 15). The list can contain more than 100 individual requirements; in order to be able to handle large numbers of requirements these should be grouped according to the planning processes that are in the scope of the APS implementation project. The SCP Matrix shown in Fig. 4.3 (page 87) can be used to identify the planning processes that are

to be supported by the APS, e.g. Demand Planning, Master Planning, Demand Fulfillment and Production Planning and Scheduling. All requirements should be assigned to one or multiple of the selected planning processes.

As APS are a relatively new software technology, most systems are only partially developed with respect to the full functional scope announced by the APS vendors. In many areas, the APS have to be further developed, either by adding additional functionality, resolving issues within existing functionality or better integrating the functional modules. The latter issue – lack of integration – is especially a problem for APS vendors that have acquired another APS vendor in order to include the systems of that vendor into their own product suite. To reflect the coverage of the functional requirements and the plans of the APS vendors to further develop their systems, the following assessment scheme has been developed, consisting of five levels:

- *Level 1:* The functionality is not available; there is no plan to develop this functionality.
- *Level 2:* The functionality is not available; it is planned to develop this functionality in the future.
- *Level 3:* The functionality is partially available; there is no plan to develop this functionality further.
- *Level 4:* The functionality is partially available; it is planned to further develop this functionality in the future.
- *Level 5:* The functionality is currently fully available.

There are three options to evaluate the functional requirements according to these five levels. The easiest and fastest way to get an assessment is to hand over the detailed requirements list grouped by the planning processes to the APS vendors and ask them to provide a self-assessment of their respective systems. For requirements being evaluated to be at levels 2 and 4, a date for the availability of the future development must be provided by the APS vendor; for requirements being evaluated to be at levels 3 and 4, details about the degree to which the functionality is currently available must be provided by the APS vendor.

The second option is to ask the APS vendors to demonstrate the required functionality in a live demo. As this takes more time and effort than the first option – on both sides, the potential customer and the APS vendor – only key functionality should be selected for demonstration. Typically, the second option is combined with the first option: Based on the self-assessment of the APS vendor, critical functional requirements are selected to be shown in a live demo. In order to prepare this, the APS vendor can be asked to state for each functional requirement his ability to demonstrate that functionality, according to the following scheme:

Level A The functionality can be demonstrated with an existing demonstration set with less than 24 hours lead-time.

Level B The functionality can be demonstrated, but requires changes to the standard demonstration models (no changes to the software).

Level C The functionality can be demonstrated at another customer's site.

Level D The functionality cannot be demonstrated easily.

The third option is to implement a prototype, to assess in detail to what degree a specific functional requirement can be fulfilled by an APS. This of course creates additional effort and must be carefully planned. The following issues should be clarified before starting a prototype implementation:

- The scope and the target of the prototype must be clearly defined. Only critical functional requirements and interface issues should be prototyped. For example, the integration of the APS into an existing order entry system can be evaluated by implementing a prototype system.
- A data set for the prototype implementation may be generated or may be extracted from the operational systems, e.g. ERP-systems.
- A detailed project plan for the prototype implementation and a budget (cost and time) must be set up. This includes the decision of what portion of the effort is taken over by the APS vendor.
- In relation to this it must be decided which APS shall be included into the prototype implementation effort. Normally, the number of systems that are prototyped is restricted to one or two. Otherwise, too much effort is invested into development work that cannot be reused in the real implementation project after the selection process.

Based on the prototype implementation(s) it must be possible to answer all open questions that have been included in the scope of the prototype.

The results of the self-assessment by the APS vendors, the results of the system demonstrations and the results of the prototype implementation are summarized in a report on which the selection decision will be based.

16.3 Implementation and Integration

The estimated effort for the implementation of the system and the integration of the APS into the existing IT landscape has to be considered upon the selection of an APS, in order to match budget restrictions.

16.3.1 Implementation of the APS Functionality

The implementation tasks can be grouped into

- the modelling of the supply chain, including the definition of the locations, sites, material flows, operations, buffers, resources etc.,
- the customization of the planning procedures and the optimization algorithms (e.g. the parameters of a scheduling heuristic),
- the setup of internal data structures and databases,
- the realization of organizational changes and
- the training and project management activities.

Typically, APS use specific modelling techniques and representations of the supply chain and employ system specific planning and optimization techniques. Thus, the implementation approach and the implementation effort strongly depends on the selected APS.

Based on the initial estimate of the implementation effort for each of the APS modules that are in the scope of the project, a rough-cut project plan is created. This is done for those APS that are on the top of the short list; in order to keep the planning effort low the creation of rough-cut implementation schedules should be restricted to the top two or three systems. The plans have to account for the availability of the required APS functionality. If one of the vendors has announced that a specific functionality is available at a certain point in time, all related implementation tasks have to be moved out accordingly. In the next step the functional implementation plan is extended by the required integration tasks.

16.3.2 Integration Technology

The integration approaches for APS range from vendor specific integration techniques to standard middleware systems (see Fig. 16.1 for an overview; a detailed description of integration and communication approaches for supply chain planning is given in Chap. 13).

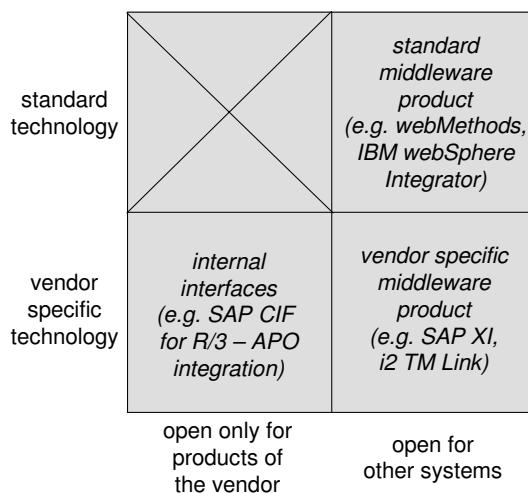


Fig. 16.1. Classification of APS integration technology

As an example, SAP provides a tight integration of their Advanced Planning System APO into SAP's ERP system R/3 via the Core Interface (CIF). i2 Technologies provides its own middleware product TradeMatrixLink that is open to a large variety of other systems including SAP R/3 and relational

databases. SAP also provides an open integration tool SAP eXchange Infrastructure (XI) that is open for other systems. Another option is to use an independent Enterprise Application Integration (EAI) system like webMethods or IBM webSphere Integrator (formerly CrossWorlds).

There are advantages and disadvantages for each of the three integration approaches. Internal interfaces like those between SAP R/3 and SAP APO are the easiest to implement. Base data and dynamic data are transferred between R/3 and APO via the internal interface without the need of further interface implementation. However, this holds true only for data *that is already maintained by the ERP system*. Data that is provided by external systems, for example a shop floor control system, requires extra interface programming. Furthermore, the aggregation of data required to map operational data from the ERP system to a master planning model is only partially supported (see also Chap. 8).

APS vendor specific middleware products are open to external systems. Interfaces between the APS and external systems are customized; programming is normally not required to setup the data transfer. In the case of i2 Technologies' TradeMatrixLink product, so-called *copy-maps* are created, mapping the fields of a data source to the fields of a data target. For example, the source could be the master plan as maintained by the APS, the target could be a table in the ERP system. SAP provides a full EAI system called XI that is open to SAP and non-SAP systems. Standard middleware products and EAI systems provide a similar functionality as APS vendor specific middleware products, with the additional advantage that the system is not proprietary technology of the APS vendor, but is supported by a wider range of applications.

Both APS vendor specific and standard middleware products support the creation of *data interfaces* between a source and a target system very well. Aggregation and filtering rules can easily be implemented on top of these middleware products. But note that data integration is only the first step. In order to fully integrate an APS with other systems, the integration must be extended to the *functional level*. Consider as an example the transfer of the master plan from the APS to some ERP system. The transfer of the master plan into some table of the ERP system is just the first step. The full integration requires that appropriate transactions in the ERP system are invoked to further process the data. For example, demand data could be created in order to drive purchasing decisions within the MRP module based on the master plan (see also Chap. 11).

16.3.3 Integration Mode, Performance and Availability

Besides the integration technology, the integration mode has to be assessed. In general, a *full data upload* into the APS is distinguished from a *netchange of the data* (refer to Chap. 13). The decision whether upload of the full data set is acceptable or whether a netchange interface is required, depends on

the planning processes that will be supported and on the performance of the data load (see Fig. 16.2).

incremental planning procedure	<i>netchange required for incremental plan adjustment</i>	<i>high performance data netchange required</i>
initial planning procedure	<i>upload of full data set appropriate</i>	<i>netchange data upload may be required for performance reasons</i>
	performance of data upload not critical	performance of data upload critical

Fig. 16.2. Integration modes dependent on performance and planning process requirements

If an initial plan is created and the performance of the data upload is not critical, a full data upload is appropriate. If performance is critical or if the planning process incrementally maintains a plan, only the changes of the data have to be uploaded into the APS. Some APS like SAP APO even provide an online interface between the APS and the ERP system. For example, new production orders and changes to existing production orders are continuously transferred to the production planning and scheduling module of APO, and the netchange to the plan is computed by APO. This enables the continuous update of the production plan and quick responses from the APS to the shop floor.

For some planning processes, e.g. order confirmation, not only the performance, but also the availability of the interface and the integrated system must be considered. It may be crucial for the business performance that *every order* gets a quote in nearly real-time, i.e. within milliseconds, even in case of system failure. In order to guarantee a high availability of the order promising system, some APS vendors employ highly available transaction systems like the TIBCO data bus that has been developed for use in highly available, on-line transaction environments as for example the finance sector (Tibco, 2000). The mean time between failures can be used as a measurement for the availability of the integrated system.

16.4 Post-implementation Effort and Support Model

The fourth step in the selection process is the assessment of the expected post-implementation effort and the support model of the APS vendor. The efforts – and costs – that are created after the completion of the implementation can be classified into

- the yearly maintenance and support fees requested by the APS vendor,
- the costs for a release update and the typical frequency of release updates,
- the costs for the system administration and
- the costs for the user support.

Most APS vendors charge a specific percentage of the license fees per year for the continuous support services they provide to their customers. Typically, the yearly support fees are in the range of 15–20% of the license fees. However, the availability of the support centers, the languages in which support can be given and the range of the support services differ. Some APS vendors offer the full range of their services online via the Internet, while others rely more on telephone support. It is especially useful if the APS vendor is able to login remotely to the APS in order to detect and resolve issues.

As APS are still evolving very rapidly APS vendors offer several updated and extended releases per year. According to the guiding rule *Never change a running system* one should not follow every release change immediately. However, some APS vendors offer support services only for the latest release. Thus, the APS of these vendors should be upgraded on a regular basis to the latest release (e.g. every second release). In order to get a rough idea about the effort for an upgrade of the system to a new release other customers of the APS vendor should be interviewed about their experience related to release changes. Especially, the question whether external support in addition to the support of the APS vendor is required or not has to be answered, as external support would require a higher budget for release changes.

Besides release changes, general administration tasks have to be assessed. Examples of these tasks are

- administration of databases used by the APS,
- rollover of a rolling monthly or weekly plan to the next planning cycle,
- administration of the APS servers on operating system level (in most cases, Unix or Windows servers are used),
- extension and/or adaption of the APS, e.g. creation of new reports, installation of new clients, modification of models, user administration etc.

For each of these administration tasks it must be decided whether it will be managed internally or whether the task will be outsourced. In both cases, the skills required and the effort generated have to be assessed for all APS considered.

The fourth post-implementation task that should be evaluated in order to compare APS is user support. In practice a three level support structure is

often setup: First level support is given by so-called *super-users*. A super-user is an especially skilled and trained end user, who is able to receive descriptions of issues from other end users, explain and resolve simple issues and transmit a complete description of a complex issue to second level support. Typically, the super-users have already been members of the implementation team and have supported the APS implementation project in a leading role. Second level support is normally embedded into the standard IT support organization. Some issues, especially those related to system administration, will be resolved there. Internal issues of the APS will be forwarded to third level support, i. e. the support of the APS vendor. APS differ in the tools for issue detection and resolution they provide. Again, it might be useful to ask other customers of the APS vendors about their experience with costs and effort related to the end user support for the product of the APS vendors.

References

- Busch, A.; Dangelmaier, W.; Pape, U.; Rüther, M. (2003) *Marktspiegel Supply Chain Management Systems*, Wiesbaden
- Enslow, B.; Gartner Group (Eds.) (1998) *Evaluating a vendor's SCP market strategy*; Integrated logistics strategies, 24 November 1998, Doc. # DF-06-6980
- Felser, W.; Kilger, C.; Ould-Hamady, M.; Kortmann, J. (1999) *Strategische Auswahl von SCM-Systemen*, in: PPS Management Vol. 4, No. 4, 10–16
- Goldratt, E. (1990) *Theory of Constraints*, Croton-on-Hudson, N.Y.
- Hellingrath, B.; Gehr, F.; Braun, J.; et al. (1999) *Marktstudie Supply Chain Management Software*, Fraunhofer IPA and IML, Stuttgart et al.
- Kortmann, J.; Lessing, H. (2000) *Marktstudie: Standardsoftware für Supply Chain Management*, Paderborn
- Laakmann, F.; Nayabi, K.; Hieber, R. (2003) *Market Survey 2003: Supply Chain Management Software – Detailed investigation of supply chain planning systems*, Fraunhofer IPA Stuttgart, Fraunhofer IML Dortmund, ETH Zürich BWI
- Philippson, C; Pillep, R.; Wrede, P. von; Röder, A. (1999) *Marktspiegel Supply Chain Management Software*, Forschungsinstitut für Rationalisierung, RWTH Aachen
- Schragenheim, E.; Ronen, B. (1990) *Drum-buffer-rope shop floor control*, in: Production and Inventory Management Journal, Vol. 31, No. 3, 18–22
- Silver, E. A.; Pyke, D.; Peterson, R. (1998) *Inventory management and production planning and scheduling*, 3rd ed., New York et al.
- Tibco (2000) *Homepage*, URL: www.tibco.com, Date: July 19, 2004

17 The Implementation Process

Ulrich Wetterauer¹ and Herbert Meyr²

¹ j & m Management Consulting AG, Kaiserringforum, Willy-Brandt-Platz 5, 68161 Mannheim, Germany

² Vienna University of Economics and Business Administration, Institute of Transport Economics and Logistics, Nordbergstraße 15, A-1090 Vienna, Austria

A successful implementation of the selected APS is the obvious goal of any organization that has decided to go for a supply chain project. The first section of this chapter details an approach to ensure the success of supply chain initiatives based on the experience of several realized projects. In the second section, an APS implementation project will be considered from a modelling point of view.

17.1 The APS Implementation Project

As an SCM project affects multiple functional areas like sales, production, procurement, distribution or order management (see Chap. 15) the risks involved in such an implementation are considerable. Many enterprises have experienced spectacular project failures due to a number of reasons, surprisingly few of which have to do with the technology involved. Reasons that show up consistently include:

- The business strategy did not drive process design and deployment.
- The user expectations were not met.
- The time to implement was much longer than expected.
- The cost to implement was much greater than expected.

In the following, a proven approach to ensure the success of supply chain projects is detailed. It provides guidance on the five major implementation phases (see also Fig. 17.1):

- project definition,
- solution design,
- solution details,
- execution and deployment and
- close.

For each phase we will show the necessary organizational tasks and some proven ways to avoid the major pitfalls. The most important deliverables of the involved activities will be mentioned as well.

The company which has decided to implement the APS will subsequently be called *client* organization or enterprise.

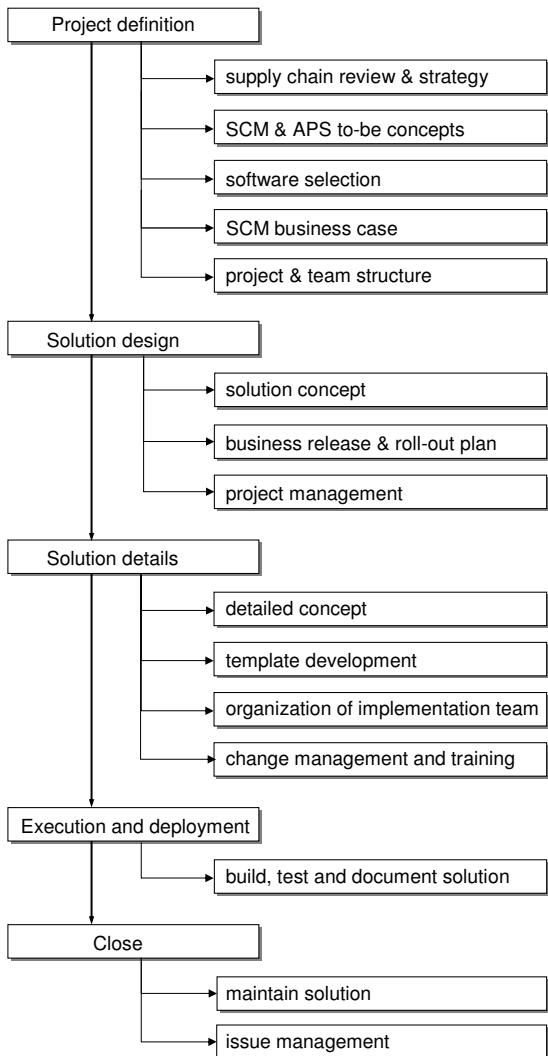


Fig. 17.1. Phases and activities of an implementation project

17.1.1 Project Definition

In the *project definition* phase the company's business vision and SCM strategy have to be related to the drivers for supply chain management, which are benefit realization and cost reduction, and subsequently to the goals of the SCM initiative. Examples for goals are inventory reduction, profit optimization or improvement of customer service.

The main deliverable of this phase is the *business case* that will be presented to senior management. A systematic approach to defining a supply

chain project including the business case methodology is given in Chap. 15 (see also Fig. 15.2). In this chapter the focus is on the main tasks and deliverables for this process.

- **Supply chain evaluation:** The current processes, organization and systems in use are analysed and documented. Possible benefits of a supply chain project can arise in different functional areas like sales, production or procurement and include additional revenues by attracting new customers, reducing inventory and procurement cost, decreasing lead times in production and much more. All major improvement opportunities have to be identified during the as-is analysis process, specified and documented according to the company goals and taken into account in the to-be model development.
 - **Business strategy and vision:** Based on the company's business vision and SCM strategy the scope of the project has to be defined and documented carefully, both for the required functions as well as for the business processes which have to be changed. No amount of advanced technology can offset the problem of inefficient business processes. The time needed for this activity is well spent as it prevents cost overruns due to scope extensions in future stages of the project. In addition, areas which are out of scope should be documented as well to set the expectations for the implementation results and to avoid later discussions about what has to be included in the project.
 - **SCM to-be concepts & APS to-be model:** Solution options and a high-level To-be model are developed based on industry best practices (see also Sect. 2.2.2) and employee suggestions. The project scope is refined and detailed. This activity may include the selection of a suitable software package (for a detailed description of the APS selection process see Chap. 16 and Sect. 17.2). Best practice processes and APS functions are explored to determine the best fit to the future business model. The functions and processes that are to be addressed should be decomposed into lower-level activities (e.g. process: demand management → activities: collect historical sales data, determine forecast proposition, manage forecast entry, achieve consensus about forecast, release forecast to production) to allow the mapping of these activities to the modules and functions of the selected APS.
- Note that the team conducting this phase should avoid to jump right into looking at software functions — the analysis may even yield the result that for major benefits no technology implementation is required at all.
- **Supply chain potential analysis:** Benefit areas are identified and the baseline information is calculated for these benefit areas using a suitable, mutually agreed calculation method.
 - **Implementation plan:** Based on the To-Be model (possibly including several solution options) a transition strategy including project phases and activities is determined and constitutes to a high-level implementation plan.

- **Time-phased benefits:** The benefits associated with the solution options are calculated over the project timeline based on the implementation plan.
- **Time-phased implementation costs:** Direct and indirect cost are determined including resource requirements and risk estimations. The management should realize that supply chain improvements to achieve strategic business opportunities almost inevitably require a redefinition of business processes. Potential changes may address any aspect of the current organization, including process, technology and people.
- **SCM business case:** The assumptions are verified, combined to the business case and presented to the management.

The project definition phase usually requires a combined effort of internal resources which cover the enterprise-specific requirements, specialists with detailed software know-how and consultants which contribute their experience in industry best practices. The management should carefully assess the availability of internal and external skills and knowledge. It has to ensure that all skill gaps are closed already in this early phase. As a suitable project structure and team staffing is a critical activity in every project, these topics will subsequently be addressed in more detail.

The design of the *project structure* requires several activities. A project sponsor and the initial contractors have to be found, the team organization has to be determined and the project control and reporting processes have to be defined as well as the project rules. These topics will be discussed in detail below.

The *project sponsor* must have the authority to make changes happen within the enterprise and to maintain a sense of urgency for completing the implementation activities on time. To implement supply chain management strategies successfully traditional cross-functional barriers and contradictory performance measurements (supporting local optima as opposed to a global optimum, e.g. local capacity usage) have to be aggressively removed. In addition the solution strategy must have the support of senior management and all departmental heads affected. Obtaining and maintaining this support is a major responsibility of the project sponsor.

The initial *contractor* relationships must be established. Consulting firms are usually required to provide resources with experience in best practice processes, software features and project management. Software firms can provide resources with detailed technical know-how. As APS are complex software products, the commitment of the software provider in case of package changes or programming efforts has to be ensured.

The *project control and reporting* processes must be defined, e.g. steering committees, escalation procedures and project management. Clear reporting structures and responsibilities are crucial for the success of the project, especially if several parties are contributing to the project leadership, e.g. different departments or internal and external resources.

Projects can only be executed successfully with an efficient implementation team. In building a team, the technical skills of the team members as well as their characters and organizational needs (e.g. coverage of different departments) must be considered. The structure of the *project team* usually reflects the distribution of responsibilities among the parties involved, i. e. client organization, software provider and consulting firm:

- **Project management:** Full-time resources, both internal and external, are required for project management, quality assurance and guidance. Special emphasis must be put on an efficient integration of the different sub-projects.
- **Team leaders:** Each major process area, usually represented by a sub-project (*functional*: demand planning, master planning or *organizational*: different departments or business areas), requires both an internal and external team leader. They have the responsibility to ensure that all business requirements are covered as well as to supervise the design and configuration of the solutions.
- **Coordinators:** Typically the implementation of an APS radically changes the way in which people do their jobs and interact. With the exception of small pilot projects, successful change projects cross organizational boundaries. The most important team members next to the project leader are therefore the so-called coordinators, full-time team members acting as experts for the enterprise-specific processes. Each coordinator represents a distinct business unit or group affected by the project. Without their participation and buy-in during every single step of the project, an APS project cannot succeed. They have the responsibility to support the design and validation of solution concepts, to improve the communication between the project team and the client organization, to prepare the organization for the necessary changes and, in the end, to achieve the final goal of the project. It is therefore essential to keep the motivation of the coordinators at a high level, by monetary or other means.
- **Functional and process team:** Each functional area's scope has to be addressed by experienced resources. Internal users provide the knowledge with respect to enterprise processes while consultants act as best-practice and application specialists, usually also having an integrative role between the different workstreams. The selection criteria for these team members will be described in more detail in Sect. 17.1.3.

The project management reports to the steering committee which meets on a regular basis, for example every two weeks. The task of this institution is to supervise the whole project based on the project reports, to make decisions about major changes in the project plan and to approve or “sign-off” the project results. The steering committee should be composed of senior management representatives of all organizations and departments involved.

17.1.2 Solution Design

In this phase, the high-level design of the proposed solution is refined and adapted to the selected software in more detail, utilizing the available solution options, if necessary. Key processes and functionalities are validated to identify the potential risks and constraints to the implementation. It is essential that all organizational units which will be affected by the implementation project participate in this task to avoid resistance against the necessary changes. This participation is typically managed by the coordinators mentioned in the last section. Any anticipated constraints to implementing the proposed design have to be assessed. The *solution concept* developed in the previous phase is refined in the three areas *concept*, *activities* and *scope* (see also Fig. 17.2):

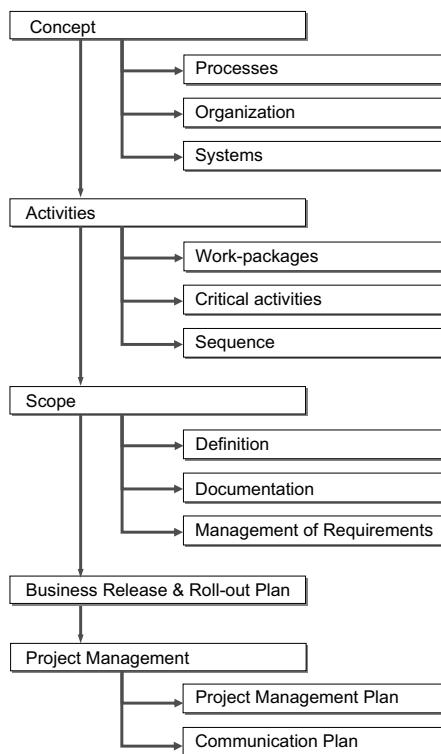


Fig. 17.2. Tasks of the Solution Design phase. Instead of the task *solution concept* the detailed tasks *concept*, *activities* and *scope* are shown.

Concept

The solution has to be designed considering the future processes, organization and systems:

- **Processes.** As far as the processes are concerned it can be distinguished between *Supply Chain Planning*, *Supply Chain Execution* and *Supply Chain Controlling*. Supply Chain Planning refers to the SCP-Matrix as introduced in Chaps. 4 and 5, including the integrative and collaborative processes. As Supply Chain Planning directly affects operative processes like production (e.g. by generating start and completion lists) or order management (by calculation of ATP quantities and due dates), the transactional processes related to SCM (called Supply Chain Execution) have to be considered as well. Finally, KPIs relevant for SCM have to be monitored and analyzed by Supply Chain Controlling to implement a continuous improvement process.
- **Organization.** The realization of SCM processes will likely require some organizational changes, e. g. the shift of responsibilities, the implementation of Supply Chain Planning groups or even the foundation of a new SCM department.
- **Systems.** The system landscape and the development of interfaces and add-ons have to support the planning and controlling processes while balancing the process needs (which might require a deviation from supported software standards) with the related one-time and on-going cost for development and maintenance.

Mapping the refined solution to the selected software typically results in an adaptation of the concept for two reasons. First, every enterprise has its own specialties as far as their processes are concerned, and no APS will be able to cover all details, resulting in functional gaps. Second, to limit the time and cost for the project very often industry-specific, preconfigured templates are implemented (if available, see also Sect. 17.2). Consequently, most companies have to accept compromises and trade-offs, for example a redefinition of the project scope or a change of internal processes.

To identify the functional gaps, the key APS functionalities which are needed to support the solution are validated, for example by building a small pilot using a very limited set of data. It is very important to determine the major gaps in this early phase of the project in order to have a solid basis for the cost and risk estimate.

Activities

The detailed concepts are organized into work-packages or activities which are the basis for the high-level project plan. The critical activities have to be identified and sequenced to determine the critical path. Wherever possible, milestones should be included to enhance the visibility of the project development.

The development of a reliable estimation of implementation cost and duration is a key requirement for successful projects. Gartner estimates that, for 40 percent of enterprises deploying ERP or ERP II systems through 2004, the actual time and money spent on these implementations will exceed their

original estimates by at least 50 percent (Strategic Analysis Report, Zrimsek et al. (2001)). Considering that SCM projects are typically even more complex than ERP implementations due to their cross-organizational character, it can be safely assumed that similar figures can be applied there as well. It is therefore essential that the cost and risk management process is supported by experienced personnel (possibly external consultants) with sufficient know-how in this area based on successful projects.

Scope

Special emphasis should be put on the topic of scope. The scope has to be defined, documented and communicated carefully to limit the expectations to the feasible. How to deal with requirements and expectations will be discussed in more detail later in this section.

Business Release & Roll-Out Plan

To ensure a smooth implementation of the SCM concepts in the *execution and deployment* phase, a suitable business release & roll-out plan has to be developed. A business release denotes a set of functionalities covered by the solution whereas the implementation of a solution or business release is called roll-out.

Business release planning should consider the following approach:

- Start small and simple with transparent standard supply chain processes.
- Develop functionalities for less complex regions first.
- Increase functional complexity and automated planning after learning phases.
- Avoid non-beneficial functional enhancements by stringent scope management.

The roll-out strategy typically has to cover regional areas and/or functional aspects, e.g. roll-out of a business release first to Italy, then to France and Eastern Europe or roll-out of the demand planning and master planning module of an APS (first business release) in a first stage and of the detailed scheduling module (second business release) in a following stage.

Project Management

The implementation plan (including all tasks related to concept, activities and scope) is completed by adding the efforts needed for project management and for a suitable communication plan. Because of their importance these topics will be discussed now in more detail.

Experience has shown that many SCM projects exhibit the same characteristics: budget overruns, missed deadlines and failed organizational expectations. Not all of these can be related to project management problems, but too many can. The importance of project control can be appraised by the multitude of literature that exists about this topic (see e.g. Kerzner (1997)

and Meredith and Mantel (2000)). In this section we will give a general introduction and highlight some of the major pitfalls.

Project management activities must be planned to span the complete live time of the project. The control and reporting procedures are essential for the effective management of any project, regardless of size or scope. Key aspects of any reporting procedure are deadlines and an *early warning* capability. The procedures must effectively deal with project work progress, issues and risks. The main project management activities in order to minimize cost and risks are

- to structure the scope,
- to plan and control the project activities,
- to organize the resources and
- to assure quality.

To *structure the scope* means to divide the project scope into parts and sub-parts and to assure the integration between the different workstreams. This is especially important for SCM projects where several initiatives are realized simultaneously. An example is the implementation of a SCM concept including demand planning and master planning where the result of the demand planning process is the starting point for the master planning. In such a common scenario the project would typically be organized into the two workstreams “Demand Planning” and “Master Planning”. The project management team has to assure that the solution for the master planning process considers the demand planning concepts and the timelines of this workstream.

The main tool to control a complex project is the *project plan*. It should consist of a project master plan for the whole project, with a limited level of detail, and detailed project plans for the different sub-projects (e.g. Demand Planning and Master Planning in the example above). It is necessary for the project plan to be broken into easily definable phases. Regular updates are mandatory tasks for the project managers. Experience has shown that weekly project management meetings are required to keep the different parts of the project under control.

There are several ways to *organize a project plan*, typically involving software tools like Microsoft Project or Excel. Every project manager should consider the difficulty to change the way the advancements of the project are controlled and should therefore carefully choose a suitable kind of method for himself. Most methods are based on monitoring the critical path, although the major problem, the combination of unexpected delays and dependencies of tasks, is often not addressed with the right emphasis. Given the nature of implementation projects, unanticipated tasks will come up that must be completed without revising the final deadlines. Therefore buffer times included in the project plan have to be used with a great sense of responsibility and must not be wasted, otherwise delays in the project time lines are unavoidable (see Goldratt (1997) and Leach (2000)). Communicating this fact has a top

priority for the project leaders, especially with an inexperienced, temporary project team.

The project plan has to consider the required resources for the different implementation phases. Typical implementation phases are

- the creation of an enterprise-wide template,
- the roll-out of this standard solution at a pilot site, validating the standard concepts and
- the roll-out of the solution to all sites including site-specific adjustments and enhancements.

To ensure the availability of these resources is the responsibility of the project management team.

Quality assurance is based on periodical reviews of the concepts, on the implementation of approval processes and on proactively looking at potential risks. The expected result of all these project management activities is a reduction of cost and risks for the project.

It has been mentioned above that, next to the project management processes, a proper *communication plan* has to be installed to ensure the success of the project. All goals and expected benefits should be communicated to the relevant people, starting with a carefully organized kick-off workshop to create an atmosphere of anticipation and motivation. These activities have to be considered in the cost estimation.

Although the communication of the goals and expected benefits is important, this is not sufficient to create the atmosphere which is required to successfully implement SCM processes. Additional trainings are needed to explain the basic concepts of SCM to the key users in all areas affected by the project. It is essential to create this acceptance, commitment and enthusiasm in the team, in the environment of the project (production planning, order management, sales etc.) and in the supporting management (department managers etc.) in a very early phase of the project by applying the following principles:

- Let the participants experience the benefits of the SCM concepts, e.g. via interactive simulations.
- Demonstrate the (basic) functions and features of APS.
- Provide a clear understanding of the risks involved, especially for the project team.
- Focus on acceptance and commitment rather than on mere knowledge.

The cost for these activities have to be included in the cost estimation as well.

In addition to the kick-off meetings and trainings, periodical workshops with the users should be organized to show the progress of the project and to preserve and improve their commitment. Especially in long-term projects

people tend to lose focus on the goals and expected benefits which might result in discouragement and even resistance.

In every SCM project unexpected issues will arise which cannot be solved by the implementation team itself. Examples for these issues are serious software bugs or unexpected resistance within the organization. In addition to project management it is therefore essential that an issue management process is established, clearly understood by the project management team and then implemented during the early stages of the project. The procedures to be defined are analysis, assignment of responsibility, tracking and resolution.

A very important aspect of issue management is to deal with requirements and expectations. Unrealistic expectations, for example concerning the difficulty of implementing a new concept, and a short-term focus can lead to a shift from a planned implementation to “quick fixes” that do not solve the fundamental business problems. These conditions in combination with the lack of a formal process for defining business requirements often lead to a loss of focus and scope creep, thus drastically increasing the implementation time. This can be avoided by rigidly using a formal process to incorporate user requirements or change of requirements, preferably using the coordinators to filter out less important requests.

The implementation plan is combined with the benefit, cost and risk estimation to form the business plan which is the basis for the final proposal presented to senior management. After approval, the senior management should demonstrate and communicate its commitment and buy-in to the proposed high-level solution throughout the organization.

17.1.3 Solution Details

In this phase the details of the proposed solution are defined and software templates are developed, if appropriate. The project plan is refined and a detailed description of the work packages necessary to complete the project is prepared. This also includes the roll-out and training plans for the sites and users as well as change management activities. The packages are assigned to the required team members and the resulting activities are scheduled considering the availability of the resources.

Detailed concept & template development

The concepts from the previous *solution design* phase are reviewed to gain a full understanding of the implications that the implementation of the solution will have on the affected units and on the organization as a whole. It has to be ensured that the available functionality of the selected software is applicable also on a detailed level, although this should have been tested already in the selection phase by the use of mini- and maxi-prototypes (see Sect. 17.2). All functional gaps have to be identified in this phase.

Typically a *template* covering the standard processes is developed including system customizing and the resolution of functional gaps (see also

Sect. 17.2). An example for this approach is the implementation of a detailed production planning system for a multi-site company. The standard processes which are used in all plants would be considered in the template whereas site-specific enhancements would be developed in the next phase, *execution and deployment*, during the roll-out of the solution.

Organisation of implementation team

Team staffing is not required in full at the start of the project but typically ramps up in this and in the next phase. If some members are not assigned full-time to the project, this must be considered in the project plan. Even for full-time resources no more than 80% of the available work time should be planned to allow for travel time, administration and vacation. It is essential that the coordinators are assigned full-time as they should have a start-to-end responsibility for the success of the project.

Experience has shown that SCM projects are typically staffed with roughly an equal share between internal and external resources. The selection and provision of internal project team members is an important step. External experts are essential to provide experience and know-how, but only those who live within the organization can carry the project to a successful end. As the implementation of an APS is an inter-disciplinary effort, the criteria which should be applied to select the right internal people for the core team include:

- experience: All critical aspects of the project should be covered, e.g. sales, product management, order management, production planning, IT services etc. People with influence in the key areas will be very valuable.
- skills: Required are advisers who know the business very well and internal consultants who build up know-how which remains in the client company after the project is finished.

For external personnel who are to participate in the project, a similar scheme of criteria should be applied. The requirements include

- experience: in project management, in change management, in best-practice processes and with the software product, and
- skills: in programming and customization of the software product and in the development of requirement specifications, roll-out activities and training of users.

Although it is not possible to staff every project exclusively with experienced people, especially for long-term projects with the inevitable replacement of team members it is important to insist on and control a certain level of experience and skills of the internal and external project members.

Change management and training

An important part of any SCM project where external resources are typically employed is the area of change management. The implementation of an APS

almost inevitably requires change to an organization's structure and culture. The dynamics of change processes which have to be addressed during the realization phase are listed in the following (for further information about change dynamics related to processes and teams see e. g. Hayes (2002), Belbin (2004) and DeMarco (1999)):

- *Shock*: Confrontation with an unexpected event or environment.
- *Refusal*: No acceptance of the need for changing the own behavior to react to the changes.
- *Rational understanding*: The need for change is recognized, but the willingness to change the own behavior does not yet exist.
- *Emotional acceptance*: New chances and risks are identified and the necessity for change is accepted.
- *Training*: Readiness for training and to change the own behavior. New forms of behavior are tested.
- *Knowledge*: Gained experience helps to decide what behavior fits best to the according system.
- *Expertise*: The new behavior is fully integrated in the daily work, accepted and evident.

In addition to the usual problems (resistance to change in general, satisfaction with the status quo, threats to job security and career objectives, etc.), there are two more barriers specific to the implementation of an APS: The acceptance of *automation* and the *shift of responsibility*.

APS are based on problem solvers and optimization algorithms which help to rapidly respond to changing conditions by automatically generating proposals and alerts or even by automated decisions. This has an impact on the daily work of sales people, production planners and other people concerned with the planning process, as the responsibility for a successful planning shifts from these people to the software tool (and indirectly to the people concerned with the maintenance of basic data).

An example for this is a scenario where a production planner of a plant now has to trust the production plan of an APS, thus only reviewing and solving the problems indicated by the software (usually via alerts or messages). The part of the production plan without problems might get directly transferred to the execution. The planner has to assume that the software calculations are correct (which is generally the case) and that the foundation of this calculations, the input data, is accurate (correct lead times, yield figures etc.). This is typically only the case if the people responsible for this input data know about their influence on the overall planning process.

Resistance against the planning tool is an obvious consequence. This problem can only be solved using an appropriate change management approach (communication plan, involvement of employees, rewards and recognition etc.).

There is another aspect to the shift of responsibility, from local to central, which has to be addressed as well. Typically SCM processes require a central

planning organization, for example in the areas demand management (consolidation of forecast from different sales organizations, central management of allocations) or master planning (coordination of material and capacity constraints across the supply chain). As a result the scope of the local planners becomes restricted, which is usually not appreciated by the people affected.

The availability and quality of basic data are further major problems in APS projects. An APS has more extensive requirements to the quality of basic data than the old processes and legacy systems which have evolved on-site and which are therefore more adapted to the current basic data situation.

As far as the availability of basic data is concerned, the project team will face the problem that SCM is executed across the borders of departments (or divisions or companies). This typically involves the integration of data from diverse application systems on different databases running on multiple hardware platforms. The consequence is a common situation: The people needed to maintain the basic data do not have the overall responsibility. As a result improvement of data quality is a slow and painful process.

Lack of basic data or poor data quality inevitably leads to delays in every stage of the project plan: software development becomes very difficult, professional tests of software releases are almost impossible and a productive use of the final solution is unlikely. To avoid the pitfalls associated with basic data, the process for basic data maintenance has to be revised and, if necessary, has to be set up right from the start of the implementation project.

The requirements for basic data have to be communicated to all relevant people within the organization. In general, the activities connected with the communication plan have to be continued and intensified. These include newsletters, workshops and preliminary trainings to make the concepts of SCM as well as the selected software functionalities accessible to the users.

17.1.4 Execution and Deployment

In the *execution and deployment* phase the key components of the detailed solution are constructed, tested and documented. This includes software development and customization, implementation of best practice processes and user training. The template designed in the *solution details* phase is enhanced to include specific organizational requirements and eventually rolled out to the different sites. To limit the time and cost for this phase, it is essential to establish and retain the following success factors:

- focus on the objectives and benefits,
- limit the implementation to the predefined scope,
- show constant support by senior management and
- ensure effective communication between everyone in the project.

The complexity of APS projects is the reason for one of the most dangerous pitfalls in this phase: scope creep or, in other words, loss of focus. This tendency to model and implement every detail of emerging user requirements,

in contrast to the approach based on the carefully designed solution defined in the previous phases, leads to drastically increased implementation times or even to the failure of entire projects.

The only way to avoid scope creep is to install a rigid change-request-management process with the objective to validate every new user requirement and to reject (or at least postpone to a later release) the ones that are not critical for the success but mere enhancements. Only the part of the requirements that still remains after this process has to be developed and included into the model.

The development activities have to be supported by a well designed testing environment and a defined test base for the validation of software releases and ongoing enhancements. Especially with regard to a final approval by the client management staff it is necessary to implement a formal test-plan management system with a sign-off process.

The IT infrastructure typically includes:

- a development system,
- a test and training system,
- a quality assurance system and
- a productive system.

The hardware for each of these systems has to be configured to allow sufficient performance, even for the development system. The processes to transfer functional developments from one system to the other, e.g. from the test environment to a quality assurance system and subsequently to the productive system, have to be designed carefully in an early stage of this phase.

To avoid problems during later stages of the project, it is necessary to set up a sufficient documentation system as well as to insist on a complete and precise documentation from the start. A professional *document and knowledge management system* supports the implementation efforts as well as the development of training materials and, in the next phase, the setup of a support and maintenance organization. Although this statement seems trivial, it is much too often ignored, especially in the first phases of the project where the complexity is still limited and the need for a rigid documentation as well as for an extensible documentation system is not yet coercive. In addition to the technical documentation the minutes of every important or official meeting should be maintained in the documentation system as a future reference.

The user training is based on the project documentation developed in the phases *solution details* and *execution and deployment*. It has to address the to-be processes as well as to cover all software functionalities required for the daily business of the users and should include hands-on exercises using a training system. In addition the training team should provide support materials such as desk reference manuals and self paced training exercises. The user training has to be performed timely and with sufficient effort, especially in APS implementation projects, as insufficient knowledge transfer can impede the progress of the desired business solution.

As mentioned in the *solution design* phase the work progress as well as the budget have to be monitored and controlled carefully. To maintain and increase the acceptance for the project within the organization it is important to communicate all success stories, goals achieved and milestones reached on a regular basis as part of the overall communication plan.

17.1.5 Close

This is the period late in the project life cycle during which the post-implementation processes are planned and organized:

- maintenance of IT environment,
- maintenance of solution and
- issue management.

The maintenance of the IT environment (e.g. productive system, quality system, interfaces etc.) can be carried out by the IT department of the client organization. Alternatively, an outsourcing solution can be considered.

The team responsible for the maintenance of the solution (i. e. the functionalities covered by the APS system as well as the stability of the software itself) should already be established at the end of the *execution and deployment* phase. It should consist at least in part of experienced team members who have participated in the implementation of the solution. This team should also manage all issues which might arise after the go-live of the project, for example user requests, bug-fixes or performance problems.

The documentation has to be finalized and signed-off by representatives of the client, typically the coordinators.

Ultimately, the measures for the performance benefits of the business solution have to be installed and the solution has to be officially approved by top management, closing the implementation project.

17.2 Modelling Phases of an APS-Project

The last section, introducing the five major implementation phases, was primarily focusing on project management and change management issues. In the following, an APS implementation project will be considered from a *modelling* point of view. The different types of models will be sketched that implicitly and explicitly represent the supply chain and its planning system during the implementation life time. A “guideline” for modelling and integrated planning with APS will be given by referencing the chapters of this book that are related to the respective modelling phases.

17.2.1 Major Phases

As a quintessence of Sect. 17.1 one can state that two decisions of general principle have to be made in an APS implementation project: The first one is to check whether a computerized support of planning is useful and necessary at all and — in case it is regarded beneficial — to decide that the APS market should be evaluated thoroughly. The second one is to select a certain APS for the implementation. With respect to these decisions, from a modelling perspective three major phases of an APS implementation project have to be distinguished:

Evaluation phase: In the evaluation phase a concept for the company's (client's) future planning activities has to be developed *independently of a particular APS*. This concept gives a first impression of the crucial planning tasks and coordination links.

Selection phase: In the selection phase there is a common belief that an APS might be helpful for supply chain planning. However, among the many systems on the market the one has to be chosen which best fits the company's needs. From a modelling perspective, it is essential to evaluate the planning capabilities of *each potential APS* as carefully as possible in order to recognize functional gaps *before* an APS is bought. As Chapt. 16 has shown, there are several options for evaluating the functionality of an APS, differing with respect to the evaluation time, the evaluation costs and the reliability of the insights gained. In the following, we will concentrate on the most reliable option, the prototyping, which should only be executed for one or just a few "hot candidates".

Introduction phase: In the introduction phase the decision for a single APS has already been made and this decision usually will not be revised because of the high investment costs necessary. Thus, for a given APS executable supply chain models have to be designed which support the long-term to short-term planning tasks introduced in Fig. 4.3.

Note that only at the end of the introduction phase "executable models" of the supply chain exist. Supply chain models of the preceding evaluation and selection phases are not designed for a final application. They represent preliminary, aggregated views of the supply chain and merely support the decision processes concerning the introduction of an APS in general and the selection of a particular APS. Nevertheless, after these jobs have been accomplished successfully, the insights gained will not be lost but can provide guidelines for the final implementation process.

17.2.2 Steps of Supply Chain Modelling

Figure 17.3 shows the various models and modelling steps that lead to the final application of an APS. The models are assigned to the three major phases mentioned above. During the following discussion of the respective

modelling steps a reference to the corresponding book chapters that offer more detailed descriptions of the modelling processes is given.

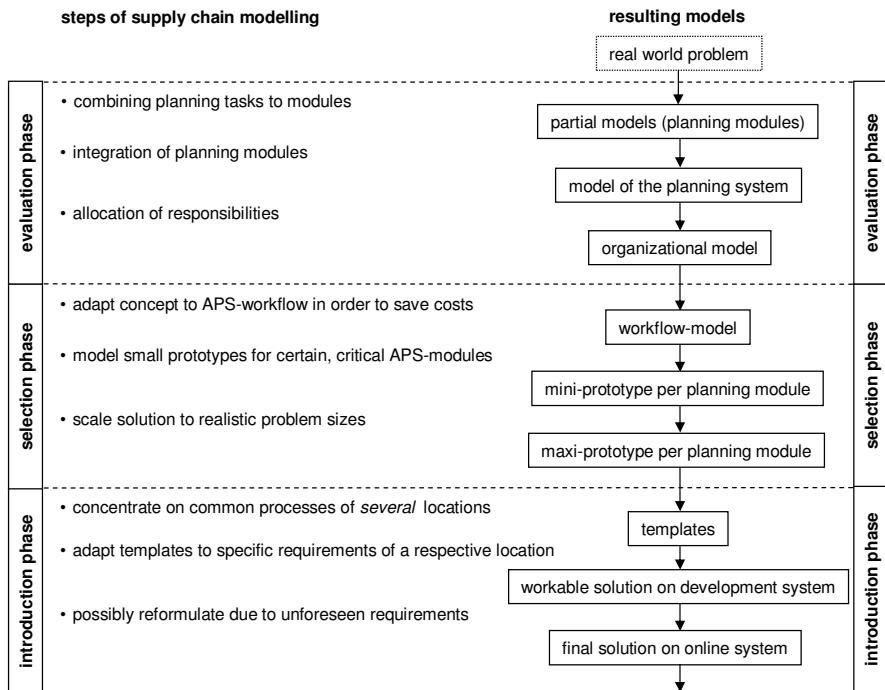


Fig. 17.3. Modelling phases of an APS implementation project

In a first step the main planning tasks with a high potential for improvement have to be identified. The ones that show strong interrelations and a similar planning horizon have to be considered simultaneously and should be combined to manageable planning modules, each of them being in the responsibility of a specific planner or planning department (see Chap. 4). Since supply chain planning occurs on several planning levels with different planning horizons, the resulting planning modules are merely *partial models* of an overall planning system, in which the individual planning modules are in a (weaker) mutual interrelation.

These planning modules have to be coordinated so that the supply chain as a whole shows the best possible performance and the highest possible degree of integration. In order to achieve this coordination, various information has to be exchanged between the planning modules (see e.g. Fig. 13.1). For example, directives of higher-level, coordinating planning modules (e.g. Master Planning) are sent to lower-level, coordinated planning modules (e.g. Production Planning or Distribution Planning) or feedback of lower-level modules is sent the other way round. Therefore, a *model of the planning system*

(i.e. of the planning modules and their mutual informational links) has to be developed which defines the basic data flows between the different planning modules. Guidelines for installing these information flows by means of hierarchical planning are given in Chap. 4. At this early stage of a business process re-engineering project important aspects of planning, like the technical practicability, can only be assessed to a certain extent. Thus, the resulting model is a high-level design of a planning system that has to be adapted in later stages though its essential features should remain unchanged.

In order to prevent potential solutions from being excluded too early and to ensure that an “ideal” planning system can be developed, existing organisational structures have been assumed to be changeable during the first two modelling steps. Only in the subsequent step, the model of the planning system should be adapted consciously. Within this third modelling step responsibilities for the individual planning tasks and planning modules have to be assigned to the existing or future organizational departments. All in all, the resulting *organizational model* of the planning system may deviate significantly from the preceding rough planning concept. Note that hierarchical planning, as introduced in Chap. 4, can also respect such organizational constraints, and thus, can be used to derive and compare both the “ideal” and the organizational model of the future planning system.

This organizational model is the basis to decide whether the *selection phase* is started or not. In case it is, the following three steps of SC modelling should be executed for only the few APS that are short-listed.

In order to save time and money during implementation, APS vendors sometimes provide (industry-specific) “workflows” for their systems, i.e. planning concepts with pre-configured data flows between the software modules. These planning concepts are designed to meet the requirements of “typical” companies of a respective industry or line of business. To be generally applicable they also are, as far as possible, independent of the organizational structures of different companies. A basic way to generate such workflows has already been sketched in the course of this book (see Sects. 3 and 4.3). To give an example, Fig. 18.3 shows a workflow from the APS vendor Peoplesoft for the process and consumer goods industries. If in the selection phase an APS vendor is tested which offers such a workflow for the company’s type of supply chain, it has to be reviewed how the organizational model had to be adapted to the pre-configured data flows of the workflow. In case the gap between the originally desired organizational model and the “*workflow model*” (resulting from this adaptation) is limited, the existence of a time- and cost-saving workflow is an indication to select the respective APS vendor.

The expenditures in cost and time forbid to implement the complete workflow model physically during the selection phase and to test the complete planning system by means of a prototype. However, some presumably critical planning modules can usually be implemented as prototypes, either by the company itself or by the APS providers. This is necessary in order to

check whether a software module is capable of representing all functional requirements appropriately and to test whether high quality solutions can be achieved within an acceptable time frame. In case of failures, both the “modelling gap” and the “solution gap” constitute the functional gap introduced in Sect. 17.1.2. In order to identify the modelling gap of a software module, a small and easily solvable “*mini-prototype*” of the corresponding planning module should be implemented whose structure shows all (presumably) critical features of the desired final application. If some of the requested features cannot satisfactorily be represented, it has to be decided whether an adapted model would also be acceptable. It should be mentioned that the willingness of an APS vendor to introduce a missing function in a new software release is typically very limited (see Chap. 16).

Solely after the basic structure of the *mini-prototype* has been verified, the model should be scaled to a realistic, i.e. practically relevant, problem size which may reveal potential solution gaps. If a solution gap exists, a re-modelling, e.g. by using general principles of hierarchical planning like aggregation and decomposition (see Sects. 4.1 and 8.2), may be helpful. In the best case, the “*maxi-prototype*” that results after scaling and re-modelling increases the solution time and/or decreases the solution quality only slightly. In the worst case, it has to be recognized that the real-world planning problem cannot satisfactorily be solved by the software module tested. As Chap. 27 shows, this can happen e.g. by increasing the number of integer variables of a Mixed Integer Programming problem. In order to limit an investment in the wrong software, a careful test of the solution capabilities is already important in the selection phase, *before* buying a software module.

The *maxi-prototypes* of critical software modules help to estimate the risks and costs of an implementation of the short-listed APS. At the end of the selection phase, it has to be decided whether an APS should finally be used, which APS should be chosen and which software modules should be employed to support the various planning tasks. The following *implementation phase* can also be subdivided into three modelling steps.

If a planning and software module can be used for similar purposes at several sites, in order to save implementation time and costs, it may be useful to create a standard *template* that can be used as a basis for the implementation at all of these sites. This template has to subsume as many common features of the various sites as possible.

As a next step the templates are installed at the various sites. At each site, *workable solutions* are created on a development and quality assurance system that are used to prepare and test the finally aspired operational solutions. In order to build the workable solutions, the templates have to be adapted to the particular requirements of the respective sites. There is a general trade off to be balanced regarding the use of a template. On the one hand, using a template saves time and money of installation. On the other hand, a workable solution resulting from adapting a template usually differs from a solution

that would be tailor-made for a respective site. Thus, for each site it has to be decided whether the savings gained by a template are worth the compromises that have to be made in the operational use later on.

Finally, the APS has to be connected with the (in most applications) already running ERP system or with other OLTP systems (see Sect. 13.2). If the testing has been done thoroughly, the workable solution can be transferred directly to the operationally used online system. However, if undesirable surprises occur like an abnormal behaviour because of missing or obsolete data or an (despite of all tests) insufficient scalability of hard- and software, the *final solution* has to be adapted to such new requirements. During the regular use, a continuous monitoring and controlling of KPIs (see Sect. 2.3) is necessary in order to evaluate whether the planning system finally installed yields the desired effects.

The sequence of the three major phases “evaluation”, “selection” and “introduction” is pre-determined by the decisions to be taken. The proposed sequence of models within a major phase is a proven but certainly not the only useful one. Of course, some of the individual modelling steps can also be combined into a more comprehensive overall model.

It should be noted that modelling skills remain important even after the final solution has been installed successfully. They are, for example, needed when a new product is to be launched or when a new technology has to be introduced that fundamentally changes the underlying planning requirements. Thus, it has to be checked regularly, whether the currently applied models of the supply chain and the currently applied planning models are still up to date or have to be re-formulated.

References

- Belbin, R. M. (2004) *Management Teams*, Butterworth-Heinemann
- DeMarco, T.; Lister, T.; (1999) *Peopleware: Productive Projects and Teams*, 2nd ed., Dorset House
- Goldratt, E. M. (1997) *Critical Chain*, North River Press Publishing Corporation
- Hayes, J. R. (2002) *The Theory and Practice of Change Management*, Palgrave Macmillan
- Kerzner, H. (1997) *Project management: A systems approach to planning, scheduling and controlling*, 6th ed., New York et al.
- Leach, L. P. (2000) *Critical Chain Project Management*, Artech House
- Meredith, J. R.; Mantel, S. J. (Jr.) (2000) *Project management: A managerial approach*, 4th ed., New York et al.
- Zrimsek, B.; Phelan, P.; Karamouzis, F.; Frey, N. (2001) *Gartner Best Practice and Case Studies: Estimating the Time and Cost of ERP and ERP II Projects — A 10-Step Process*, available at www.gartner.com, Date: July 19, 2004

Part IV

Actual APS and Case Studies

18 Architecture of Selected APS

Herbert Meyr¹, Jens Rohde², Michael Wagner³ and Ulrich Wetterauer⁴

¹ Vienna University of Economics and Business Administration, Institute of Transport Economics and Logistics, Nordbergstraße 15, A-1090 Vienna, Austria

² Darmstadt University of Technology, Department of Production & Supply Chain Management, Hochschulstraße 1, 64289 Darmstadt, Germany

³ Paul Hartmann AG, Corporate Materials Management, P.O.Box 1420, 89504 Heidenheim, Germany

⁴ j & m Management Consulting AG, Kaiserringforum, Willy-Brandt-Platz 5, 68161 Mannheim, Germany

This chapter will introduce the APS used in the case studies from *i2 Technologies, Peoplesoft and SAP: Six.One, Supply Chain Planning, and Advanced Planner and Optimizer (APO)*. As these tools regularly consist of a multitude of software modules and special add-ons, only a brief survey without claiming completeness can be given. Furthermore, different lines of business can use different modules of an APS. It is also possible to use an APS only partially, e. g. without modules for scheduling or only using modules for demand planning and demand fulfilment. For each individual case the composition of modules has to be evaluated and selected (see Chap. 16).

18.1 i2 Technologies – i2 Six.One

i2 Technologies, based in Dallas, Texas, with European branch offices in London, Munich, Paris, Milan, Helsinki and Barcelona offers a broad range of APS software modules to synchronize and optimize activities involved in successfully managing supply and demand. i2 was established in 1988, when it offered its first software solution called *Factory Planner*, initially successfully used in the Metals industry. Over time i2 developed decision support modules for complex customer problems, which reflected the shift in focus from isolated decision support to integrated SCM solutions, collaborative and continuous planning. i2 offers a comprehensive solution package. The current release i2 Six.One covers the entire value chain for all kinds of business in all major industries including Automotive and Industrial, High-Tech, Retail, Metals and Transportation. i2's core competencies include products for the five optimization areas Production, Revenue and Profit Management, Supplier Relationship Management, Fulfillment and Logistics.

Partnerships with IBM, Accenture, ROCE Partners, Tata Consultancy Services and i2's own (over a 1000 people strong) operations in India service global international corporations as well as regional mid-market customers (see e. g. i2 Technologies (2004)).

18.1.1 i2's Software Modules

Hence, Fig. 18.1 shows only a partial survey of some modules with respect to the case studies of this part of the book (Chaps. 21 and 23). The i2 software modules can be grouped into the five categories of optimization mentioned above. Together with Content and Supply Chain Operating Services (SCOS) they complete the i2 offering. In the following, we will focus on the core SCM modules in more detail, assign them to the different planning processes of the SCM Matrix and give an overview of additional i2 solutions.

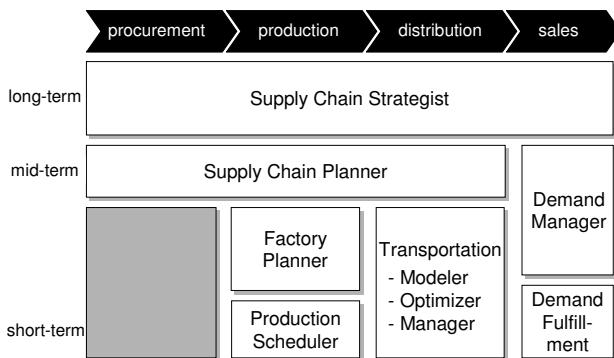


Fig. 18.1. Software modules of i2 Technologies Six.One

Supply Chain Strategist supports strategic what-if-analysis across the entire supply chain. The optimal combination and location of production sites, distribution centres and other entities is determined. Material flows with all related costs as well as constraints can be modelled in different scenarios.

Demand Manager supports the forecasting process through statistical methods, inclusion of causal factors and management of multiple inputs from different organizational units. POS (point of sales) data can be integrated and different views on demand data can be offered. Furthermore, OLAP tools enable efficient access to relevant data.

Forecast data can be aggregated and allocated across sales hierarchies, product hierarchies and time, or can be segmented into further definable hierarchies as for manufacturing locations, packaging or financial measures. So-called "attach rates" support the creation of dependent forecast. Dependent forecast is used in the consumer packaged goods industry, for example, to derive a forecast for individual finished goods (e.g. shower gel and shampoo) from a forecast for a bundle of finished goods (e.g. shower gel plus a small package of shampoo that is bundled for marketing purposes). This functionality is also used in industries like the high-tech

industry where a forecast on component level is derived from the forecast on finished goods level to drive purchasing decisions (see Chap. 21).

Supply Chain Planner enables modelling and optimizing supply chains with respect to material, capacity, transportation and customer-service constraints. It generates optimized, feasible plans across several factories and independent ERP systems. The *Strategy Driven Planning* (SDP) allows planners to define types of problems and strategies to solve them. Furthermore, it is possible to apply appropriate solvers like Linear Programming, heuristics and genetic algorithms. The module provides overall visibility over the supply chain model and the generated plan.

Factory Planner provides detailed visibility over the production plan and helps to reduce manufacturing cycle times. It generates optimized production plans by scheduling backward from the requested date, as well as scheduling forward from current date while considering material and capacity constraints simultaneously. After a first infinite planning step where demand and supply is matched, a finite capacity plan can be determined by i2's proprietary *Constraint Anchored Optimization*. However, the planner can manually interact by analyzing capacity shortages and performing what-if-analysis. In the last planning step a detailed schedule can be generated for the factory.

Production Scheduler supports complex manufacturing environments requiring detailed scheduling and sequencing. It builds detailed sequences and schedules based on genetic algorithms. The decoupling of constraint definition and optimization algorithm allows the handling of a large number of complex constraints. These constraints include shop floor capacities, workload balancing, material availability etc. Additionally, an interactive schedule editor allows manual changes.

Demand Fulfillment provides functionality to assign scarce supply to customers and to quote feasible delivery dates for customer orders. Allocation of supply can be based on current and projected inventory positions or available capacities. It respects customer priorities modeled in a sales hierarchy (e.g. global pool of available supply versus restricted supply for important customers only), fair share and profitability based strategies with the assignment of supply to particular customers. Orders are promised in real-time based on the current status of inventories and capacities in the supply chain network. It typically is integrated to existing order management systems and respects all supply and its allocations.

Enterprise Project Planner plans projects along the entire value chain from design and procurement, manufacturing and transport to installation on site. It links project management and traditional production optimization by shortening project planning cycle time, by improving visibility between manufacturing and installation, and by sending feedback about the current project and supply chain status. The goal is to be able to communicate more reliable project delivery dates to customers

and to optimize the utilization of critical resources for multiple projects. It is typically closely coupled with i2 Factory Planner.

Transportation Modeler, Optimizer and Manager are tools for support of distribution planning processes. Transportation Modeler helps to improve the efficiency of the transportation network. Real world data are used to perform what-if-analysis. Transportation Optimizer automatically builds and routes loads and determines pick-up and delivery times with respect to delivery, equipment and personnel constraints. Furthermore, cross-docking opportunities can be selected dynamically and grouping constraints can be accounted for. The third tool, Transportation Manager, supports activities to execute and manage the transportation process from order management through customer service and financial settlement.

18.1.2 Content

In a collaborative environment it is necessary to speak a common “language” across the Supply Chain to reduce the risk of errors, to speed up communication and to discover potential for improvement by the standardization of criteria for product descriptions. i2 provides content software to administer catalogs for standard data, analyze the company's data management, and provide classification schemes for legacy system data spanning several different systems.

18.1.3 Coordination of Modules and System Integration

Data Integration, application integration and a common user interface embedded in a portal are supported by *i2 Supply Chain Operating Services* (SCOS). SCOS is an open and service based architecture that incorporates several modules for data integration, managing data, for establishing workflows across systems and for measuring and comparing performance indicators throughout the supply chain.

Data integration is based on the *i2 Operational Data Storage* (ODS). The ODS data warehouse incorporates all data relevant for planning in a single data model. Between the i2 planning modules (e.g. Supply Chain Planner, Factory Planner, Demand Planner, etc.), the ODS and towards external systems the functionality can be exposed both as an API and a service. A new forecast, for example, that is created using i2 Six.One Demand Planner, can be automatically transferred via the ODS to the i2 Six.One Supply Chain Planner as input for the master planning process. This is achieved by i2 Common Infrastructure Services (CIS), which supports existing Enterprise Application Integration (EAI) software such as Webmethods for message-based integration and Extract-Transfer-Load (ETL) concepts such as Informatica for mass data transfer.

i2 Business Process Execution (BPE) uses services of multiple i2 modules and integrated external systems to define workflows and logic across these applications. Based on a common data model for the workflow it connects business relevant services and enables a synchronization of the underlying systems. Supported by a design environment, the *i2 Studio*, logic across applications and detailed workflows can be adapted to changes in business processes.

i2 Master Data Management (MDM) is a process-oriented data management tool that enables enterprises to create a common business vocabulary for all their data but also manages business rules, like the relationship of an order to its shipping details. It also manages data synchronization processes between applications - if data changes in one application, it will synchronize the impact of that change with all affected applications. *i2 Performance Manager* (PM) enables business analysts to create reports of key performance indicators of the supply chain, to compare multiple plans and simulations and to alert users to potential problems so they can be resolved proactively.

Finally SCOS provides a common web-based user interface for all i2 applications. It supports role-based user authentication and authorization.

18.1.4 Optimization and Collaboration Modules

Fulfillment Optimization connects planning and execution systems and focuses on customer service (see also Part II). It provides tools to plan inventory and replenishment, manage orders and provide visibility to the execution of processes. Based on SCOS, the backbone for Fulfillment Optimization is *i2 Distributed Order Execution*, which supports processing of transactions across multiple systems and monitoring of supply chain events and process disruptions. Several cross application solutions are offered such as *i2 Customer Program Management* for managing collaborative and VMI workflows, or *i2 Customer Order Fulfillment*, which integrates order configuration, pricing and delivery date quoting of distributed orders. Fulfillment optimization also includes *i2 Replenishment Planner* and *i2 Service Parts Planner*, which e.g. plans multi-level spare part inventories under consideration of demand and supply fluctuations, as well as breakdown probabilities.

Revenue and Profit Optimization is a suite that contains merchandising, planning and pricing tools to manage revenues and achieve the most profitable product mix and pricing strategy. *i2 Merchandise Planner* and *i2 Markdown Optimizer* support strategies to analyze effects of promotions, to manage sequences of price markdowns and to quote profit optimal prices.

Supplier Relationship Management encompasses strategic sourcing, collaborative design and manufacturing, and collaborative e-procurement both for direct and indirect goods. It enables to create and sustain sourcing strategies across areas of design and strategic sourcing responsibilities.

During procurement, Request for Quote processes and negotiations are supported, as well as multi-currency and multi-language transactions. i2 Six.One Supply Chain Collaboration, for example, can be used to integrate suppliers and customers into the planning process through an Internet based platform. It improves the exchange of information through a standardized communication process and the definition of problems to flag alerts.

18.2 Peoplesoft – EnterpriseOneTM Supply Chain Planning

Peoplesoft, founded in 1987 and headquartered in Pleasanton, California, traditionally offers ERP software for non-productive and manufacturing industries. In 2003 Peoplesoft acquired J. D. Edwards, a Denver-based provider of ERP and Advanced Planning software for medium-sized companies. This acquisition turned Peoplesoft into the second largest enterprise application software company in the world. On this occasion Peoplesoft took over “Supply Chain Planning”, the Advanced Planning part of the J. D. Edwards 5 software suite, which is now within the Peoplesoft EnterpriseOneTM product family (see e. g. Peoplesoft (2004)).

18.2.1 Peoplesoft’s Software Modules

In the following, the APS software modules of Supply Chain Planning are briefly introduced. An overview is given in Fig. 18.2.

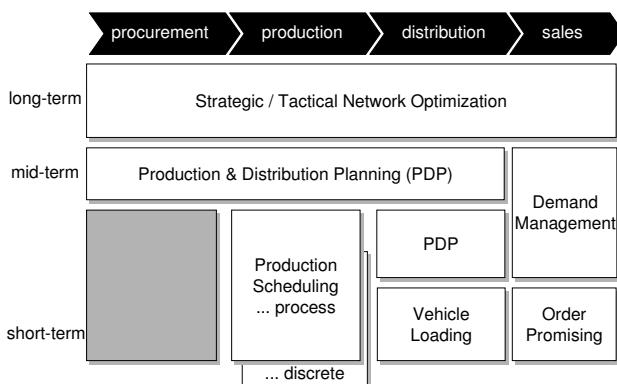


Fig. 18.2. Software modules of Peoplesoft’s Supply Chain Planning

Strategic Network Optimization (SNO) is intended to be applied on the strategic planning level. Optimization methods of Linear and Mixed Integer Programming (CPLEX; see ILOG (2004)) and special purpose heuristics (e.g. for capital asset management and single sourcing) support the choice of appropriate supply chain structures. The most striking feature of SNO is its visualization. Even complex supply chains can graphically be “designed” without any knowledge of mathematical modelling being necessary. The case study of Chap. 20 will show that SNO is not restricted to strategic planning, but can also be used for Master Planning. Tactical Network Optimization is a “light” version of SNO where the strategic features of SNO, like capital asset management, have been disabled.

Production & Distribution Planning (PDP) and Vehicle Loading.

PDP is a versatile tool. Its major focus are the mid-term Master Planning and the shorter-term Distribution Planning. For this, Linear Programming and some heuristics are applied. Here, ATP (see Chaps. 9 and 12) quantities can be computed which serve as an input for order promising (in the form of so-called “supply events”).

Additionally, a more detailed planning of transportation means is aspired for the Transport Planning software module as already described (see Chap. 5). Thus deployment and shortage planning functionalities are also addressed by further components of the PDP module. Since version 4.0 of PDP multi-stage matching of anonymous supply and customer orders (“pegging”) and component/material substitution are supported. The current version 5.0 features a newly designed modelling interface, offers additional manufacturing planning capabilities and introduces a new heuristic called “Supply Planning” in order to execute an in-memory MRP run.

The *Vehicle Loading* module adds capabilities for optimizing the usage of the storage space of vehicles.

Demand Management (DM) consists of the two modules *Demand Forecasting* and *Demand Consensus*, the latter one will be described in the subsequent Collaboration section. *Demand Forecasting* provides statistical forecasting methods, e.g. exponential smoothing and ARIMA, as well as tools to analyse and incorporate causal factors like promotions and marketing campaigns. Furthermore, the identification of unexpected demand changes, trends, seasonality or correlation, and the measurement of forecast accuracy are supported. Additional features are, for example, the introduction of new products, dynamic product pricing and simulation/what-if-analysis.

Production Scheduling Process (PSP). Two alternative software modules for the short-term production planning are offered. The older one, *Production Scheduling Process*, is one of the rare software tools that are especially designed for continuous production processes, common in consumer goods industries or the process industry in general. Its main focus

are parallel continuous production lines with up to two stages of production. As opposite to most other scheduling modules which pursue time-based objectives like minimizing the lateness, PSP aims at minimum production, inventory and changeover costs. Such monetary objectives fit better to supply chains having a deliver-to-order decoupling point where products are made to stock without the knowledge about definite customer orders.

Production Scheduling Discrete, however, is dedicated to multi-stage production processes with floating bottlenecks and complex BOMs which can frequently be found in discrete parts' production. As solution method *texture-based programming* (Beck et al., 1997) is used. Texture-based programming builds on similar principles as constrained programming (see Chap. 29). However, as opposite to constrained programming no problem specific knowledge is needed because the algorithm is self-configuring.

Order Promising. Besides ATP and CTP checks, *Order Promising* provides also “*Profitable To Promise*” (PTP) functionality, i.e. different fulfilment options can be assessed by their profit margin. *Order Promising* has two modes of operation, the Auto Promise mode and the Scenario Management mode. The first one promises due dates automatically using predefined business rules. In the latter mode several fulfilment scenarios can be compared and evaluated interactively.

Execution-oriented components of both the former J.D. Edwards and the former Peoplesoft ERP system like *Warehouse Management*, *Requirements Planning* or *Sales Order Management* intend to supplement the above APS modules.

18.2.2 Coordination of Modules

Peoplesoft provides rather general workflows (see Chap. 5) for different kinds of industries in order to give some advice how to design the information flows between software modules. As an example Peoplesoft's workflow for process industries is shown in Fig. 18.3. When comparing it with Fig. 4.4 (page 97), one can see that this workflow also fits well the planning requirements of consumer goods manufacturing supply chains. Thus, instead of describing the information flows in Fig. 18.3, the reader is referred to Sect. 4.3.1.

18.2.3 System Integration

The *Advanced Planning Agent* (APAg) software component is a graphical tool to constitute and control the data flow between different software modules and to start batch execution of the respective software components. Usually, a common repository of databases is kept by APAG for consistent data storage. Both, vertical (e.g. between mid-term PDP and *Production Scheduling*) and horizontal (e.g. between *Demand Forecasting* and mid-term PDP) information flows are supported.

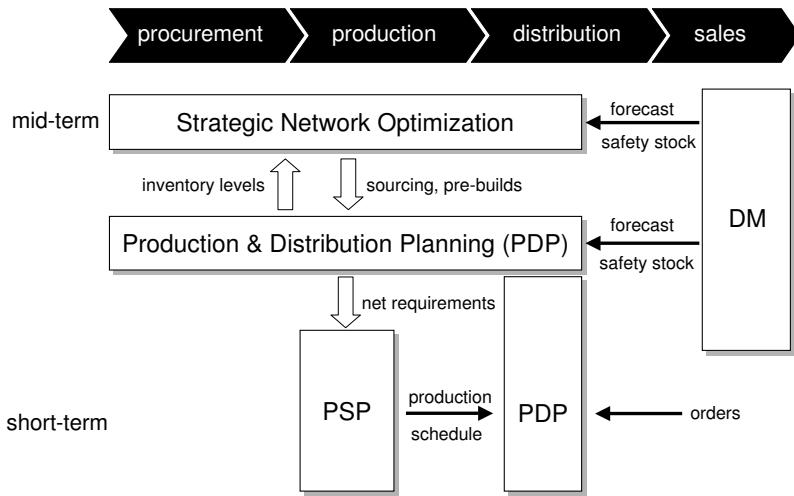


Fig. 18.3. Peoplesoft’s workflow for process industries

Soon APAg will be replaced by the more modern *Supply Chain Business Modeller* (SCBM). SCBM provides data to APS modules at the required level of granularity and enables batch integration to ERP, CRM and other legacy systems. A real-time integration with Peoplesoft’s ERP system is implemented by the *XML-based eXtended Process Integration (XPI)* interoperability engine. This functionality is of particular interest for the *Order Promising* module in order to guarantee an online response to customer requests.

18.2.4 Collaboration Modules

Peoplesoft’s Demand Consensus provides the capability to generate common, consensus-based forecasts via the Intra- / Internet. A reconciliation engine calculates the relative weight of each demand planner on the basis of his historical forecast accuracy. The *Demand Consensus* thin clients can even download worksheets for working off-line. *Demand Consensus* can be used in combination with a statistical forecasting application, such as *Demand Forecasting*. For this *Demand Consensus* and *Demand Forecasting* use the same data base scheme.

PDP provides the infrastructure for collaboration of different users in several locations. This PDP functionality is built upon the Peoplesoft *Distributed Object Messaging Architecture (DOMA)*. It allows several users to work on synchronized data by continuously propagating data changes. The most important features of DOMA are:

- multi-user mode and user specific profiles,
- decision-making distributed across the whole supply chain,

- real-time information sharing,
- database surveillance and event-oriented triggering of data changes and
- alert monitoring and messaging.

Furthermore, PDP features a DOMA-enabled Excel client and a web interface (“Collaborative Web Client”).

18.3 SAP – APO

SAP AG (Walldorf/Germany) has been active in the APS market since 1998. The *Advanced Planner and Optimizer* (APO) was originally intended and sold as an independent software suite. Together with the *Inventory Collaboration Hub* and *Event Management* it now sold as the *mySAP Supply Chain Management* solution. Further mySAP solutions are, for example, *mySAP Business Intelligence* including the *Business Information Warehouse* (SAP’s Data Warehouse), *mySAP Customer Relationship Management* and *mySAP Supplier Relationship Management*. This section will provide an overview of selected APO components. For more information, see the current APO documentations (e.g. SAP (2004) and Knolmayer et al. (2002)).

18.3.1 SAP’s Software Modules

APO is a fully integrated APS. All modules can be accessed through the *Supply Chain Cockpit* and have an identical look-and-feel. The paragraphs below give a brief description of the individual APO modules illustrated in Fig. 18.4 based on APO 4.1.

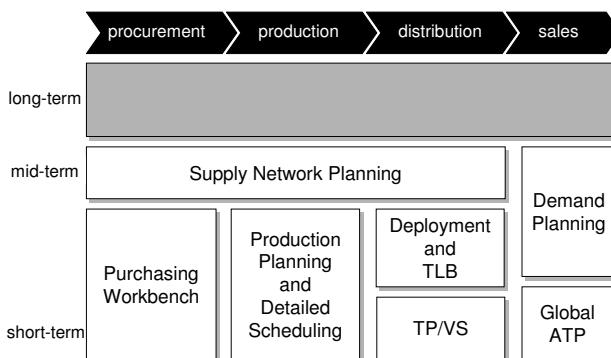


Fig. 18.4. Software modules of SAP APO

Demand Planning offers – in addition to conventional statistical methods – promotion planning tools, life cycle concepts, what-if-analysis, phase-in

planning for new product initiation and collaborative forecasting methods. Reports on forecast accuracy can be generated and alerts can be raised. Furthermore, this module provides OLAP tools for Data Warehouse integration. The current version provides new functionality called *characteristics-based-forecasting* having special aggregation and disaggregation procedures for components of configurable products (e.g. personal computers).

Supply Network Planning provides planning and optimization functionality that take into consideration capacity constraints and costs, for example. Optimization is based on automatically generated *Linear and Mixed Integer Programming* models (see Chap. 27) which use ILOG CPLEX (see ILOG (2004)) for the solution process. Decomposition rules regarding time, resources and priority can be applied to speed up the solution process at the price of a lower solution quality. Additionally, proprietary heuristic approaches are used, such as *Capable-to-Match* (CTM) – a rule based approach on the basis of customer orders. Simulations of different supply chain configurations as well as matching of supply and demand with respect to alternative production sites, substitution of products, prioritizing customers, durability etc. are supported. Alerts can be raised in case of late deliveries and violation of bottleneck capacities or further constraints. Supply Network Planning contains the modules Deployment, Transport Load Builder (see below) and VMI (vendor managed inventory; see Chap. 14). Supply Network Planning provides functionality to manage different planning scenarios, and thus, can be used for evaluation in a strategic planning.

Global ATP performs a rule based multi-level component and capacity check based on current data. It provides product substitution methods, alternative site selection for production and purchasing, and methods for allocating scarce products and components to customers, markets, orders etc.

Production Planning and Detailed Scheduling provides methods for optimizing detailed capacity and material planning simultaneously. It performs multi-level forwards and backwards scheduling. Different constraints can be considered in simulations and interactive scheduling using gantt-charts is provided. Current, short-term data are integrated into optimization runs. Production Planning and Detailed Scheduling uses constraint programming (based on ILOG Scheduler and ILOG Solver) and proprietary genetic algorithms (see Chaps. 28 and 29). These approaches can be combined with decomposition rules regarding time, resources and priority.

Deployment and Transport Load Builder (TLB) allocates inventory by matching actual supply (produced quantities) to planned supply. This allocation is controlled by *push* and *pull* strategies, predefined quotas and priority rules. For example, results from the Supply Network Planning optimization run can be used for defining such quotas. Inventory and al-

location plans are displayed graphically. Transport Load Builder ensures that vehicles are loaded within a specified minimum and maximum range. Iterations are used to derive a feasible deployment plan respecting vehicle loads.

Transportation Planning and Vehicle Scheduling (TP/VS) is SAP APO's short-term planning module for the transportation processes. On the basis of daily shipment plans optimal vehicle loadings and routings can be derived. The optimization is driven by ILOG components. A proprietary genetic algorithm (see Chap. 28) and additional heuristics supplement TP/VS's solution process. TP/VS models allow inclusion of multiple picks and deliveries, multiple depots, time windows as well as net-changes of existing plans (e.g. adding vehicles, modification of time windows etc.).

Purchasing Workbench is used to make (automated) decisions on multiple supply sources and replenishment. A web-based interface provides a platform for exchanging e.g. confirmed delivery dates as well as sharing and adjusting delivery plans.

18.3.2 Coordination and Integration of Software Modules

APO offers a graphical user interface, the *Supply Chain Cockpit*, that gives an overview of the supply chain being modelled and from which all APO software modules can be accessed. The Supply Chain Cockpit also provides the *Supply Chain Engineer* to graphically build a macro-model of the supply chain. This model can be shown in detailed views, and special information can be extracted for each entity. The *Alert Monitor* is also part of this module. APO planning modules use a common database. To enable fast access for all software modules this database is kept memory resident (the so-called *liveCache*).

APO's *Plan Monitor* provides predefined Key Performance Indicators (KPIs) to observe the performance of generated plans. With calculation rules, additional user defined KPIs can be generated from existing predefined KPIs.

mySAP *Supply Chain Event Management* provides automated collection and tracking of information such as order status, shipments and inventory using Internet and mobile technologies. In response to exception-based events activities in planning and execution system can automatically be triggered.

18.3.3 System Integration

APO provides two different options for integrating OLTP systems. The *Core Interface* (CIF) allows direct access to SAP R/3's data objects and vice versa. Integration to non R/3 systems is achieved through so-called *Business Application Programming Interfaces* (BAPIs). By using BAPIs the objects of APO can be accessed by a kind of programming language. Thus, it is possible to map, for example, ASCII-files to APO data objects. SAP also provides

the *Business Information Warehouse* for storing historical data. APO is able to receive these data, which are particularly relevant for Demand Planning, using predefined queries and OLAP tools.

18.3.4 Software Modules for Collaboration

SAP uses Internet and associated technologies, such as XML, to enable the collaboration between business partners. Using conventional Internet browsers APO can be accessed online. The SAP APO *Collaborative Planning* modules enable this collaboration. They support consensus based planning processes for collaboration on shared plans within demand planning, procurement planning etc. (e.g. Inventory Collaboration Hub, web-interfaces of the Purchasing Workbench and Demand Planning). They further provide read-write data access as well as access to planning activities for authorized users using Internet browsers, user specific negotiation processes, user defined screens and workplaces, visualization of alerts, the connection to multiple systems, and links to partner systems.

References

- Beck, J. C.; Davenport, A. J.; Sitarski, E. M.; Fox, M. S. (1997) *Texture-Based Heuristics for Scheduling Revisited*, in: Proceedings of the National Conference on Artificial Intelligence (AAAI-97), AAAI Press, Menlo Park, California
- i2 Technologies (2004) *Homepage*, URL: <http://www.i2.com>, Date: July 19, 2004
- ILOG (2004) *Homepage*, URL: <http://www.ilog.com>, Date: July 19, 2004
- Knolmayer, G.; Mertens, P.; Zeier, A. (2002) *Supply Chain Management based on SAP systems – Order management in manufacturing companies*, Berlin et al.
- Peoplesoft (2004) *Homepage*, URL: <http://www.peoplesoft.com>, Date: July 19, 2004
- SAP (2004) *Homepage*, URL: <http://www.sap.com>, Date: July 19, 2004

19 SCM in a Pharmaceutical Company

Frank Altrichter and Tanguy Caillet

j & m Management Consulting AG, Kaiserringforum, Willy-Brandt-Platz 5,
68161 Mannheim, Germany

Competitive advantage in the pharmaceutical industry is driven by first class research and development and by optimised supply chain operations. Harmonised SCM processes, systems and organisations will lead to reduced inventories, increased capacity utilisation, reduced order lead time, less obsolescences and lower IT system maintenance costs. Critical decisions can be made faster resulting in an improved customer service level. Based on common, standardised data, error rates are reduced and most importantly, full FDA CFR 21 part 11 and GMP compliance can be guaranteed and sustained.

The case study described in this chapter is based on a project in a European pharmaceutical company, that initiated to implement best practice supply chain operations for five European manufacturing plants and the European logistics organisation (active ingredients supply, distribution centres, affiliate customers and third party manufacturers). The project scope includes SCM planning processes, supporting the production planning and detailed scheduling within the pharmaceutical plants as well as the network planning across the company's supply chain to optimally match supply and demand. The planning processes are implemented based on SAP R/3 4.6C and APO 3.1. The case study focuses on the implementation of the APS components SAP APO PP/DS to model the production planning and detailed scheduling in the manufacturing plants, and SAP APO SNP to model the supply network planning of the supply chain. The main results and benefits of the project will be highlighted as well as the major hurdles encountered in the implementation of the SAP APO PP/DS and SNP solution.

19.1 Case Description

19.1.1 Topology of the Pharmaceuticals Supply Chain

The supply chain consists basically of three main levels: chemical plants, pharmaceutical plants and marketing affiliates.

The *chemical plants* deliver the active ingredient (AI). The production of the AI within the chemical plants has not been tackled within the project as the manufacturing process differs significantly from the one of the pharmaceutical plants. The chemical plants, either part of the same company or third party suppliers, are treated as suppliers. Material requirements are planned by the Supply Network Planning module for the entire planning horizon (24 months).

The five *pharmaceutical plants*, spread over Europe, manufacture a wide range of product types. Solids (coated and uncoated tablets, capsules), liquids and creams, biotech medicaments, medical devices, consumer and OTC (“over the counter”) products, sterile products and patches. All manufacturing processes consist of two main steps, formulation and packaging. The output of the formulation step is bulk material (unpacked tablets, liquids, etc.). The bulk material is packed in the packaging step into different put-ups (e.g. blister sizes, country specific packaging). As an order of magnitude, 50 active ingredients are formulated into 500 bulk materials, those are packed into 10 000 finished products.¹

The *marketing affiliates* represent the biggest customer group in terms of volume. Other customer types, e.g. tender business, small countries, wholesalers, government agencies, non-governmental agencies complete the demand picture. All customers forecast their future demand. Depending on the customer type, these forecast figures are converted into sales orders according to Service Level Agreements (SLA) within a certain horizon (on average 9 – 12 weeks). Demand assigned to one plant can also result from a dependent requirement of another plant. For example a bulk material is produced in one plant, but packed by a second plant.

The supply and demand flows are handled by a *sourcing company*. The sourcing company, headquartered in a tax-optimised country, owns the valuable products. The ownership of finished products is transferred immediately upon the quality release to the sourcing company. The active ingredient, as the most valuable part of the product, is always owned by the sourcing company. Thus, the plant acts as a contractor for the sourcing company, transforming the active ingredient into finished products, only invoicing the manufacturing fee.

Distribution centers and warehouses are mainly located close to the plants. In most cases, products are directly shipped from the DCs and warehouses to the customers. In other distribution scenarios finished goods are shipped from one manufacturing plant to another plant due to regulatory reasons and are then delivered to the customer. The distribution itself is not considered as critical, as the value of the finished goods is rather high compared to physical volume and transportation costs. From a master planning perspective the distribution and transportation lead times have to be considered.

Fig. 19.1 gives a high level overview of the pharmaceutical supply chain, including processes and IT systems. Tab. 19.1 summarises the supply chain typology of the pharmaceuticals supply chain.

¹ In the remainder of this chapter the shortform “plant” is used to denote pharmaceutical plants.

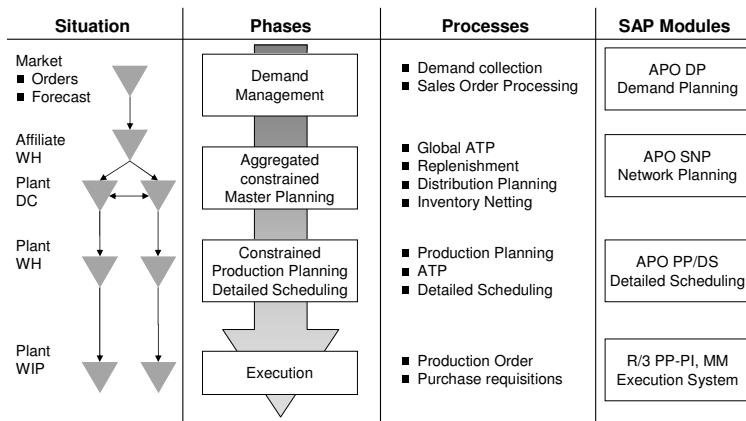


Fig. 19.1. Supply Chain, Systems & Processes

19.1.2 The As-Is Situation

At the start of the project, a very heterogeneous environment, grown over the last decades, was in place. The following list highlights some key aspects of the company's as-is situation:

- Four SAP R/3 systems (two running R/3 PP, two running R/3 PP-PI), one BPCS system,² one R/2 system were used. Data integration between the systems was low, data structures not harmonised. The same product existed with several material numbers in different systems. Information sharing as well as synergies out of a common system were not achievable.
- No central supply chain network planning system was available, resulting in basically no central visibility of the supply chain constraints and problems.
- Production Planning and Detailed Scheduling of the manufacturing processes were performed in various stand-alone systems and spreadsheets, interfaced with local ERP systems. This resulted in massive manual planning effort and sub optimal capacity utilisation.
- There was no central statistical forecasting system in the as-is environment.
- KPI measurements were not consistently defined and did not support common targets.
- The business processes were rather complex, without uniquely defined responsibilities for core planning tasks like Materials Planning, Detailed Scheduling and Master Planning.

² Business Planning and Control System, an ERP-system sold by Systems Software Associates (SSA).

19.2 Objectives of Project

The objectives of the project were the creation of

1. a *to-be supply chain vision* with a clear objective to implement best practice processes enabling the future growth of the company
2. a *to-be system landscape* supporting the to-be supply chain vision and following a global strategy.

Table 19.1. Typology for the pharmaceuticals supply chain

Functional attributes	
Attributes	Contents
number and type of products procured	few (Active Ingredients, AI) specific (Packaging Materials)
sourcing type	single (Active Ingredients) multiple (Packaging Materials)
supplier lead time and reliability	long, based on forecast (AI)
materials' life cycle	short, reliable (Packaging Materials) long
organisation of the production process	two steps (formulation & packaging)
repetition of operations	batch production
changeover characteristics	sequence dep. setup times & costs
bottlenecks in production	known, almost stationary
working time flexibility	frequently used, additional shifts
distribution structure	two stages
pattern of delivery	cyclic with specific country demand dynamic with standard export demand
deployment of transportation means	unlimited compared to cost of products & stock-outs
availability of future demands	forecasted
demand curve	seasonal for medications linked to winter illnesses for example static for others
products' life cycle	several years
number of product types	several (solids, creams, liquids, steriles, patches, biotechs, medical devices)
degree of customisation	standard products (country specific)
bill of materials	divergent in formulation step
portion of service operations	divergent in packaging step tangible goods
Structural attributes	
Attributes	Contents
network structure	divergent
degree of globalisation	Europe
location of decoupling point(s)	assemble-to-order (country specific) deliver-to-order (standard export)
major constraints	capacity of formulation lines manpower in packaging lines
legal position	intra-organisational
balance of power	customers
direction of coordination	mixture
type of information exchanged	forecasts and orders

19.2.1 The To-Be Vision

To achieve the targeted goals of harmonised processes, systems, data and organisational units, the heterogeneous as-is environment was reengineered, and a streamlined and integrated to-be environment had to be designed. In the to-be environment the six ERP systems are integrated into one central SAP R/3 system. Based on the central ERP system for the entire company one central APS will be setup, representing the entire supply chain. The list below summarises the major features of the to-be environment:

- Change the entire organisation from local, function-oriented thinking to a common European company, sharing the same targets and commitment in true collaboration between the business functions and the supporting IT function;
- building of an European team to support that challenging vision on both, IT and business side;
- base the project founding on expected benefits, proven by a business case performed before the implementation started;
- buy-in of all involved stakeholders right from the beginning to propagate the new vision and to support its implementation;
- setup a collaborative forecasting;
- visibility of the demand and the supply through the complete network of the supply chain based on one global, constrained master plan;
- one common detailed scheduling system used by all plants, customised to support local specificities and process inherent constraints;
- installation of a common European reporting and controlling process, supported by common key performance indicators (KPIs);

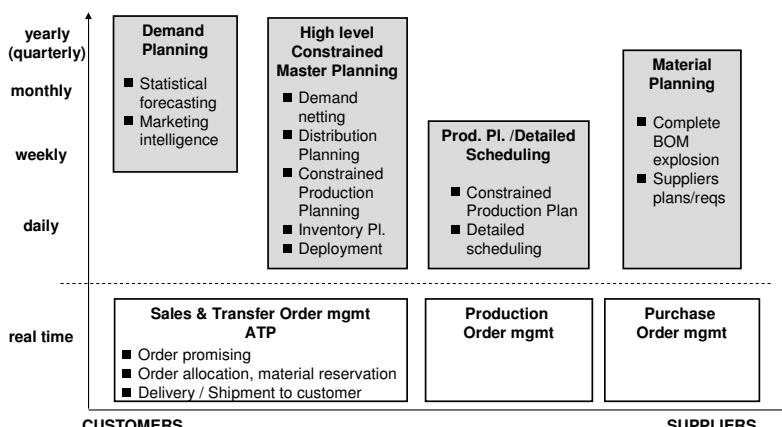


Fig. 19.2. Simplified Process Map

- integration of suppliers into the master planning processes;

- implementation of VMI processes for the major affiliates and customers;
- transportation planning and vehicle scheduling done by the third party logistics providers.

The to-be process map shown in Fig. 19.2 illustrates the vision set at the start of the project.

19.2.2 The To-Be Landscape Implemented

Fig. 19.3 visualises the new IT system landscape supporting the to-be vision of the project. The IT system landscape is based on SAP R/3 4.6C, SAP APO 3.1, the standard core interface (CIF) to integrate R/3 and APO, SAP Business Warehouse (BW) and SAP Enterprise Portals (EP).

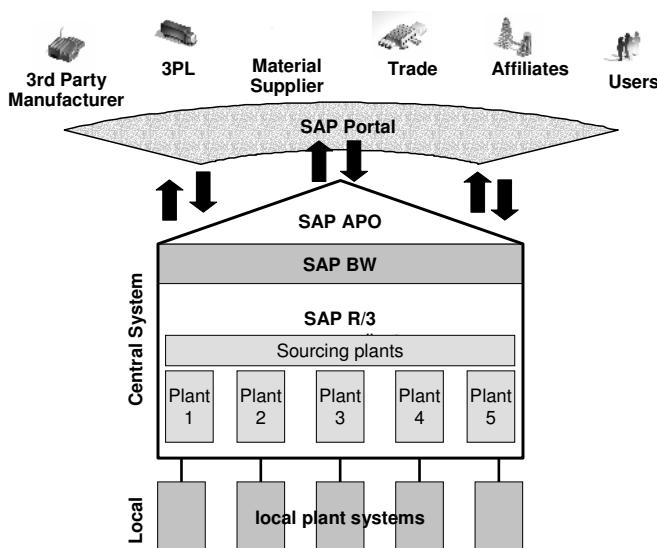


Fig. 19.3. Simplified System Landscape

The central SAP R/3 system covers functionalities provided by the following modules: Production Planning / Process Industry (PP-PI), Materials Management (MM), Sales & Distribution (SD), Quality Management (QM, for batch management, quality inspection lots only), Controlling (CO, for product costing and budget planning), Warehouse Management (WM). After all relevant functionalities were migrated from the old (local) ERP systems of the plants to the new central ERP system, the remaining local non-ERP systems had to be interfaced to the central environment. These are mostly execution control systems, laboratory information management systems (LIMS), material handling systems (MHS) and warehouse management systems (WH).

The central R/3 system provides the integration basis for the APO system. From APO, the modules Demand Planning (DP), Supply Network Planning (SNP) and Production Planning / Detailed Scheduling (PP/DS) are used. The process coverage of the APO modules is shown in Fig. 19.1. SAP BW is the foundation of a common reporting and performance measurement system. SAP EP is used to integrate customers into the demand planning process and to enable customers to access sales orders and delivery confirmations.

19.3 Planning Processes

The planning processes introduced in Fig. 19.2 were mapped to the following ERP and APS modules:

Demand Planning	SAP APO DP
Master Planning	SAP APO SNP
Detailed Scheduling	SAP APO PP/DS
Materials Requirements Planning	SAP R/3 MRP
Production Order Management	SAP R/3 PP-PI
Inventory Management	SAP R/3 IM-WM
Procurement Direct Materials	SAP R/3 MM
Master data management	SAP R/3 MM
Supply Chain Controlling	SAP R/3 CO and SAP BW

The implementation of SAP APO PP/DS and SAP APO SNP, being the main enablers of the supply chain benefits envisioned in the initial phase of the project, are described in the next two sections.

19.3.1 Production Planning and Detailed Scheduling - APO PP/DS

The planning model As mentioned before, the product range of the company is very heterogeneous, including solids, liquids & creams, steriles, patches, biotech and medical devices. This variety requires APO PP/DS to support a broad range of constraints and planning scenarios:

- finite capacity of manufacturing resources,
- availability of labour force,
- prohibition to operate defined resources in parallel,
- priorities for a production or procurement alternative,
- availability of critical components,
- maximum holding time of products before next processing step,
- minimum waiting time between manufacturing steps and

- various lot sizing rules.

The production plan is generated automatically by APO PP/DS. Based on the production plan a sequenced schedule is computed. This requires the optimisation of the following parameters:

- Reduction of machine set-ups by producing products of equal set-up groups in a sequence;
- production sequence ordered by increasing compound concentrations to avoid intermediate cleanings;
- production sequence grouped by product and ordered by increasing order quantity, such that in case of quantity deviations of the packed bulk batch, the yield is on the biggest order (the relative deviation will be the smallest for the biggest order);
- in case of bulk material with an intermediate storage in a holding tank, the holding tank has to be emptied as soon as possible, requiring to group products consuming the same bulk material together.

In Fig. 19.4 an example is given for the creation of a sequenced schedule. The production planning run creates, changes or deletes unfixed planned orders. The planning run is executed bottom-up starting with the finished goods demand. The finished goods demand is covered by corresponding planned orders. The planning run creates dependant demand for components and the half finished goods bulk, which are covered either by planned orders or by purchase requisitions.

The resulting production plan considers all material and capacity constraints and tries to keep the due dates, but the sequence of the orders is not good. The optimisation of the sequence is done interactively by the planner as not all parameters for an optimal sequence can be considered by the system automatically. To obtain an optimal sequence, the planner selects orders for a certain time period on one or multiple resources in the detailed scheduling planning board (DSPB) of PP/DS and calls the PP/DS optimiser. Using the genetic algorithm provided by the system, set-up times and delay costs are optimised as visualised in Fig. 19.4. The resulting sequenced plan respects due dates and avoids unnecessary set-ups.

The PP/DS optimiser does not take all of the mentioned sequence parameters into account. Therefore, the planner has to further improve the plan manually. The first option is to manually select a group of orders (e.g. a sequence of orders of the same set-up group) and improve their sequence using a sorting heuristic of APO (e.g. by order quantity or compound concentration). As a second option the planner may improve the sequence by manually moving orders and operations in the Gantt chart. Sequence optimisation is performed usually weekly for a planning time horizon of 5 – 9 weeks (four weeks is the frozen horizon).

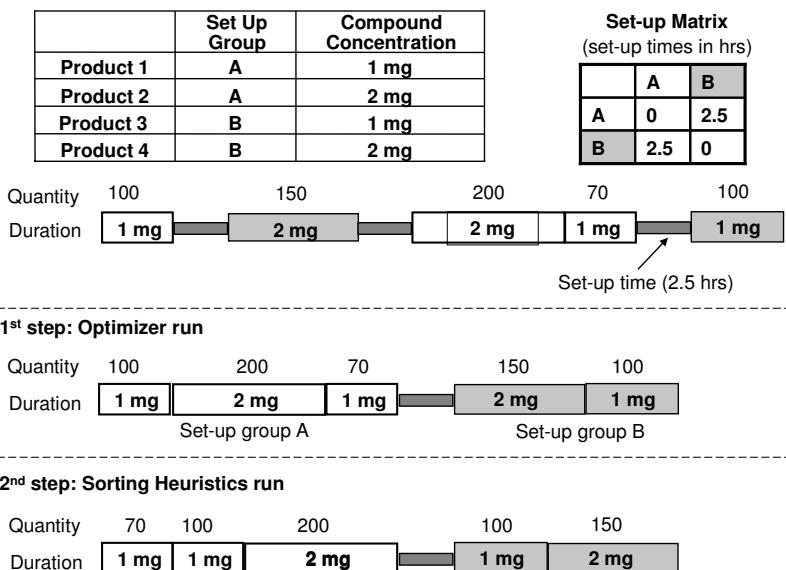


Fig. 19.4. Overview of Sequence Optimisation

Process Integration The production planning and detailed scheduling process transforms demand into feasible production proposals and feeds back information about the actual production execution into the planning processes. Out of the master plan generated by the Supply Network Planning module (SNP) planned orders (fixed and unfixed ones) are created to be further processed by PP/DS. As a first step in the overall planning cycle (see Fig. 19.5), the SNP planned orders are manually converted into PP/DS planned orders by the planner. Compared to the bucketised SNP orders the PP/DS orders are time continuous, i. e. they have a precise start and end time.³ The PP/DS production orders are planned with respect to critical materials only that may constrain the plan. Uncritical materials not constraining the plan are planned by the MRP process in R/3, based on the PP/DS production plan.

Close to the execution date, the sequence of planned orders is fixed. Within a horizon of usually 2 weeks, the PP/DS planner triggers the conversion of planned orders into process orders. While the PP/DS planner may still change some attributes of created and released process orders (e.g. limited changes of quantities and adjustment of the plan to the actual production progress), the planning object “process order” is within the responsibility of SAP R/3 as only there GMP relevant tasks can be performed (e.g. allocation of batches, keeping track of the process order life cycle, QA approval). The PP/DS planner can easily identify the status of the execution progress

³ The SNP orders are planned based on weekly time buckets.

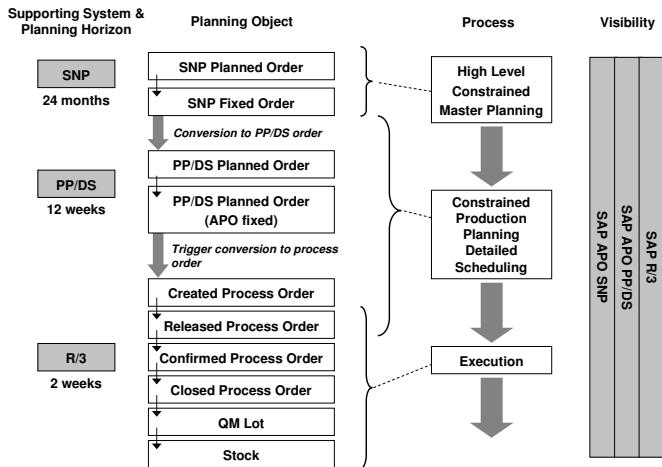


Fig. 19.5. Planning Object Lifecycle

within the PP/DS planning board by a graphical code (shapes and colours), reducing the coordination effort with the shop floor significantly.

Set-up of the master data model As the effort spent on master data for the implementation exceeded 60% of the total project volume, this topic will be discussed in more detail. All relevant master data (see also list below) is mostly already existing in SAP R/3:

- R/3 material master data (= products in APO)
- R/3 resources & capacities (= resources in APO)
- R/3 recipes and bill of materials linked by production versions (= production process models/PPMs in APO)
- Existing in APO only: Set-up matrixes indicating the duration and cost of a transition of one set-up group to another (from one product to another)

For the integration of master data from SAP R/3 to APO PP/DS the standard core interface (CIF) was used.

A commonly underestimated effort though is hidden in the fact that the R/3 master data is usually set-up in a way to support execution processes, but do not take planning aspects into account. Prior to this project master data was existing in various systems like SAP R/3, SAP R/2 and BPCS, and planning was performed in stand alone systems with proprietary master data and complex interfaces. The re-usability of this master data for the APO implementation was low. Effort had to be spent for redefining all MRP parameters to support the technical integration aspects of R/3 with APO and to support the new planning processes. The material master data of R/3 had to be completed with MRP and APO parameters, conflicts with established MRP processes had to be resolved. The R/3 resources define the basis for all

capacity planning in APO. Resources had to be adjusted or newly defined in R/3 in order to create the resource models in APO correctly. For example, it must be defined whether a resource can be used by APO PP/DS and SNP or not, whether multiple operations can be executed in parallel or not, and whether resources are alternates. The standard core interface between R/3 and APO was extended by additional fields and additional rules and default values for data objects used by APO.

Probably the most complex master data element in APO is the production process model (PPM). The PPM combines R/3 information from the production recipes, the bill of materials and the production versions. As the name “PPM” is already indicating, the PPM models the production process. Information like on which resources a production is to be executed, what are the relationships and time dependencies between operations, how much time does an operation need to run, which set-ups are to be performed, etc. are defined in the PPM. The PPMs in APO are generated mainly based on standard R/3 data objects like recipe, bill of materials and production version. Due to the GMP compliance, recipes, bill of materials and production versions in R/3 are critical elements; any change to the objects requires an approval from quality assurance. In the project, basically all recipes had to be restructured and additional information was added to support the planning processes in APO:

- Phases and operations in the R/3 recipes had to be split into individual production steps that require an individual visualisation in APO.
- Dummy phases or dummy resources were added to model specific constraints in APO.
- Recipes were split into APO-relevant and non-relevant phases.
- Sequence dependant set-up informations (the set-up groups) were added to recipes.
- Additional production versions were created to support time dependant changes in the recipe.

In addition to the changes required, the general quality of the existing master data required rework, too. Interdisciplinary skills were needed in the project team to cope with all impacts that an R/3 master data change may have.

A significant set of R/3 master data was made available already at the beginning of the APO modeling phase. This helped to assess the quality of the existing master data for all relevant production processes. However, the improvement of the data quality took longer than expected and consumed more effort than originally planned.

19.3.2 Master Planning - APO SNP

The SAP APO SNP solution was designed to take full advantage of the standard functionalities of the tool, and to prepare the company for the future

steps in the evolution of their supply chain. The list below is a summary of the main requirements that were to be implemented:

- Master Planning performed by APO SNP on a 24-months horizon.
- Global network representing the complete flow of materials from affiliates to distribution centres, the pharmaceutical and the chemical plants.
- Integration of critical suppliers, for direct purchasing as well as third party manufacturers with monitoring of their available capacity.
- Utilisation of the Vendor Managed Inventory scenario with 20 affiliates.
- Weekly release of the forecast from Demand Planning to SNP.
- Demand is constrained by the supply plan based on the global SNP optimiser run.
- Monitoring of the supply network with generation of alerts depending of the planning situation.
- Deployment of the available supply at the plants and at distribution centres to the VMI customers.
- Transport Load Building to propose an optimised loading of the different transportation modes (truck, air cargo, sea container, parcel) for the VMI customers only.

SNP planning process The SNP planning process is based on a weekly planning run of the SNP optimiser for the complete supply chain except the third party manufacturers.⁴ Based on the weekly release of the forecast from the Demand Planning module at the VMI customers and at the shipping distribution centre for non-VMI customers, the optimiser is planning the network while respecting the following constraints:

- Production constraints at the plants,
- due dates of the demand,
- safety stock levels at the VMI customers,
- availability of the critical components like active ingredients from the chemical plants,
- fixed lot size of the formulation batches and
- distribution and transportation lead times.

By including the VMI customers in the optimiser run, master planning became more stable. VMI customers may use their safety stocks to prevent shortages. This increases the flexibility for optimizing the plan. Stability of the master plan also helped the plants to stabilise their production plan. Non-VMI customers usually create more demand fluctuations than VMI customers. Those could be better served due to the additional flexibility gained by the VMI customers' safety stocks and the stable demand signal from the

⁴ Third party manufacturers were excluded from the SNP optimisation run due to technical reasons of the SAP APO 3.1 release.

VMI customers. As a result, the delivery on time was improved and inventories could be significantly reduced.

After a planning run of the SNP optimiser (usually performed over the weekend), the master planners look at the alerts generated by SNP and will solve the issues during the week. For the issues involving production capacity, the master planners will need to decide on the shift pattern to use for the given weekly bucket in order to tune the available capacity. For a constrained capacity planning on a mid to long term horizon, a weekly planning frequency is sufficient.

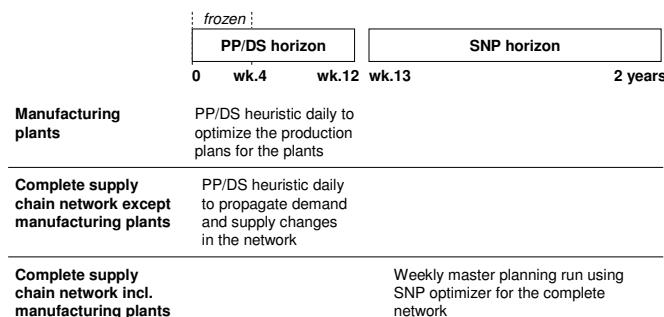


Fig. 19.6. Planning Horizons of PP/DS and SNP

For the near term horizon, the planning system must be able to react quickly to demand and supply changes: supply information is changed daily based on the production plan, sales orders are changed online (whereas the forecast is released only once per week by APO DP). To propagate short term demand and supply changes through the complete supply chain, a further planning run was designed to optimise the complete supply chain network except for the plants. This planning run operates in the PP/DS planning horizon and is based on the PP/DS heuristics (see Fig. 19.6). However, note that this short term planning run is executed by the master planners as it covers the complete network except the plants. The demand (forecast and sales orders) and supply information is propagated daily through the complete network using the planning heuristics of PP/DS. The main result of this planning run are production requirements from the distribution centres to the plants. By that, the production planner gained additional visibility, as the orders received from the distribution centres were not anymore aggregated to weekly buckets, but represented individually in the plan. This visibility was a major business requirement, as many production orders are customer specific orders (e.g. country specific packaging).

The supply generated by production orders within the frozen PP/DS horizon (week 0 to 4) is deployed by the master planner to the distribution centres and to the VMI customers using the deployment heuristic of SNP. The de-

ployment heuristic generates confirmed stock transfers. Based on these stock transfers the master planner creates the VMI sales orders for each VMI customer by running the Transport Load Building module of SNP. This module aggregates for a given transportation lane (a DC to a VMI customer) all the products to be shipped at a certain date, and optimises the load of the trucks or containers based on their weight, volume and forms of transportation. The generated VMI sales orders are transferred to R/3 SD and are sent to the VMI customers as confirmation of the future replenishment deliveries through the SAP Enterprise Portal (see Fig.19.3).

Set-up of the master data model The data model for APO SNP was based on the PP/DS data model. By that, corrected data for materials, resources and PPMs of the plants was provided to the SNP project. For the remaining master data at the distribution centres level, we had to rely on data maintained in the central R/3 system transferred by the standard CIF interface from R/3 to APO. However, two important types of master data could not be transferred from R/3 to APO using CIF, these are the SNP PPMs and the transportation lanes.

The SNP PPMs are normally generated from the PP/DS PPMs by a standard programme within APO. One useful feature of this programme is to take into account the different modes of PP/DS PPMs: Operations in a PP/DS PPM may be run on alternate resources, called *modes*. The PP/DS planning algorithms consider the availability of alternate resources in order to optimise the production plan. Unfortunately, the planning algorithms of SNP cannot consider PPMs containing alternate resources for the same operation. Therefore, based on the given standard programme in APO, an “enhanced” conversion programme was written that

1. splits a PP/DS PPM with alternate resources into multiple SNP PPMs, and
2. creates the SNP PPMs in such a way that the structure of the PP/DS PPMs are kept and the users see the same structures in both modules, PP/DS and SNP.

The second type of master data that could not be transferred from R/3 via standard CIF are the transportation lanes. Transportation lanes are the backbone of the SNP solution, as they represent the connections in the supply network. SAP provides a standard tool for mass creation of transportation lanes. This tool had two weaknesses preventing its use in the project:

- With the tool, transportation lanes for two locations can easily be created for *all* products. In this project we needed *product-specific* transportation lanes, that cannot be created easily using that tool.
- Transportation times are computed by the tool based on the distance between the start and end locations. Here, a more complex transportation

time computation was needed, depending on the forms of transportation, internal lead time of the third party logistics provider, etc.

In order to overcome the weaknesses of the SAP standard tool, a project specific toolset for the creation of transportation lanes was designed and implemented. A *flow database* was setup containing all flows in the network from finished goods at VMI locations to the active ingredient stock at the chemical plants. This database proved to be a very useful tool to gather knowledge about product flows forming the network that is usually scattered across many people in the organisation. The flow database was then interfaced with R/3 to store for each flow the product code and the sending and receiving location in R/3. Based on these fields, R/3 complements this data with the means of transportation and the transportation lead time generated by the standard route determination logic of the R/3 SD module. Finally, for one sending location and one receiving location, the product codes and their respective means of transportation and transportation lead times are aggregated to form a transportation lane that is uploaded to APO SNP through the standard BAPI.

19.4 Results and Lessons Learned

19.4.1 Achieved Results

The main benefits envisioned in the business case prior to the APS implementation were achieved. These are in detail:

- The visibility and problem solving capabilities of the entire organisation improved by the use of a common data basis and a common visualisation tool, allowing better and faster decisions.
- System based finite capacity scheduling and fast simulation capabilities improved the plan stability and resource utilisation significantly.
- Collaborative demand planning with the customers allow for a proactive stabilisation of the demand as changes in the demand by the customer are compared with a constrained demand from the previous master planning run and exceptions are generated.
- The master planning run enables the company to better foresee the future capacity issues and plan accordingly future investments.
- By reducing the order to cash cycle, as well as pushing for more collaboration with the affiliates through a VMI process, the inventory levels were reduced.
- The collaborative concept and the reduction of supply chain steps have increased the overall reliability of the processes. The reduced time to market has lessened the risk of obsolescences which is an important driver to improve the cash flow of a pharmaceutical company.
- By consolidating the system landscape, the IT maintenance costs were reduced significantly.

- Standardisation of the master data enables the visibility and interchangeability of information faster across the supply chain.
- The overall administrative workload for tasks performed previously manually or based on wrong information was reduced significantly.
- Seven local organisations grew together into one European organisation.

19.4.2 Lessons Learned

Without organisational changes and business process reengineering, the risk of implementing the as-is situation within a new tool is real. The to-be vision needs to be propagated by the upper management towards the different organisations of the company to reduce their resistance to change. This has proven mainly true wherever boundaries between the organisations had to be broken down. The KPI measurements should follow the new processes and be revised accordingly. Keeping previous KPIs will not facilitate change as the old way of working is imposed. The planning processes have to be seen and defined in an integrated way. Demand planning – master planning – production planning – materials requirements planning have to be integrated technically and from an organisational point of view.

VMI processes based on Collaborative Planning, Forecasting, and Replenishment (CPFR) could only partly be realised. The marketing affiliates, although part of the same company, were granted control of the replenishment demand towards the plants. They were accepting the result of the replenishment planning run and/or changing the result according to their needs, allowing manual override of the VMI concept. Therefore, contingencies have to be built up in form of inventory to cover uncertain changes in the affiliate's requirements, annihilating the benefits of the VMI process. The success in the implementation of a VMI scenario is not determined by the technical integration of different systems, but strikingly driven by the relationship of the partners and their ability to rely on one another.

A global master planning run as foreseen and supplied by SAP APO SNP will be implemented gradually only. The change management efforts from the function oriented supply chain organisation to a single and integrated European supply chain is estimated as a project risk due to the expected resistance within the organisations. Reactivity of the supply chain will only increase once the master planning can be executed in one integrated step. Resistance to a central master planning in planning and production departments of the plants has to be anticipated and avoided by an involvement of all users, right from the beginning.

Availability and consistency of R/3 Master Data is crucial for succeeding an APO implementation, more than 60 % of the APO PP/DS and SNP implementation time had to be spent actually on R/3 master data definition and revision. Double maintenance of data in different systems (R/3 and APO), especially in a validated environment is not acceptable. Enhancements had to be foreseen to minimise the maintenance effort and the risk of discrepancies.

20 Food and Beverages

Michael Wagner¹ and Herbert Meyr²

¹ Paul Hartmann AG, Corporate Materials Management, P.O.Box 1420,
89504 Heidenheim, Germany

² Vienna University of Economics and Business Administration, Institute of
Transport Economics and Logistics, Nordbergstraße 15, A-1090 Vienna, Austria

Everyone knows the situation: You go shopping in your favourite supermarket and all the items on your list are available, except for one. Therefore, you have to drive to the next store and hope that you can get the product there.

Situations like that decrease customer satisfaction quite noticeably and affect both retailer and producer. That is why customer service has become the main objective of consumer goods supply chains. However, too often higher customer service is accompanied by higher investments in inventory. State-of-the-art Advanced Planning Systems are basic tools for achieving both conflicting management goals, high customer service *and* low stock levels.

This chapter provides some insights into the implementation of the *Peoplesoft EnterpriseOne™ Supply Chain Planning* software (see Chap. 18 and Peoplesoft (2004)) for the food and beverages division of a large European consumer goods company. The aim of the project was to build the infrastructure for more flexible, accurate and faster planning processes.

After introducing the most important attributes of the supply chain in Sect. 20.1.1, the planning architecture for the company's German food branch is described (Sect. 20.1.2). Furthermore, model-building with the Peoplesoft module *Strategic Network Optimization* is specified (Sects. 20.3 and 20.4). This model served as a template for the whole division and therefore had to integrate requirements of other product lines of the company, too.

20.1 Case Description

After introducing the supply chain itself a description of the architecture of the planning system will follow.

20.1.1 The Supply Chain

The supply chain under consideration consists of production plants, a distribution network and some key suppliers. Each of the three production sites produces up to 70 different products, but only a few products can be made in all three plants.

Procurement The material procured for food production can be divided into two classes. The first type of materials are the raw materials and ingredients which the food is prepared from. They are usually bought on the free market and therefore the ordering decision noticeably depends on the current market price. Packaging and labelling material on the other hand is usually single or double sourced. For those items annual basic agreements are made with the suppliers. Therefore, the ordering decision neither causes additional costs nor depends on the order size or the price.

Production The production process consists of three consecutive stages: pre-blending, blending and packaging. The first one is physically separated from the others and only available at plant 3 (see Fig. 20.1). From plant 3 all plants are supplied with pre-blend just-in-time. The pre-blend process comprises a complex production network consisting of batch processes. As more than one product is based on the same pre-blend, the number of pre-blend types is less than the number of end products.

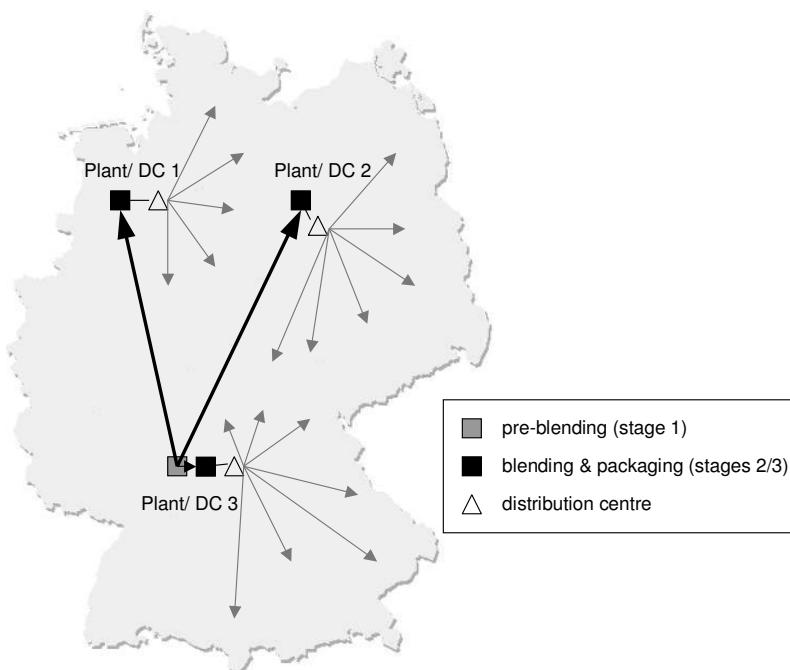


Fig. 20.1. Supply chain structure

If the production equipment has to be switched from one product to another, significant sequence-dependent changeover costs and times occur on

the first and second stage. In each production plant more than five blenders/packers (dedicated to a specific range of products) operate in parallel. A fixed 1:1 relationship usually holds between blenders and packers, but it is also possible that one blender feeds two packers (simultaneously or alternatively). As the pre-blending stage supplies the blenders just-in-time, there are only small buffers for storage of one day's demand at the most. Between stages two and three there are no buffers, as there are fixed pipes which connect both stages (Fig. 20.2). While most products are produced once per week, a few are set up every second week (cyclic schedule). Therefore, the reorder lead-time for the distribution centres equals one to two weeks, depending on the product.

The regular working time in the plants is based on a three shift pattern (24 hours) on five days per week. This regular time can be extended by one to three shifts overtime on Saturdays, resulting in additional costs for higher wages.

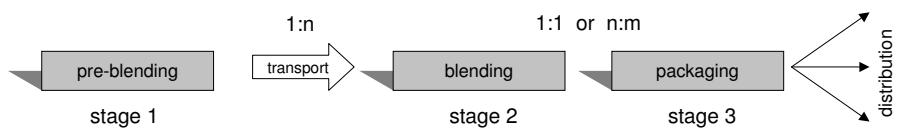


Fig. 20.2. The production processes

Distribution After packaging, the products are transported to the distribution centre (DC) next to each production site. Those items which are not available at the local factory are supplied by one of the other two plants. Each customer is served from the next DC either directly or via cross docking in one to two days (three-stage distribution system). Therefore, each DC has the whole assortment of products available. Only if stock-outs are impending, emergency transports between the DCs are initiated to balance the stocks.

Sales Most products are made to stock and only a few items are made according to customer orders. These make-to-order products are supplied exclusively for one customer under his own brand name. The customer service level is determined by matching orders and available stock at the DC daily. The “customer” ordering the product is usually not the consumer, but mostly a retailer (chain) which operates the retail outlets. Even though the sales for some products are quite constant over time, most of the products are influenced by seasonality and promotional activities. Storage of finished products is limited in time due to shelf-life restrictions.

This type of supply chain has already been introduced in Chap. 3. The food and beverages case considered here is additionally summarized in Table 20.1.

Table 20.1. Typology for the food and beverages supply chain

Functional attributes	
Attributes	Contents
number and type of products procured	few, standard (raw materials) and specific (packaging materials)
sourcing type	multiple (raw materials) single/double (packaging materials)
supplier lead time and reliability materials' life cycle	short, reliable long
organization of the production process	flow line
repetition of operations	batch production
changeover characteristics	high, seq. dep. setup times & costs
bottlenecks in production	known, almost stationary
working time flexibility	low, partially additional costs
distribution structure	three stages
pattern of delivery	dynamic
deployment of transportation means	unlimited, routes (3^{rd} stage)
availability of future demands	forecasted
demand curve	seasonal
products' life cycle	several years
number of product types	few
degree of customization	standard products
bill-of-materials (BOM)	convergent (blending)/ divergent (packaging)
portion of service operations	tangible goods
Structural attributes	
Attributes	Contents
network structure	mixture
degree of globalization	Europe
location of decoupling point(s)	deliver-to-order
major constraints	capacity of flow lines
legal position	intra-organizational
balance of power	customers (retailers)
direction of coordination	mixture
type of information exchanged	forecasts and orders

20.1.2 The Architecture of the Planning System

The architecture template had to take into account the specific requirements of different production processes in the whole food and beverages division. Furthermore, the existing planning systems and the IT-landscape needed to

be integrated into the new architecture. Also, obsolete spreadsheet solutions which had been developed by the planners only for their specific purposes should be replaced by the Advanced Planning System of Peoplesoft.

For the project it was decided to focus on the following planning processes (for a description of general planning tasks in the consumer goods industry see also Chap. 4):

- long-term production and distribution planning,
- mid-term master planning (production and distribution) and
- short-term production scheduling.

For demand planning and short-term distribution planning third-party systems already existed before implementing *Peoplesoft Supply Chain Planning*. As these modules have tight connections to the new Peoplesoft Supply Chain Planning modules, they needed to be integrated accordingly (see Fig. 20.3). Procurement processes were only integrated if they had been identified as potential bottlenecks.

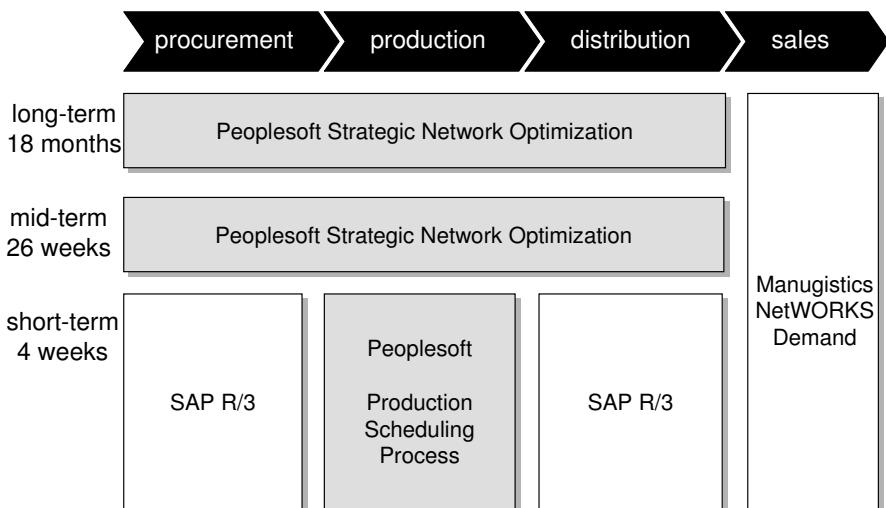


Fig. 20.3. The architecture of the planning system

Long-term Production and Distribution Planning Long-term planning tasks in the food and beverages business comprise a time horizon between one and five years. In this range strategic decisions on the product programme to be offered, the opening or closing of production lines or plants, and the distribution network are made. But in case of the company described here, the long-term production and distribution planning is restricted to 18 months, as in this time horizon production lines can be built up or moved

from one site to another. Furthermore, changes in the shift-pattern have to be coordinated with the works committee and therefore need to be initiated months in advance.

This planning task is based on monthly time buckets and aggregated products and resources. The *Strategic Network Optimization* module is utilized to simulate different scenarios. The software calculates a capacity-constrained, optimized flow of goods and the respective costs for the 18 months horizon. Therefore, the planning process consists of the following steps, iteratively executed:

1. evaluate the status quo,
2. change the network structure manually, e. g. close a line, allow three instead of two shifts etc. and
3. evaluate the new model.

Mid-term Master Planning (Production and Distribution) The Master Planning module integrates all decisions on materials and capacities concerning the whole network of plants and DCs. Therefore, the connecting material flows between the different locations need to be planned on this level. The more detailed Production Scheduling module considers only local resources and assumes the inflow to and the outflow from the site as being given.

The Master Planning decisions taken in this case study are:

- weekly transportation quantities from production sites to DCs and between the DCs,
- weekly material requirements (i. e. packaging material) which have to be ordered from suppliers,
- necessary overtime on the production lines,
- assignment of products to production lines per time bucket and
- weekly inventory levels in DCs.

As the number of product types is relatively small in this application, the products and resources do not have to be aggregated. The time horizon of half a year is divided in 26 weekly time buckets. *Strategic Network Optimization* was selected as the premier solution for this task because it provides easy to use graphical modelling capabilities and a powerful optimization engine (CPLEX; see ILOG (2004)). The model has tight data integration to the short-term Production Scheduling module via the *PeopleSoft Advanced Planning Agent* (APAg) and an Oracle database (see Oracle (2004)) which holds all relevant planning data.

The objective of the Master Planning model is to minimize all costs which are influenced by the decisions described above. Therefore, transportation costs, production costs, costs for overtime and storage costs have been considered.

Short-term Production Scheduling The production scheduling task is implemented using the Peoplesoft *Production Scheduling Process* module. It covers both the lot-sizing and the scheduling task and therefore integrates the modules Production Planning and Scheduling (see Chap. 5). Production Scheduling only has to plan stages two and three of production as the first stage is decoupled by transport and therefore can be planned independently by an additional scheduling model.

The objective is to create a cost-optimized schedule for the production facilities of a single factory. Inventory holding costs, setup costs and penalty costs for not meeting the desired minimum inventory levels form the objective function.

The model covers a planning horizon of four weeks rolled forward once a week. However, the plan may even be revised daily on an event-driven basis. These events are caused by machine breakdowns or impending stock-outs. As some materials and pre-blend have to be ordered two days in advance, the frozen horizon only covers the next two days. The “demand” (requirements) which drives the scheduling model is calculated considering

- updated daily forecasts,
- actual and planned shipments to all DCs,
- safety stocks which have to be held at the DCs,
- actual inventory levels at the DCs, in transit to the DCs and at the plant and
- a “sourcing-matrix” which states the quota to be sourced from a specific plant (calculated from the results of Master Planning; see Sect. 20.4).

20.2 Aim of the Project

The aim of the project was to automate the planning processes by providing a decision support system

- which is able to make planning proposals on its own
- and can be used to interactively simulate several planning alternatives
- thus enabling the planner to select the best one with respect to supply chain costs and constraints.

Expected results were an increased supply chain visibility, reduced inventories and supply chain costs and shorter planning cycles.

20.3 Model Building in Peoplesoft Strategic Network Optimization

Models built in Strategic Network Optimization do not use any mathematical notation. A production and distribution system is modelled graphically

(e.g. by means of drop down menus) and interactively within the system. A Strategic Network Optimization model consists of the following basic elements (see e.g. Günther et al. (1998), Kolisch (1998) and J.D. Edwards (2001)):

- time periods,
- commodities,
- nodes and
- arcs.

Their most important properties will now be introduced.

Time Periods

An optimization model considers a certain planning horizon that may be subdivided into several time buckets. Since the model structure has to be the same in all periods, it has to be graphically defined only once. This structure is then copied for each period to be considered and the period-specific data (e.g. demand varying over time) have to be filled in each copy. In an optimization run all periods are considered simultaneously. Thereby, each period is linked with the preceding one by the stock that is held at the beginning of the period (of course being equal to the stock at the end of the previous period).

Commodities

Two different kinds of “commodities” may occur. First, commodities denote distinct types of (*physical*) goods like raw materials, work-in-process or final products – no matter in which stage of production they are. In this case study the various kinds of pre-blend and packaged goods are examples of such goods. Secondly, commodities represent the *time* spent in production, transport or storage processes. For example, the commodities *regular blending time* or *blending overtime* may be used to distinguish between the cheaper and the more expensive variant of the blending process.

Nodes

“Nodes” represent the processes themselves, e. g. all activities supplying, storing, consuming, transforming or simply controlling any type of commodity. Therefore, nodes model the critical components and constraints of a production and distribution system, but not the material flow within the system. Different kinds of nodes have been launched to represent different types of activities. Generally, nodes can have several input and several output commodities. Nodes without either input or output commodities are available as well. However, nodes of the same kind share common input and output characteristics.

A few kinds of nodes – later on used in this case study – shall illustrate the function and dominant role of nodes. For sake of clarity only the main attributes of each kind of nodes are presented:

Supply node: A *supply node* supplies a single commodity (usually of the physical goods type) and therefore does not have any input commodities. Upper and lower bounds of the amount to be supplied can be specified as given data. Also unit costs for supplying the commodity can be particularized in order to consider total costs of supply. The result of an optimization run is the amount of the commodity actually (and optimally) to be supplied.

Machine node: A *machine node* has a quite similar function, but is intended to supply the commodity *machine (or personnel) time*. Therefore, the capacity and unit costs of a machine have to be specified. The optimal run length of a machine (of both regular time and overtime, depending on the type and costs of the commodity supplied) results.

Process node: A *process node* transfers input commodities (goods and/or time) into output commodities and therefore may have several input and output commodities. This transformation is done with respect to fixed rates of input and output. For example, two units of intermediate product, one tub and two seconds machine time are combined in a packaging process to obtain one unit of a final product.

Batch node: *Batch nodes* restrict the flow of a commodity. A batch node can e. g. be used to specify a *minimum run length* of a process or a *minimum lot-size* – depending on the type of commodity considered. Note, by this a binary decision is implied: either nothing or more than the minimum lot-size have to be produced. As Chaps. 8 and 27 show, a model containing batch nodes is quite hard to solve because the underlying optimization problem has changed from a simple LP to a more complex combinatorial MIP. The user should therefore utilize a batch node only if it is absolutely indispensable to represent reality correctly. Other types of integer/binary decisions like *batch sizes* (predefined amounts of commodities with only integer multiples being allowed) or *setup times* can be modelled by a batch node, too.

StorageCoverLocal (SCL) node: The *StorageCoverLocal node* calculates the desired stock level of a commodity (physical good) at the end of the time period considered. This calculation is due to the inventory balance equation:

$$\begin{aligned} \text{stock at the end of the current period} &= \\ &= \text{stock at the end of the previous period} + \text{inflow} - \text{demand}. \end{aligned}$$

Thereby, the *demand* (e. g. market demand for a final product) is not computed within the model but has to be pre-specified. *Inflow* and *stock levels* (except for beginning inventory) are results of an optimization run.

The inflow is an input commodity of the SCL node whose amount is usually further restricted by another node, e. g. a process node representing some production process.

The resulting amount of stock is influenced by three types of (increasing) “target” stock levels: the minimum, safety, and maximum stock level. While minimum and maximum stock levels usually are hard constraints that must not be violated, the safety stock level is a soft constraint which will be punished by penalty costs if fallen short of. These stock levels have to be specified by the number of periods of future demand, they are expected to cover. Each stock keeping unit – independent of the stock level – is priced with actual per unit holding costs.

Monitor node: A *monitor node* takes different types of commodities as input and limits the total number of (stock keeping) units consigned. This is useful for modelling the storage capacity of a warehouse where stocks of distinct products share a common inventory space. For example, the total stock held in several SCL nodes can jointly be controlled by a single monitor node.

Arcs

Arcs connect nodes, thereby carrying exactly one commodity, each. Therefore, arcs represent the material or time flows within a production and distribution network. The amount of the commodity to be carried is a decision variable. It can be restricted by predefined upper and lower bounds (e. g. minimum or maximum transport capacities if the commodity denotes deliverable goods). Again, the unit cost of the commodity can be specified.

For aesthetic purposes and sake of clarity a further kind of nodes, a *working node*, has been introduced in Strategic Network Optimization. It bundles arcs connecting two bipartite sets of nodes, but carrying the same commodity. So an $n : m$ relation of nodes is replaced by an $n : 1 : m$ relation, thus reducing the number of arcs significantly.

After analyzing the decision situation (see Sect. 8.1), the elements of the Strategic Network Optimization model have to be defined and the necessary data (e. g. resource capacities, production rates, costs) have to be filled in. This may either be done graphically and interactively or via import files. A trial solution run has to be started, checking whether predefined solver parameters are set adequately or refinements have to be made.

These may especially be necessary if a MIP model has been defined, for example by use of some batch nodes. In the most lucky case such a refinement just requires some changes in the parameters of the solver heuristics that Strategic Network Optimization provides for MIP problems. In the worst case a redefinition of the optimization model is necessary, possibly inducing serious modifications in the design of the planning module (see Chap. 4).

20.4 Implementing the Master Planning model

20.4.1 Model Structure

In the following section the elements of the mid-term Strategic Network Optimization model and the solution approaches actually used in this project are described. The model covers the food supply chain from some important external suppliers to the final distribution centres. Master Planning is usually done every Thursday starting with the following week over a rolling horizon of 26 weeks. Four different groups of commodities are considered: raw/ packaging material, intermediates, finished products and time. All nodes and arcs utilized in the model are described in the following subsections. A graphical overview of the node structure is given in Fig. 20.4.

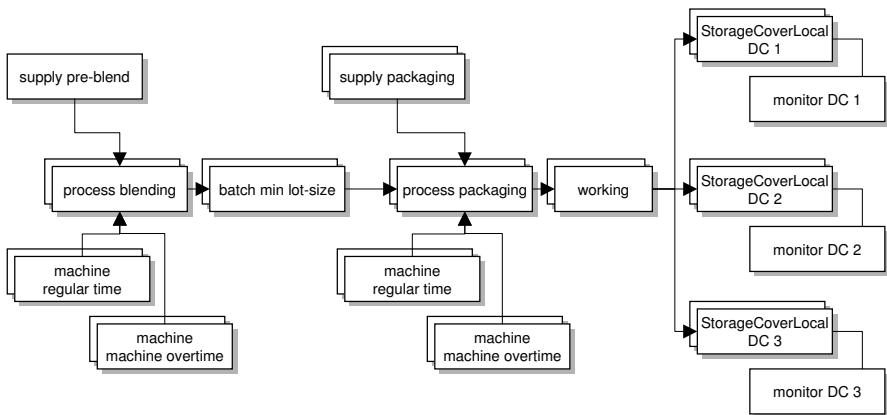


Fig. 20.4. Overview of the Strategic Network Optimization Model

Procurement

Procurement processes are implemented in Master Planning only if the supply of material is restricted in some manner. This applies to the internal supply of pre-blend from plant 3 and to some critical suppliers of packaging material whose weekly production (or supply) capacities are known. As already mentioned, the production processes of pre-blending are not explicitly considered in the Master Planning model, but their capacities are restricted and therefore have to be modelled as material bottlenecks.

In Strategic Network Optimization the internal and external suppliers are represented as supply nodes. For each material and supplier a respective supply node is needed (only a small section of the procurement part is shown in Fig. 20.5).

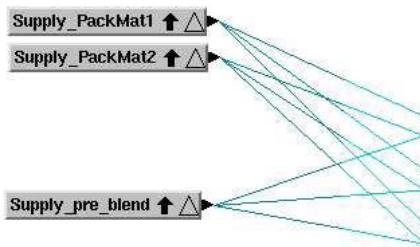


Fig. 20.5. Example of procurement processes in the Strategic Network Optimization Model

The fields of the node are filled with the data described in Table 20.2. The costs in the supply node are required only if a specific material is double-sourced or the prices are changing over time. Costs have to be filled in for all materials since the objective function calculates the whole procurement costs in this case.

The supply nodes are connected with process nodes representing the blending stage. If transportation costs for the transports to the plants occur (esp. for pre-blend from plant 3 to plants 1 and 2), these are modelled as a linear cost rate penalizing the flow on the arc.

Production

Since both production stages, blending *and* packaging, are potential bottlenecks, they cannot be considered as a single planning unit. Furthermore, the production model has to guarantee minimum lot-sizes on the production lines. For each production line and product a single process node is needed (see also Fig. 20.6). The blending process node combines the commodities pre-blend and (regular and over-) time for producing the intermediates. Therefore, the input rates are modelled as follows:

- pre-blend: quantity (in tons) needed for production of one ton intermediate product,
- time: capacity (in hours) needed for production of one ton intermediate product.

Arcs connect the blending process nodes with the batch nodes which ensure that either nothing or more than the minimum lot-size is produced per week.

Table 20.2. Supply node data fields

Data field	Case specific input
Max	Maximum material supply (of pre-blend or packaging materials)
Costs	linear costs per unit procured

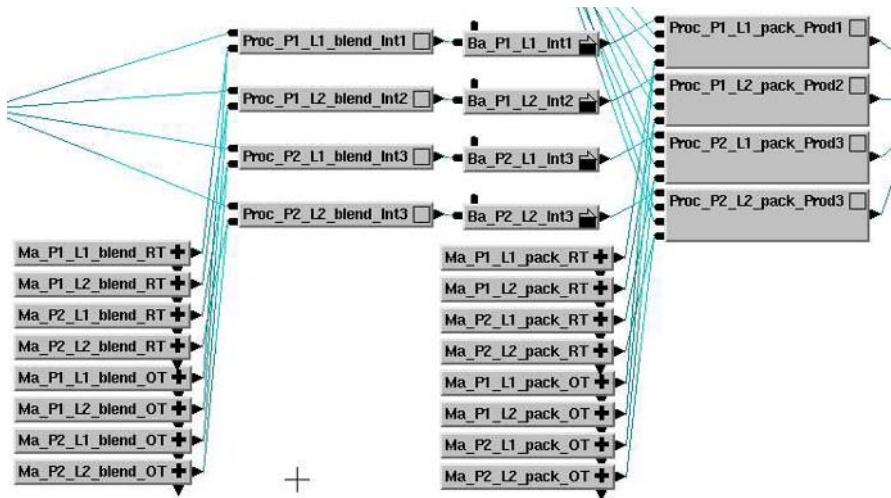


Fig. 20.6. Example of production processes in the Strategic Network Optimization Model

This is the only kind of node in the model which induces integer variables (more precise: semi-continuous variables). This additional complexity (for further information see Chaps. 8.2.2 and 27) has to be balanced carefully against the higher accuracy achievable by considering the minimum lot-sizes. The field “min-run-length” is filled with the minimum lot-size (in tons) which is given by technical restrictions of the machines. The arc (intermediate product) leaving the batch node is directly connected to the packaging process node. Variable production costs (excluding costs for material supply and personnel taken into account otherwise) are modelled as a cost rate on the arc. Input rates of these process nodes are set according to:

- intermediate product: quantity (in tons) required for one ton of finished products,
- packaging material: quantity (in units) needed for one ton of finished products,
- time: capacity (in hours) needed for the production of one ton of finished products.

Both types of process nodes (blending and packaging) consume capacity (time) which is supplied by *machine nodes*. For each process node two machine nodes are required: one provides regular capacity (RT) and the other overtime capacity (OT). Both machine nodes are connected with the process node. Table 20.3 shows the field entries which have to be made in machine nodes and arcs. The maximum capacity is calculated by reducing the total capacity (e.g. five days \times 24 hours per day = 120 hours) per week by an efficiency factor. This value is retrieved from historical data by considering

Table 20.3. Capacity model data fields

	Data field	Case specific input
Machine RT	Max	regular maximum capacity (in hours) available for production on a respective blending or packaging machine
Machine OT	Max	maximum overtime capacity (in hours) available for production on a respective blending or packaging machine
Arc Costs	Costs	linear personnel costs per hour

the following components: start-up/ shut-down time, changeover time, maintenance and repair time.

Distribution and Sales

Each production site is able to serve all three DCs. Therefore the necessary transports from packaging to warehousing are modelled by arcs connecting the packaging process node with the SCL node. As in some cases an n:m relationship between production lines and DCs exists, a working node merges the flows from different production lines producing the same product (see e.g. lines 1 and 2 of plant 2 in Fig. 20.7). This also enables quick access to the overall production quantity of a specific plant/ product combination. The transportation arcs carry linear transportation costs which are calculated from price lists of the third-party carriers.

The SCL node stays abreast of inventory tracking and demand fulfilment. Limits on inventory levels are modelled as *stock covers* (number of periods of future demand covered by stock-on-hand), as it is common practice in the

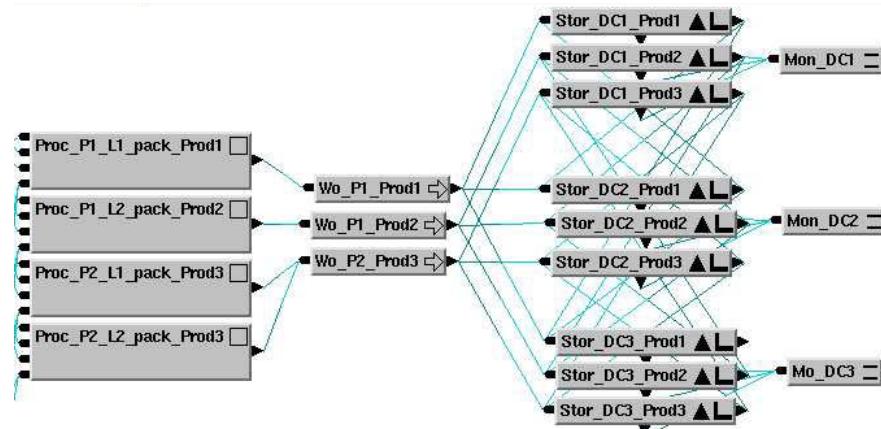


Fig. 20.7. Example of distribution and sales processes in the Strategic Network Optimization Model

consumer goods industry. Here, all stock limits (min, safety and max) were modelled as soft constraints (with penalties) in order to ensure feasibility of the optimization model which could be endangered if the forecasted demand in the first week is much higher than stock-on-hand plus available production capacity. The minimum stock (min cover), the model has to guarantee, is calculated by summing the following components:

- Lot-sizing stock (cycle stock): As most products are produced each week, the lot-size equals approximately the demand of one week. Therefore, the mean lot-sizing stock is half of the weekly demand.
- Transit stock: The delivery lead-time from the plant to the DC is about one day. As the stock for this transport is not considered on the arc, the minimum stock in the DC has to be increased by this amount.
- Quarantine stock: All products have to be kept at the plant for 24 hours. This quarantine time is added to the minimum stock cover at the DC.

Since violation of minimum stocks is punished by very high penalty costs (under costs), the safety cover is only penalized by a lower unit cost (safety cost). The safety cover is calculated by adding the min cover and some further cover buffering against demand uncertainty. Table 20.4 summarizes the SCL node options. All SCL nodes for one product are connected to each other to enable emergency shipments between DCs. These transports should be avoided as they lead to additional administrative effort. Therefore, these shipments are only performed in the short run if stock-outs impend. The respective cost fields of the arcs are filled with penalties and not with real transportation cost rates.

Monitor nodes are connected to all SCL nodes of one DC to constrain the maximum inventory level.

Table 20.4. StorageCoverLocal node data fields

Data field	Case specific input
Min cover	Minimum stock cover for lot-sizing, transport and quarantine
Safety cover	Additionally buffering against demand uncertainty
Max cover	Maximum stock cover: this bound ought to avoid large inventory build-ups which could result in obsolescence
Under cost	Cost for falling below the minimum stock level (penalty!)
Safety cost	Cost for falling below the safety stock level (penalty!)
Over cost	Cost for exceeding the maximum stock level (penalty!)
Inject	Beginning inventory position of the next week (first planning bucket)
Cost	Inventory holding costs calculated mainly from interest on capital employed
Demand	Forecasted demand of each week

20.4.2 Solution Approaches

An optimal solution can be calculated if minimum lot-sizes are not taken into account. Therefore, the model can be solved using one of the LP procedures offered by the optimization engine (*primal and dual simplex, barrier*) of Strategic Network Optimization. The optimal solution is retrieved in a few seconds or minutes.

However, if the batch nodes are taken into consideration, solution time increases up to several hours. For this kind of integer variables (min run length) special purpose heuristics guide the solution process. Feasible solutions can be computed (without an optimality proof).

20.4.3 Data Flows

All supply chain planning modules of this case study are connected to a common database (Oracle). This database stores all static and dynamic data required for planning. The planning software itself does not provide database connectivity and therefore needs to be integrated by a middle-ware product. This is enabled by Peoplesoft *Advanced Planning Agent* (APAg, see Fig. 20.8) which retrieves data from the Oracle database and converts it to flat files which are accessible for the planning tools.

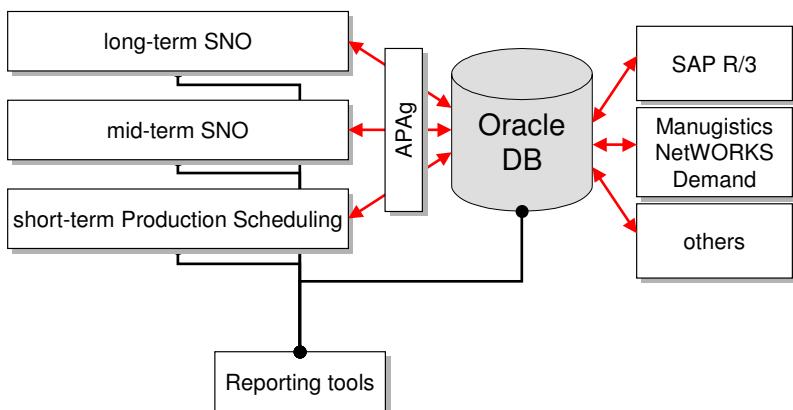


Fig. 20.8. Data flows for integration

Strategic Network Optimization is able to generate the model from a flat file describing each node and arc. Therefore the Strategic Network Optimization model is created from scratch once per week. This automatic procedure bases on the data file generated by the Peoplesoft Advanced Planning Agent component. In addition, an update of dynamic data is possible upon request of the planner. In this case the following data flows are necessary:

- update of forecasts from Demand Planning (SCL node),
- calculation of beginning inventory positions for the first planning bucket (SCL node),
- import of planned production quantities from Production Scheduling within the fixed horizon (arc between packaging process node and working node).

Input to the database is gathered from the planning output of further Peoplesoft modules and from the ERP system SAP R/3 (see SAP (2004)), the demand planning solution Manugistics NetWORKS Demand (see Manugistics (2004)), and some others. Reporting capabilities are necessary for two reasons: First of all the planner needs transparency of data available and all planning results. Therefore, the Strategic Network Optimization reporting capabilities (*smart graphs*) are used to get overviews in tabular form. These smart graphs can be customized and permit simple calculations on data available in nodes and arcs (e.g. calculation of seasonal stock from minimum stock and planned stock level). High-end reporting tools are also required, because the overall success of the project is measured by some core key performance indicators (KPIs). Every participant of the crew needs up-to-date information on these figures. The KPIs are based on planning *and* actual data which is only available in the database. Therefore, standard database reporting tools are used to build customized reports for this purpose.

20.5 Results and Lessons Learned

The Strategic Network Optimization model described above and the Production Scheduling model have gone “live” some years ago. To identify major mistakes in the implementation, both the new APS and the old planning system ran in parallel for just a few weeks.

Thus the question on benefits can now be answered with several years’ experience of practical use. The most important improvements measured are:

- *Reduced planning time:* The planning time needed to create the weekly master plan was reduced to less than 30% of the time required before the implementation of the APS. About 95% of the decisions proposed by the planning system are put into practice without changes. This shows the reliability of the planning model and enables the planner to examine what-if analysis to evaluate process changes.
- *Reduced inventory levels:* Without affecting customer service, it was possible to avoid buffers which were used to protect against planning inaccuracy and uncertainty. This also attributed to a new way of controlling inventories which is based on the inventory analysis methodology introduced in Chap. 2.
- *Reduced overtime:* The bottleneck production lines can be identified very early. Therefore, production can be shifted to alternative sites, if necessary.

- *Less emergency transports between DCs:* More accurate planning balances production and transports in advance.

The software template described so far has also been used for fast implementation of Peoplesoft Advanced Planning in other food-divisions of the company. This was possible because the template is based on intensive, collaborative analysis of several food and beverages supply chains covering the company's whole food-branch.

References

- Günther, H. O.; Blömer, F.; Grunow, M. (1998) *Moderne Softwaretools für das Supply Chain Management*, Zeitschrift für wirtschaftlichen Fabrikbetrieb, Vol. 93, No. 7–8, 330–333
- ILOG (2004) *Homepage*, URL: <http://www.ilog.com>, Date: July 19, 2004
- J. D. Edwards (2001) *Strategic Network Optimization, User's Guide*, Denver
- Kolisch, R. (1998) *Linx*, ORSpektrum, Vol. 20, No. 1, 1–4
- Manugistics (2004) *Homepage*, URL: <http://www.manugistics.com>, Date: July 19, 2004
- Oracle (2004) *Homepage*, URL: <http://www.oracle.com>, Date: July 19, 2004
- Peoplesoft (2004) *Homepage*, URL: <http://www.peoplesoft.com>, Date: July 19, 2004
- SAP (2004) *Homepage*, URL: <http://www.sap.com>, Date: July 19, 2004

21 Computer Assembly

Christoph Kilger¹ and Lorenz Schneeweiss²

¹ j & m Management Consulting AG, Kaiserringforum, Willy-Brandt-Platz 5, 68161 Mannheim, Germany

² Nestlé GLOBE Center Europe GmbH, Rohnestrassse 4, 60523 Frankfurt/Main, Germany

The computer industry is a typical example of a *material constrained* supply chain. The main bottleneck of demand fulfilment is the availability of the electronic components, e.g. disk drives, processors, memory etc. This case study is based on an actual APS implementation project at a large international computer manufacturer. Four modules of the APS system by i2 Technologies are implemented, supporting the demand planning process, the mid-term supply planning process, the short-term supply planning process and the order promising process. The following case study describes in detail

- the computer assembly supply chain, the product structure and the assembly process (Sect. 21.1),
- the scope and the objectives of the APS implementation project (Sect. 21.2),
- the planning processes being supported by the APS system, i.e. demand planning, operational planning, order planning, order promising and the integration of the applied i2 planning modules with the existing SAP R/3 system (Sect. 21.3), and
- results and lessons learned from the APS implementation (Sect. 21.4).

21.1 Description of the Computer Assembly Case

21.1.1 Computer Industry Supply Chain

The typical supply chain in the computer industry consists of five main stages: suppliers, computer manufacturers, logistic service providers, deployment partners and customers. Fig. 21.1 depicts the complete computer industry supply chain.

The suppliers supply electronic and mechanical components for the system board and computer assembly, external units like printers and monitors, accessories like keyboard and mouse, software, manuals etc. In many cases there are multiple sources for one type of components like disk drives and memory (disk drives supplied by one supplier may be substituted by disk drives from an alternate supplier if these are qualified for the corresponding configuration). The computer industry is a material intensive industry. In general approximately 15–20 suppliers constitute 80% of the procured value, 30–40 suppliers represent 95% of the procured value. Some suppliers provide

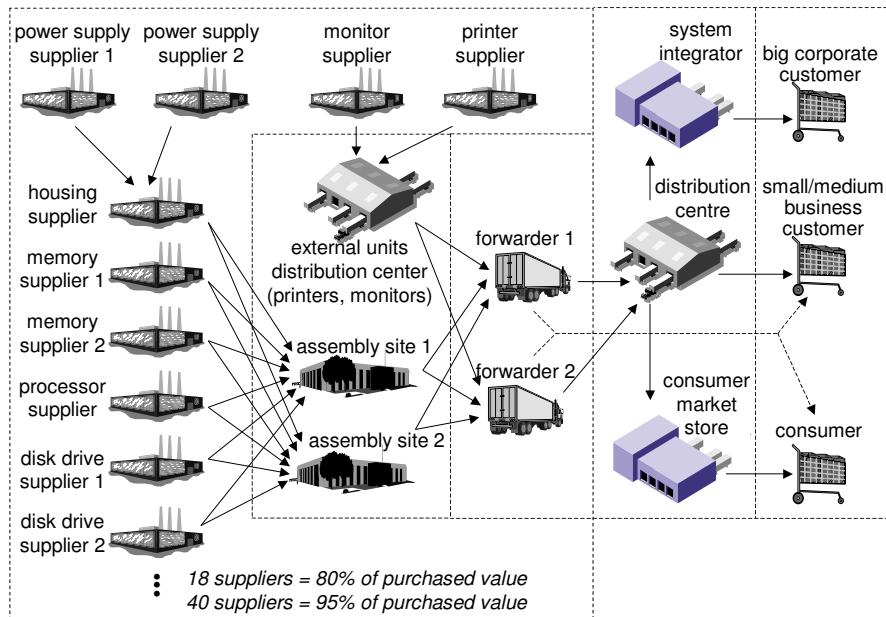


Fig. 21.1. Overview of the computer industry supply chain

simple assembly services. In the supply chain shown in Fig. 21.1 the housing supplier receives power supplies from two alternate sources and assembles the power supplies into the housings before shipping the housings to the computer assembly site.

The computer manufacturing process itself consists of two main parts: the assembly of the system board and the assembly of the system unit. There are three options to organize this part of the supply chain:

1. System board assembly and system unit assembly are done in the same assembly site.
2. System board assembly and system unit assembly are done in separate assembly sites, but belong to the same legal entity.
3. System board assembly and system unit assembly are done in separate assembly sites and belong to different legal entities.

In this case study we assume option 1, as depicted in Fig. 21.1. Typically, the computer manufacturer runs a separate distribution centre for external units like printers and monitors that are procured from external suppliers.

The transport between assembly sites and the deployment partners is executed by logistic service providers. There are three kinds of deployment partners: logistics service providers (forwarders) running a distribution centre, system integrators and consumer market stores. In most cases products are shipped by a forwarder from the assembly site and from the distribution

centre for external units like printers and monitors directly to a distribution centre, where the separate line items of a customer order are merged. From there, complete customer orders are shipped.

There are three cases for the shipment of customer orders, depending on the type of customer:

- Orders by small and medium business customers are shipped directly from the distribution centre to the customer's site.
- Big corporate customers like banks and insurance companies typically place orders with a volume of up to several thousand PCs (e.g. in order to equip all offices in a specific region). These big orders are often executed by a system integrator, who takes over responsibility for the procurement and the installation of the computing devices (PCs, servers, monitors, printers, networks, modems etc.).
- For the consumer market, department stores and consumer market stores place big orders (in the range of 10 000 to 20 000 units) that are shipped to their distribution centres and are distributed from there to the individual outlets.

As indicated by the dashed arrows in Fig. 21.1 additional direct distribution paths from the computer manufacturer to the consumer market and to small and medium businesses will be established, supported by e-business strategies.

21.1.2 Product Structure

The product portfolio contains consumer PCs, professional PCs, servers and notebooks. Within these product families, two types of products can be distinguished. *Fixed configurations* have an individual material code that can be referred to in customer orders. Normally, fixed configurations are made to order. However, in order to offer very low lead-times to the market (for example two days) a make-to-stock policy can be applied. This requires a very good understanding of future demand of these market segments, i. e. a high forecast accuracy.

Open configurations can be freely configured by the customer (configure-to-order). An open configuration is identified by the *base unit*, specifying the housing and the system board. The customer can then choose from a selection of processors, disk drives, network, video and sound controllers and can define the size of the main memory. During the configuration process specific configuration rules have to be fulfilled. Examples of hard configuration constraints are “the number of controller cards may not exceed the number of extension slots of the system board” and “the selected processor type must be compliant with the system board”. An example of a soft constraint is “the number of selected CD-ROM drives should not exceed one”. Only the base unit and the components have individual material codes. The complete

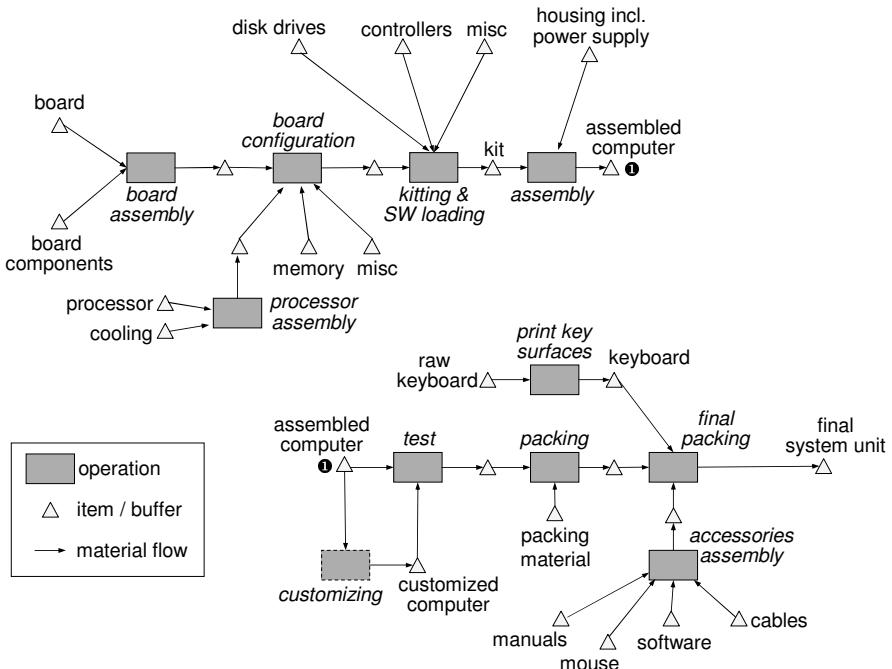


Fig. 21.2. Detailed computer assembly supply chain

configuration is either identified by the material code of the base unit (this requires a hard pegging of production orders to the customer order) or a new material code is generated as final step of the configuration process.

21.1.3 Computer Assembly Process

The computer assembly process is divided into two main parts (as shown in Figs. 21.2):

- the assembly operations and
- the testing and packing operations.

The first step is the assembly of the system board. Boards are assembled in batches of 100 to 1 000 pieces. The board assembly lead-time for one batch is roughly half a day. There are approximately 20 different system boards for PCs and another 20 for servers. The system boards for notebooks are procured.

The second step is the configuration of the board. In this step, the processor assembly – consisting of the processor and the cooling – and the memory are put onto the board.

The third step is the kitting and the loading of the disk drive with the selected software. The kitting operation collects all selected components –

disk drives, controller cards for network, video and sound etc. – into a box that is called the kit. The kit, the housing and the power supply – which are not part of the kit – are used in the fourth step, the actual assembly of the computer.

If the customer has special requests – e.g. specific controller cards that have to be assembled into the computer – a separate customization step follows the computer assembly operation. After that, the computer is tested and packed. In the final packing operation the keyboard and accessories as mouse, manuals, software, cables etc. are added.

The complete lead-time is 24–48 hours. The most time consuming operations are the software loading and the test operations. There are two production types: small batches (usually below 200 PCs) are assembled in a job shop, large batches (above 200 PCs) are assembled in a flow shop. Please note that kitting takes place only for the job shop production type. In the flow shop, the material for the complete batch is provided along the production line.

Table 21.1 summarizes the classification of the computer industry according to the supply chain typology introduced in Chap. 3.

21.2 Scope and Objectives

The target of the APS implementation project described in this case study is to improve the business performance of the computer manufacturer. For this purpose the business performance is measured by three key performance indicators (see Chaps. 2 and 15):

- The forecast accuracy shall be improved from 50 % to 80 %.
- The delivery on time shall be improved from < 80 % to > 90 %.
- The inventory turns shall be improved from 9,3 to 20.

The following planning processes are in the scope of the project and shall be supported by the APS by i2 Technologies:

- *demand planning*, consisting of unit planning and component planning,
- *operational planning* (mid-term supply planning), consisting of the Weekly SCM Workflow (forecast netting, master planning, allocation planning) and the SAP MRP run,
- *order planning* (short-term supply planning), consisting of the Daily SCM Workflow (forecast netting, master planning and allocation planning) and demand supply matching, and
- *order promising*.

The following modules of i2 were implemented:

- Demand Planner (DP) including PRO (Product Relationship Object, a module of demand management to support the component planning process) supporting the demand planning processes,

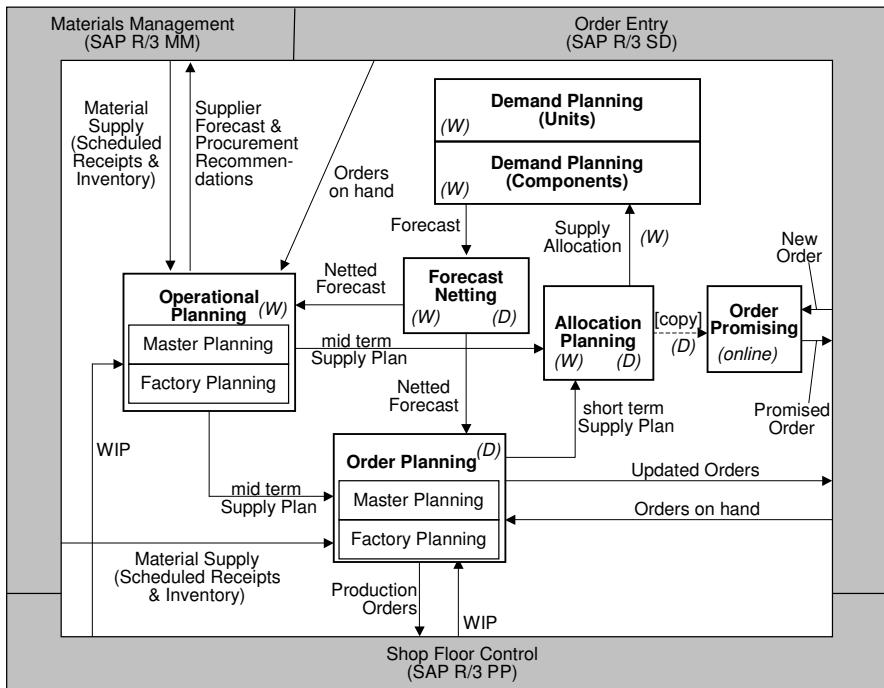
Table 21.1. Supply chain typology for the computer industry

Functional attributes	
Attributes	Contents
number and type of products procured	many, standard and specific
sourcing type	multiple sourcing
supplier lead time and reliability	long, unreliable
materials' life cycle	short
organization of the production process	flow shop and cellular
repetition of operations	larger / smaller batches
bottlenecks in production	low importance
working time flexibility	high
distribution structure	two and three stages
pattern of delivery	dynamic
deployment of transportation means	individual links
availability of future demand	forecasts and orders
shape of demand	weakly, seasonal
products life cycle	several months
number of product types	few / many
degree of customization	standard / customized
products' structure	convergent
portion of service operations	tangible goods

Structural attributes	
Attributes	Contents
network structure	mixture
degree of globalization	several countries
location of decoupling point(s)	assemble-/configure-to-order
major constraints	material
legal position	inter- and intra-organizational
balance of power	suppliers and customers
direction of coordination	mixture
type of information exchanged	orders

- Factory Planner (FP) supporting the demand supply matching process,
- Supply Chain Planner (SCP) supporting the master planning process,
- Demand Fulfillment (DF) supporting the forecast netting, the allocation planning and the order promising processes,
- RhythmLink (RL) and
- Active Data Warehouse (ADW)
- Optimization Interface (ROI)
- Collaboration Planner (RCP).

The i2 modules are integrated with the existing SAP R/3 system, i.e. MM Materials Management, SD Sales and Distribution and PP Production Plan-



(D) daily planning cycle; (W) weekly planning cycle; (online) online process

Fig. 21.3. Computer assembly planning processes

ning. Figure 21.3 summarizes the supported processes and the data flows between them.

All implementations and process designs were finished by July 2001 and are being used productively since. Now only few adaptations are made to the existing implementation to increase performance and reporting possibilities. The project started in 1999, the total implementation time was scheduled for 18 months, including a two months phase in which the APS implementation project had been defined and the APS software was selected.

21.3 Planning Processes in Detail

21.3.1 Demand Planning

The demand planning process is running weekly. It determines the forecast on unit level (i.e. finished goods) and on component level. The implementation was structured into three steps. Step 1 covered unit planning for a subset of all products, i.e. planning of complete computer systems (units),

Step 2 extended unit planning to all products and Step 3 supports component planning.

The goal of the implementation of i2 Demand Planner is a more accurate forecast: the forecast accuracy shall be improved from 65% to 75%.

The following technical “enablers” of i2 Demand Planner help to improve forecast accuracy:

1. i2 DP provides a common database to maintain all input and output data of the demand planning process.
2. i2 DP supports a collaborative planning process, where all departments participating in the demand planning process find their own planning results and are supported in the integration of the various plans to one collaborative demand plan.
3. All needed data like shipments and orders are maintained within i2 DP and can be used within the planning process.¹
4. All groups participating in the demand planning process can define their own views on the data.
5. Forecast accuracy is measured based on the i2 DP database, using a well defined uniform method that has been defined within the project.

Unit Planning

Table 21.2 shows the levels of the geographic and the product hierarchies used in the demand planning process. The numbers given in parentheses specify the number of instances on that level.

All_Geo is the root of the geographic hierarchy. The area level represents geographically defined areas in the world, e.g. Europe, Middle East and Africa (EMEA); America; Asia Pacific. The regional level represents sales regions within an area, e.g. Germany, France and the UK are regions within the EMEA area.

All_Prod is the *root* of the product hierarchy. The *product segment* level divides the product hierarchy into sub-hierarchies for PCs, servers, notebooks and the planned components (see next subsection). On the *product group* level each sub-hierarchy is split into multiple product groups, e.g. the PC sub-hierarchy is split into consumer PCs and professional PCs, and the server sub-hierarchy into small servers and large servers. The next level is the *product family* that groups products which are in the same performance class (low-end consumer PCs vs. high-end consumer PCs). The *model line* groups PCs and servers by the type of the housing. The *sub model line* groups PCs and servers within one model line by the type of the system board. The *SKU* level is normally not planned for units (refer to the next section about component planning for explanation why the SKU level is in the product hierarchy).

¹ In fact, because of the integration of all actual data and the daily update of the actuals, i2 DP is also used as a management reporting tool.

Table 21.2. Structure of the geographic and product hierarchies

Level	Geographic Hierarchy	Product Hierarchy*
1	All_Geo (1)	All_Prod (1)
2	Area (6)	Prod_Segment (4+10)
3	Region (> 40)	Prod_Group (10+20)
4		Prod_Family (20+50)
5		Model_Line (40+100)
6		Sub_Model_Line (80+200)
7		SKU (200+400)

* The number of instances on that level are given in parentheses. The product hierarchy includes the instances used for the unit planning process and those used for the component planning process, denoted as $(a + b)$, where a is the number of units on that level and b is the number of components.

The time dimension is structured into Year, Quarter, Month and Week. The weekly level is not used for the unit planning, but for the component planning. The time horizon starts two years before the start of the current fiscal year and covers 12 months into the future. Thus, it is possible to maintain two years of historic data in the i2 DP database. This provides a good basis to setup stochastic forecasting methods including seasonal patterns. However, two years of historic data are currently not available. The average product life cycle is about five months. Thus, a large manual effort is required to define the historic substitution rules that are used to map historic data of products that have reached their end of life to living products.

The following data rows have been defined and are maintained in the i2 DP database:

- *Actual data:* Three types of actual data are maintained: Shipments (quantities related to shipment date), orders (quantities related to customer requested date) and confirmed open orders (quantities related to confirmed date).
- *Budget plan:* The budget plan is updated yearly and is valid for the current fiscal year.
- *Sales forecast:* The sales forecast is created monthly and covers six months. The database contains four separate rows for the sales forecast, representing the current planning round and the last three planning rounds.
- *Plant forecast:* The plant forecast is created weekly by the planners in the production sites. The database contains four separate rows for the plant forecast, representing the current planning round and the last three planning rounds.

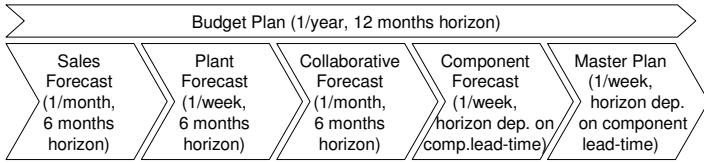


Fig. 21.4. Planning processes supported by i2 Demand Planner

- *Collaborative forecast:* The collaborative forecast is determined monthly by a collaborative process in which sales, product management, procurement and the production sites participate.

There is a yearly, monthly and weekly planning cycle as shown in Fig. 21.4. Once per year, the budget plan is created, covering the next fiscal year (12 months horizon). Every month, the sales planners update the sales forecast. For that purpose, i2 DP has been installed in all regional sales offices (approximately 50). In the second step, the sales forecast is reviewed by the planners in the plants; the result of this step is the plant forecast. In a final step, a collaborative forecast is created based on the sales and the plant forecast. All three types of forecast cover six months. The collaborative forecast is adjusted every week by the planners in the plants, resulting in a weekly plant forecast. The weekly plant forecast is the basis for the weekly component planning, that is described in the next subsection.

Component Planning

The business environment of this computer manufacturer was selected to be build-to-order and configure-to-order for the main part of the business. As the consequence of this decoupling point decision the main purpose of the monthly and weekly forecasting processes is the creation of an accurate forecast on component level. The focus of the component planning process is therefore to generate a supplier forecast for all dependent components derived from the unit forecasts. Out of the approximately 2 000 components, 600 components are considered during component planning (A-parts). The planned components belong to the material groups processors, memory, disk drives, controllers, housing and power supply.

An important aspect of component planning in the computer industry is the specification of particular components by large customers. For example, a large customer makes a contract (forecast) of 5 000 PCs which are configured to meet the IT-requirements of that customer and shall be delivered over five weeks (1 000 PCs a week). In this case, a component forecast can automatically be derived from the collaborative forecast by exploding the bill-of-materials of the particular configuration. The other standard case is that a PC is configured during order entry – in this case the sales forecast does not specify a particular configuration.

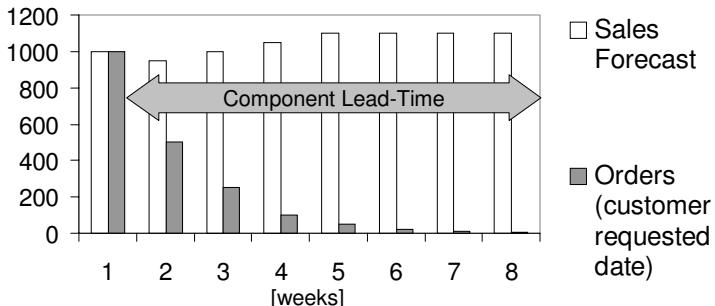


Fig. 21.5. Order lead-time vs. component lead-time

In order to meet these two requirements – fixed specification of components (build-to-order) vs. online configuration of components at order entry (configure-to-order) – the following procedure is applied to derive a component forecast from the collaborative forecast:

1. The collaborative forecast has to be split into (1) forecast related to fixed configurations and (2) forecast related to open configurations (that are still to be configured).
2. For forecasting fixed configurations, the bill-of-materials of the fixed configuration is being exploded.
3. For forecasting open configurations, the following steps are followed:
 - (a) So-called mappings are defined that map some planned instances on finished goods level (e.g. a model line) onto planned components (e.g. disk drives, processors etc.). A mapping is established between a planned item A on finished goods level and all components C that can be configured into products of type A .
 - (b) The distribution of the forecast on some planned item A on finished goods level over all components related to A by some mapping is defined by attach-rates (i. e. distribution factors). The actual planning process is to determine these attach-rate factors.
4. The total component forecast of a component is derived by adding the forecast from Step 2 and Step 3.

This component planning procedure is supported by i2 PRO (Product Relationship Object).

21.3.2 Operational Planning

Operational Planning consists of the Weekly SCM Workflow and a consecutive MRP run in SAP not described in detail here.

Weekly SCM Workflow

The Weekly SCM Workflow consists of forecast netting, master planning and allocation planning and serves two purposes:

1. It calculates the total supply and capacity needed to fulfil the demand within the planning horizon and forms the basis for negotiations with suppliers and purchasing decisions.
2. It constrains the demand based on the feasible supply and serves as a medium to communicate deviations of forecast and availability.

The demand planning process generates the forecasts for all products and components. This forecast is updated weekly and is netted against the actual orders received (forecast netting process, see Fig. 21.3). In the short-term the forecast for certain products or customers could already be realized in the form of actual orders. Therefore the remaining forecast has to be determined by a netting of actual orders against their forecast to keep a constant demand signal for the master planning process. The forecast is netted on the given level from the demand planning process, i. e. on end item level for fixed configurations and on component level for open configurations. The master planning process receives the netted forecast and actual orders and creates a fulfilment plan for the demand (actual orders and netted forecast). The fulfilment plan is then allocated to the customers according to the received customer orders and their netted forecast values.² This total number can now easily be compared to the original forecast.

The Weekly SCM Workflow is executed twice per week. In a first run the updated forecast plan from the demand planning process is taken and a fulfilment plan is generated publishing reports with defined problems in the supply chain. After the first run the exception handling starts and modifications are made to supply and demand data to solve the problems. During the Weekly SCM Workflow, the planners

- decide about sourcing options (sourcing from multiple suppliers, sourcing from multiple plants, use of alternate parts),
- generate supply requirements based on the netted forecast, including safety stock management decisions,
- generate a constrained demand plan on forecasted item level based on the netted forecast and the actual customer orders on hand and
- generate production requirements for make-to-stock forecast.
- decide about forecast shifts from one product to another due to supply constraints.

The modifications have to be finished before the second run starts. The second run then finalizes the planning cycle with updated results from the exception handling. After a final review the plan is accepted and released.³

² Refer to Chap. 9 for a detailed description of the allocation planning process.

³ Refer to Chap. 8 for more information on the master planning process in general.

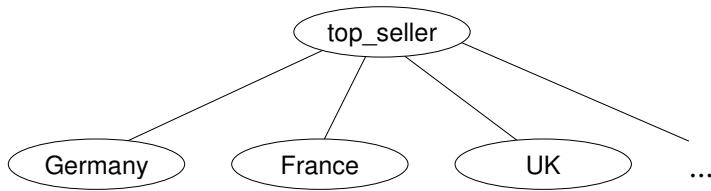


Fig. 21.6. Customer hierarchy used for allocation planning

On the one hand side the master plan defines the minimum purchasing volume that is to be ordered during the next weeks based on the constraints, i. e. demand, material supply and capacity. The capacity model that is applied to master planning is quite simple, as it is based on the number of computers that can be assembled per day. On the other hand side, the weekly master plan is the basis for order promising. Thus, the master plan captures the maximum order volume that can be promised during the next weeks.

The master planning process is implemented based on i2 Supply Chain Planner. Forecast netting and allocation planning are supported by i2 Demand Fulfillment. Only planned components, planned configurations on end item level and finished goods for which orders exist are represented in the master planning process.⁴

21.3.3 Order Planning

The Order Planning consists of Daily SCM Workflow and the Demand Supply Matching Process.

Daily SCM Workflow

The purpose of the Daily SCM Workflow is to generate latest ATP (Available To Promise)⁵ information for online order promising. It represents actual and future availability of supply and capacity that can be used to accept new customer orders. The promises should be given based on the availability of the planned components and guarantee high delivery on time. The daily planning run ensures an up to date ATP picture, reacts on changes in base data (e. g. BOMs) and handles exceptions that could not be foreseen in the Weekly SCM Workflow following pre-defined business rules. The process runs in batch mode without user interaction.

The Daily SCM Workflow – similar to the Weekly SCM Workflow – generates a fulfilment plan based on the released forecast figures from the weekly planning cycle. In the allocation planning process the fulfilment plan is allocated to the customer hierarchy (see also Chap. 9). The customer hierarchy

⁴ Purchasing decisions for non-planned components (B- and C-parts) are taken in the SAP MRP run.

⁵ Refer to Chap. 9 for more information on ATP in general.

is a sub-hierarchy of the geographic hierarchy used in the demand planning process. Currently, the customer hierarchy contains the root top-seller and one node for each sales region (Fig. 21.6). The allocation planning processes allocates the ATP to the nodes of the customer hierarchy, according to the following rules:

- The quantities planned in the master plan for pre-defined fixed configurations are allocated to the sales regions according to the sales forecast. The following table shows an example for one selected week (using the allocation rule *per committed*, refer to Chap. 9):

	Total	Germany	France	UK
Sales Forecast	10 000	5 000	3 000	2 000
Master Plan	8 000			
Allocations	8 000	4 000	2 400	1 600

- The quantities planned in the master plan for open configurations are allocated on component level at the root of the customer hierarchy (top-seller).

The allocations are called *allocated ATP (AATP)*. The ATP and the allocations are provided to the order promising process and are given as feedback to the demand planning process (Fig. 21.3). Thus, the demand planners have the overview over the ATP quantities that are allocated to them compared to what they have forecasted. This information can be used to direct demand according to the ATP situation, e. g. by suggesting alternate products to the customers driven by availability.

Demand Supply Matching

In addition to the Daily SCM Workflow there is a more detailed planning process on factory level that runs daily (see Fig. 21.3). This process is called the *Demand Supply Matching Process (DSM)*. The DSM process plans the production and material assignment of all customer orders and the net forecast for build-to-stock products based on the complete bill-of-materials as being maintained by SAP. Thus, DSM checks the demand and supply situation for all parts – whereas Master Planning only checks parts that are planned by the demand planning process as described above (A-parts).

The DSM process is executed twice per day by a group of planners, representing purchasing, order management and production. These planners review the current demand and supply situation, check options to resolve late order problems, try to improve the supply situation by moving-in scheduled receipts in coordination with the suppliers, and simulate the impact of moving-in of orders. Execution is not triggered directly by the DSM process. To close the loop to execution which is triggered by SAP a part of the late orders is transferred to SAP with a new due date. The new due date generates a new manufacturing order to a later date in SAP. The demand supply matching process is implemented based on i2 Factory Planner.

21.3.4 Order Promising

The orders enter the system through the order entry process and are promised by the order promising process. To promise a new customer order the order starts searching for allocated ATP in the dimensions time, seller and product (refer to Chap. 9). Several consumption rules define how a new order can find ATP. The promising policies assigned to the orders define if the order is promised e.g. as a whole or in several partial deliveries. Again, one must distinguish between fixed configurations and open configurations:

- Let us assume an order is received from a customer in France for x units of fixed configuration f with a request date for week w . The order promising process checks the quantity for the fixed configuration f that is allocated to France in week w ; let us call this a . If the ordered quantity x is less than the allocated quantity a the order receives a due date in week w . If this is not the case, i.e. $x > a$, then additional ATP is searched in the preceding weeks – even if ATP is available in week w at other nodes of the customer hierarchy, e.g. Germany and the UK. This consumption rule ensures that quantities that have been planned by some region are reserved for orders coming from customers of that region.
- Orders for open configurations are quoted based on the ATP for components. The order promising process searches the best ATP for each of the components required for the order. The latest ATP plus the configuration's lead-time is assigned as due date to the complete order.

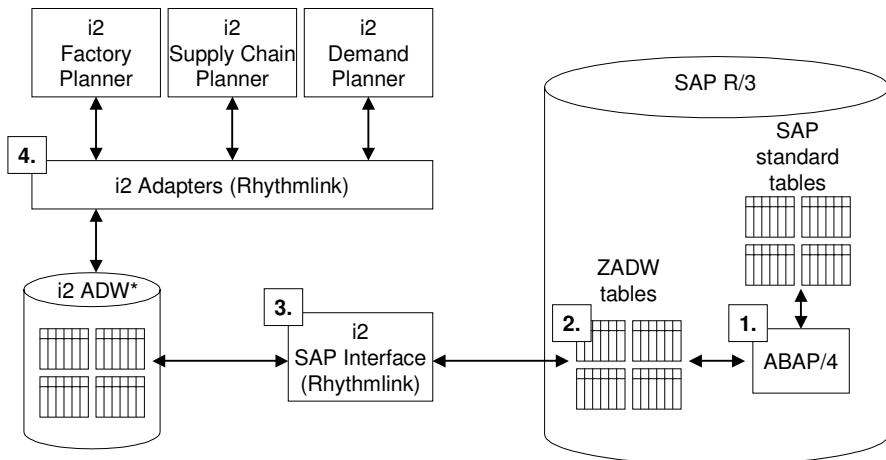
21.3.5 Integration of i2 with SAP R/3

The i2 planning engines – Supply Chain Planner, Demand Fulfillment, Factory Planner and Demand Planner – closely interact with the SAP R/3 system, particularly with the SAP modules MM Materials Management, PP Production Planning and SD Sales and Distribution. In this case the SAP R/3 release 4.6c was installed.

Figure 21.3 shows the interfaces between the SAP system and the i2 modules. There are two classes of interfaces. The first class contains all interfaces except the order entry interface. These interfaces exchange static and dynamic data in a batch mode. The second class consists of the order entry interface. This interface transfers a new order from SAP SD to i2 Demand Fulfillment and gives the order quote back to SAP SD. The order entry interface is an online interface.

Batch Interfaces

Figure 21.7 shows the architecture used for the interfaces operating in batch mode. In the following we describe the data flow from SAP to i2. The data flow back is implemented in the same way. We use the interface between SAP and the supply planning processes as example to illustrate the ideas.



*Active Data Warehouse

Fig. 21.7. Integration of the i2 planning modules with SAP R/3

1. The supply planning processes represent only planned materials, i.e. those materials that are also used in the demand planning process. Furthermore, all orders and the forecast are imported, as well as WIP quantities, scheduled receipts, inventory etc. The selection, filtering and aggregation of the SAP data according to the data requirements of the supply planning processes are executed by a collection of ABAP/4 functions. These functions have been developed specifically for the project.
Please note that the filter and aggregation functions applied represent the actual business logic on which the design of the supply planning processes is based. Thus, using a pre-defined standard interface between SAP and some APS Master Planning Module would constrain the design of the process and the interface – potentially preventing that a best-in class master planning process is achieved.
2. The ABAP/4 functions write the filtered and aggregated data for the supply planning processes in user-defined tables in the SAP database. These tables – called ZADW tables – have the same data scheme as the tables that exist in the i2 Active Data Warehouse (ADW).
3. The content of the ZADW tables is transferred into the tables of the ADW, using the i2 standard SAP interface. This interface is based on the middleware module i2 RhythmLink. RhythmLink has a specific module – the SAP-Listener – that is responsible for the technical data transfer between SAP and RhythmLink. Data streams are opened by RhythmLink copy maps that transfer the data from SAP directly into the corresponding ADW table.
4. After the complete data arrived in the ADW the standard i2 adapters are used to provide the data for the i2 planning engines e.g. the Supply

Chain Planner engine that is running the master planning process. The i2 adapters are standard software components that are shipped with each i2 module. The adapters provide all required interfaces between the i2 module and the ADW (in both directions).

Order Entry Interface

The order entry interface is an online interface. When an order is received by the order entry system SAP SD, the order is transferred to i2 Demand Fulfillment to be quoted. The quoted order is then sent back to SAP SD.

Technically, the order entry interface is based on the Optimization Interface (ROI). This interface consists of a collection of predefined ABAP/4 functions that have to be plugged into the order entry transaction as an user exit. The Optimization Interface then transfers the order to i2 Demand Fulfillment via API string and receives the quote the same way. The quote information is written into the SAP order, and the transaction is closed. An order can be quoted in milliseconds (below 100 ms per order, more than 10 orders per second can be quoted). Given that currently 800 orders per hour have to be quoted in peak load situations, the i2 Demand Fulfillment architecture scales well with an increasing number of orders – supporting even aggressive business growth strategies. The online order promising solution is used in combination with i2's High Availability architecture that is based on the TIBCO message bus system (refer to www.tibco.com). This architecture supports 24x7 (24 hours a day, seven days a week) order quoting even in case of server or network failures. This architecture consists of one primary and several secondary order promising servers that replicate all transactions on the primary server. So in case of a crash of the primary server a secondary can take over seamlessly.

21.4 Summary and Lessons Learned

The APS implementation resulted in major improvements of the planning and logistics processes and helped to improve major KPIs. Tab. 21.3 lists the improvement of logistical KPIs from 1999 to 2002 and the target value. The improvement of the customer service level and the delivery reliability resulted in additional revenue. Better forecast accuracy helped to reduce the inventory levels and by that reduced the direct material costs by approx. 0.3 % – 2 %. Through better planning support the inbound logistics costs and the process costs in purchasing and planning departments could be decreased.

In addition to these business improvements the following “Lessons Learned” can be summarized from the project work:

- The batch interfaces between i2 and SAP R/3 were easier to implement and to manage than expected. For example the adaption of the interface

Table 21.3. Improvements of KPIs due to the APS implementation

KPI	1999	2002	Target
Forecast Accuracy	50 %	70 %	80 %
Delivery On Time	< 80 %	≈ 90 %	> 90 %
Inventory Turns	9,3	14,1	20
End-to-end Order Lead Time	10 – 22 days	6 – 12 days	—

programms from SAP R/3 3.0 f to 4.6 c was accomplished within 6 weeks without support by external consultants.

- The online integration between i2 Demand Fulfillment and SAP R/3 SD turned out to be rather difficult to implement and stabilize. Especially the consideration of the SD order types and the specific customizing of SD was a source of many issues.
- The learning curve of planners and other employees working with the APS took more time and effort than expected. The use of an exception-based APS and its planning algorithms compared to the use of a transactional system like SAP R/3 required additional training and management of change activities.
- However, after these management of change activities the APS implementation lead to a better integration and synchronization of the planning and execution processes.
- The benefits from the implementation could only be realized through a clear assignment of each major KPI to one responsible manager who leads the monitoring, reporting and improvement activities related to this KPI. Based on this a continous improvement process was established that is driven by the KPI-managers (see also (Kilger and Brockmann, 2002)).

References

Kilger, C.; Brockmann, K. (2002) *Supply Chain Detectives in der High Tech Industrie*, in: Dangelmaier, W.; Emmrich, A.; Kaschula, D. (Eds.) Modelle im E-Business, Paderborn, 139-154

22 Demand Planning of Styrene Plastics

Boris Reuter

axentiv AG, Uhlandstraße 9, 64297 Darmstadt, Germany

The purpose of Demand Planning is to reduce uncertainty about what will be sold to the customer in the future. Improving forecast accuracy leads to economic benefits such as cost cutting by reducing safety stocks, and increasing sales by avoiding stock out situations. The case presented describes a Demand Planning implementation using mySAP SCM at an international company in the process industries.

The case study is structured as follows: first, we start with a description of the supply chain, then we introduce the architecture of the planning system used. Subsequently, we focus on model building with the module Demand Planning of SAP's Advanced Planner and Optimizer (SAP APO DP) as a part of SAP's cross-industries solution mySAP SCM. The process of the Demand Planning of styrene plastics shows the different tasks, responsibilities and dependencies. Finally, we finish with some concluding remarks about the benefits and lessons learned.

22.1 Description of the Supply Chain

This case study is about a project at one of the world's leading chemical companies. Here we focus on the Demand Planning of styrene plastics in the European plastics division.

Styrene plastics are all-purpose plastics and can be found in a multitude of different consumer products such as CD cases, packaging, computer chassis, monitors or printers. The main sectors for the products are electronics communication, consumer electronics and computer, leisure, household, packaging and film, and medical technology.

The main raw material to produce styrene plastics is monostyrene, which originates from crude oil. The polymerization of monostyrene leads to "high impact" or "general purpose" polystyrenes – so-called product families – depending on whether further additives such as rubber are used. The assortment is made up of approximately 500 products, from commodities to specialities, including colored or fire-resistant granulates. The products are wrapped in different packaging leading to a total of 1,500 stock keeping units. From the logistic point of view the products can be combined into approximately 50 product groups (see Fig. 22.1).

The price is an important marketing instrument because a large amount of the quantities sold is made with commodity products. Furthermore, only a few large competitors are in the same market. Thus, changes in the quantities

of one market player have noticeable effects on all others – a typical oligopoly situation. Because of the direct dependency on the rising price of crude oil on the one hand and the selling in consumer product markets with decreasing prices and quantities on the other hand, the margin has been under pressure. The sales activities are organized regionally. This means that one or more countries are combined to form planning regions, e.g. Germany, Switzerland and Austria are in the same planning region, while Portugal and Spain are in another one. A key account structure, which is made up of international business partners with a large portion of the sales volume, exists in parallel.

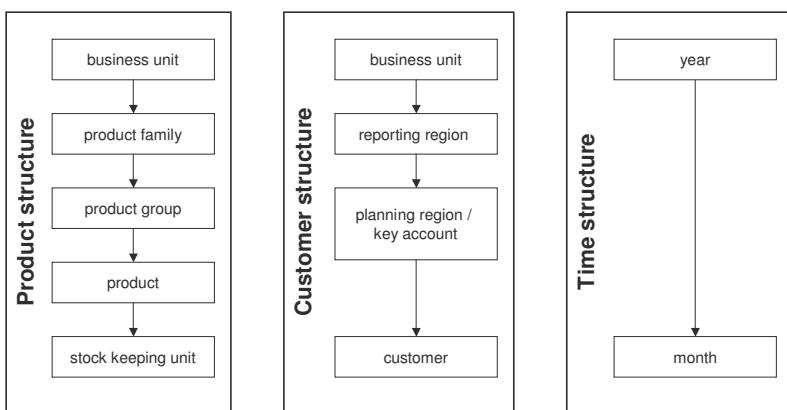


Fig. 22.1. Product, customer and time structures

Further characteristics of the supply chain type can be taken from Table 22.1

Increasing the market share is at the expense of competitors and easily leads to price reductions and thus to decreasing margins. Because of the market situation, the economic benefits can only be obtained by decreasing costs. Precise forecasts are needed to achieve adequate inventories and balanced utilization rates throughout the supply chain. The main aims of this project have been to increase forecast accuracy and responsiveness by the effective use of all information in the system, and to secure the market share and gain a higher profitability in the organization.

22.2 The Architecture of the Planning System

The demand planning of styrene plastics is embedded in a more complex planning architecture (see Fig. 22.2), which is implemented using SAP APO (see Chap. 18.3). This process consists of the following:

- demand planning using SAP APO Demand Planning,
- midterm production planning using SAP APO Supply Network Planning (SNP),

Table 22.1. Typology for the styrene plastics supply chain

Functional attributes	
Attributes	Contents
Number and type of products procured	Few, standard (raw materials, packaging) Multitude (additives)
Sourcing type	Double/multiple
Supplier lead time and reliability	Short, reliable
Materials life cycle	Long
Organization of the production process	Flow line
Repetition of operations	Batch
Changeover characteristics	Weak sequence dep. setup times and costs
Bottlenecks in production	Known, shifting
Working time flexibility	Low, but high machine throughput flex.
Distribution structure	One stage, regionally organized
Pattern of delivery	Dynamic
Availability of future demands	Forecasted (3-24 months)
Demand curve	Almost stationary
Product's life cycle	Several years
Number of product types	Many
Degree of customisation	Blending and packaging
Bill of materials (BOM)	Convergent (blending)/ divergent (packaging)
Portion of service operations	Tangible goods
Structural attributes	
Attributes	Contents
Network structure	Mixture
Degree of globalization	International
Location of decoupling point(s)	Make-to-stock/make-to-order
Major constraints	Capacities of flow lines
Legal position	Intra-organizational
Balance of power	Customers, but oligopoly
Direction of coordination	Mixture
Type of information exchanged	Forecasts, orders and contracts

- short-term production planning using SAP APO Production Planning and Detailed Scheduling (PP/DS), and additional optimization functionality of an ILOG Cartridge,
- demand fulfilment using SAP APO global Available-To-Promise (global ATP) and
- procurement Planning using SAP R/3.

In the Demand Planning (DP) module, planned sales quantities – demands – are collected from customers (e.g. collaborative forecasting), forecasted by experienced planners or calculated by statistical models. The demand planning process consists of several dependent, planning steps with different time horizons and aggregation levels.

The demands are the basis for midterm planning in SNP, where resource constraints and material availabilities in each time bucket are taken into consideration. Here, the planning horizon is 12 months with monthly time buckets.

PP/DS supports lot-sizing and the scheduling of production amounts per time bucket coming from SNP. Due to specific requirements in the process industries, an optimization extension with ILOG Cartridge Technology has been implemented to improve the lot-sizing of PP/DS in the first three months of the planning horizon.

Demand fulfilment leads to promised delivery quantities that can be shipped to dedicated key accounts or regions. Global ATP supports checking allocated product quantities and the availability of materials, using search procedures or multilevel checks (see Chap. 9).

Procurement planning is done in the transactional system SAP R/3 based on the material requirements from planned orders.

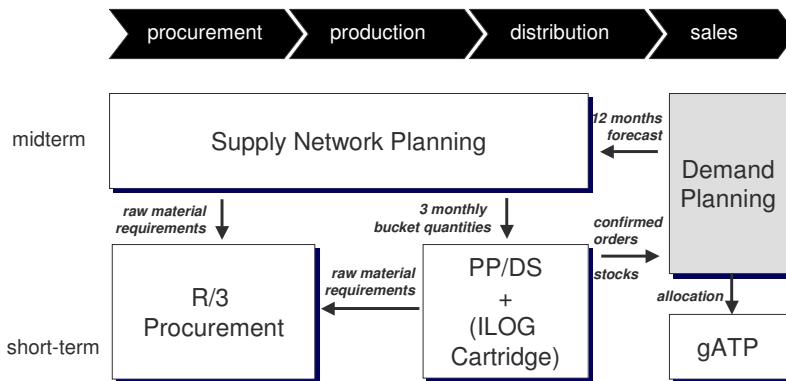


Fig. 22.2. The architecture of the planning system

22.3 Model Building with SAP APO Demand Planning

Modeling with SAP APO Demand Planning takes place in two different design areas (see Fig. 22.3):

- Historical data are provided by SAP's Data Warehouse.
- The planning environment is based on the *liveCache* (see Sect. 18.3).

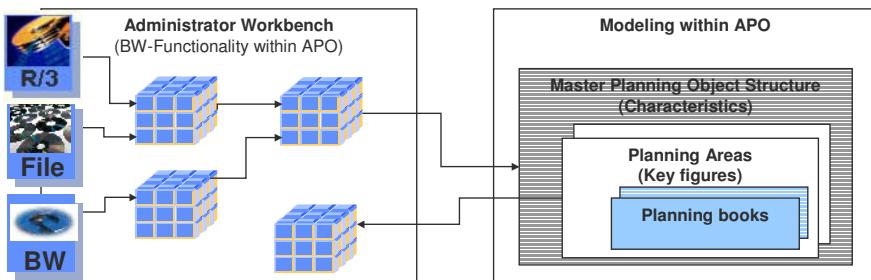


Fig. 22.3. The two design areas of the Demand Planning

The first step applies a typical data warehouse structure and uses the functionality of SAP's Business Information Warehouse (SAP BW). Thus, it supports the Extraction-Transformation-Loading (ETL) process to get data from different data sources, and cleanse and enrich them for further use. The data sources can be different transactional systems, e.g. R/3 systems of various affiliates, other data warehouses, data bases or even flat files. The structure of a data warehouse allows reporting large data sets along different dimensions on different aggregation levels (month or year, stock keeping unit or product family, customer or region, etc.). Also, selection of data and summations along hierarchies are performed very efficiently (see e.g. Berry and Linoff (1997) and Reuter (2004)).

InfoCubes, Characteristics and Key Figures

The data model of a data warehouse is often represented by a “star scheme” as shown in Fig. 22.4, where the data values are stored in one “fact table” containing all data sets.

Each data set is identified by a unique key, that is a combination of characteristics, and consists of several key figures. The characteristics (see Fig. 22.4: customer C, product P, time T) are grouped in dimensions and might have several hierarchies (e.g. day, week, month, year; or see Fig. 22.1). The characteristic combinations span a multidimensional data space – the so-called InfoCube (see Sect. 7.1).

The key figures are the quantities assigned to a dedicated characteristic combination such as actual sales, open orders, sales budget or forecast quantity (see Fig. 22.4). Key figures can be retrieved for every selection of one or more characteristics. For example, actual sales values can be displayed for one or more products for each region or each customer.

The interactive way to analyze the data is called online analytical processing (OLAP); it has the following functionalities:

- slice and dice: get a subset of data, e.g. show the sales history of the region “Germany,”

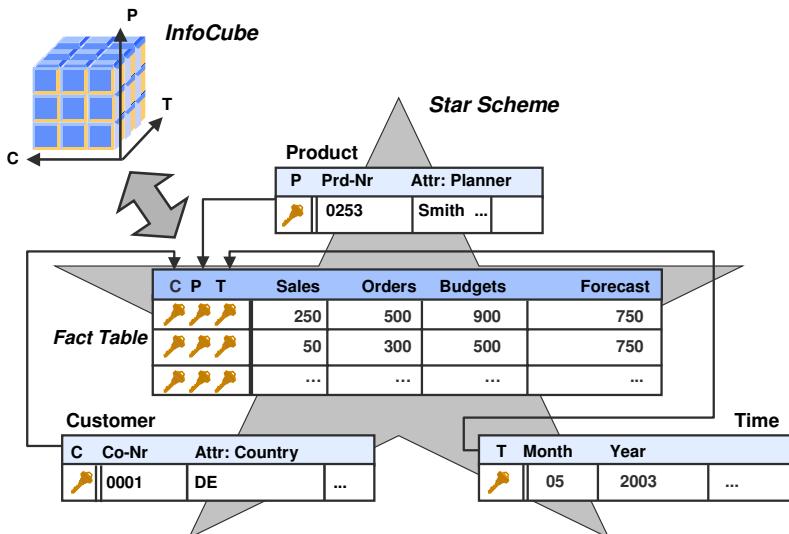


Fig. 22.4. InfoCube and star scheme

- drill down: get more detailed information from one hierarchy or dimension, e.g. show the sales history of all customers in the selected region and
- rotate: change the granularity of two characteristics, e.g. from showing the sales history of all products of the region “Germany” to showing to which region product “X” is sold.

Due to these functionalities, this structure is very appropriate for flexible Demand Planning purposes.

While a data warehouse supports reporting – retrieving data – the input of new planning data is done in the second design step. SAP provides several planning environments for Demand Planning outside the transactional system:

- quantity-based plans: SAP APO Demand Planning and
- cost-, revenue- or quantity-based plans: SAP BW-BPS Sales Planning.

In this case study, the focus is on SAP APO Demand Planning.

The process starts based on the structures of historical data that describe the sales history. Usually, Demand Planning is done on quantities of several materials sold to several customers in dedicated periods. Long-term planning is performed at a more aggregate level, while short-term planning is performed at a more detailed level. First, the granularity of planning is defined by choosing aggregation levels of the characteristics. For example, Demand Planning of stock keeping units for each customer on a weekly basis is on a very detailed level and increases the number of data sets very quickly in

contrast to a planning granularity of product groups, regions and months. Because of the vital impact on performance and workload the trade-off between detailed and aggregated data has to be investigated very carefully.

Master Planning Object Structure

Each characteristic to be planned on (e.g. region, customer, product group, product, business unit), needs to be included into the set of characteristics the so-called Master Planning Object Structure. Others, which are only for reporting or selection purposes, can be put into hierarchies or (navigation) attributes. In Fig.22.4 the “planner” is an attribute of the characteristic “product”. If the granularity-level of the characteristics is defined, a first estimate of the number of characteristics combinations can be done.

Planning Areas

As mentioned above, the key figures contain the quantities of a characteristic combination. The structure of SAP APO’s Demand Planning allows for the creation of several “bundles” of key figures – the so-called Planning Areas – that can be assigned to one Master Planning Object Structure. Reasons for multiple Planning Areas can be different time granularities in different planning scenarios or avoiding data locking in simultaneous planning activities on the same characteristic combinations. SAP APO Demand Planning distinguishes three types of key figures:

- persistent key figures: actual data with the origin of the data warehouse that should not be changed,
- planning key figures: open key figures to enter or manipulate data by the planner (data is saved to the liveCache) and
- temporarily created key figures: key figures that are used for intermediate calculations and are not saved.

Each key figure has its own aggregation rule. Thus, the selection of more than one characteristic combination displays the aggregation of the quantities of the key figure. For example, if the aggregation rule of actual sales is “adding up,” then all the total sales quantities of all customers and all products in the selected region are displayed. From the planning point of view, the disaggregation rules are more important. If the planner enters a quantity for a key figure on an aggregate level, e.g. a region containing several customers, the disaggregation rule describes how the quantity is distributed to the different customers of the region. The most important rules are

- equally distributed: The quantity is distributed equally by the number of selected characteristic combinations;
- pro rata: The distribution ratio is calculated by the portion of the key figure for a certain characteristic combination relative to the total of the key figure; and

- based on a key figure: The distribution ratio for the modified key figure is calculated by the portion of another reference key figure for a certain characteristic combination relative to the total of the reference key figure.

Fig 22.5 shows the aggregation of the forecast quantities (FCST) of three customers with 200, 400 and 600 units in Region 1 added up to 1,200 units. The equally based disaggregation of 900 units of forecast planned on the region level leads to 300 units for each customer. The pro rata rule uses the current portions, e.g. $200/1,200 = 1/6$, to assign $1/6 \cdot 900 = 150$ units to Customer 1. In case of the disaggregation on another key figure, e.g. sales, the portions are taken from the sales quantities to Customer 1, divided by the total sales of Region 1 ($400/1,000 = 2/5$) and lead to a forecast assignment of $2/5 \cdot 900 = 360$ units to Customer 1. The example shows that the choice of the disaggregation rule has a large impact on the planning results.

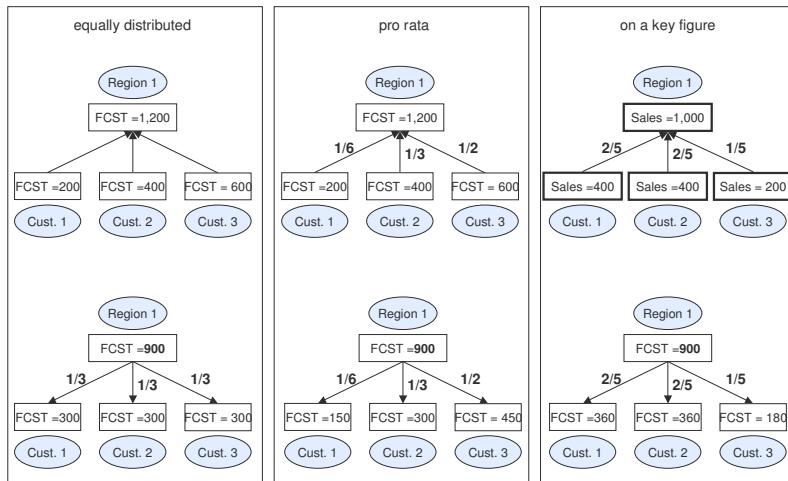


Fig. 22.5. Disaggregation rules

Planning Books and Data Views

The access to the key figures is managed by planning books and their data views. Here data like actual sales or forecasts can be shown, checked and changed. They are the “planning front-end” to the planner. The planning book is a container providing the planning functionality, e.g. forecasting, the subset of characteristics of the Master Planning Object Structure shown to the planner for selection, a subset of key figures of the planning area and the temporarily created key figures. Based on these objects, a data view additionally defines the time horizons of past and future and the layout and format

of the key figures (sequence of key figures, background color, number of decimals, etc.). Additional functionalities help the planner to create the forecast or to understand the data. The functionality (e. g. locking or highlighting data cells, calculating key figures, generating alerts) can be implemented with a programming language – the so-called macro-builder. Publishing a planning book in a web- based scenario supports collaborative demand planning (see Chap. 14).

Forecast Methods

SAP APO Demand Planning provides univariate and multivariate forecast methods and the combination of the two. They can be configured in a forecast profile. Here the type of model (e. g. constant, trend; see also Chap. 7) and the forecast methods (e. g. moving average, exponential smoothing, Winters' method, multiple-linear-regression) are also specified as the parameters of the method, the basis of historical data and the number of periods to be estimated (see Chaps. 7 and 26). As further topics, promotions planning and life-cycle models are provided to estimate singular or non-stationary effects of a time series.

The created demand forecast has to be released to production planning either to Supply Network Planning for finite capacity planning – as in our case – or directly to SAP R/3 for infinite capacity planning.

22.4 The Demand Planning Process of the Styrene Plastics Division

The goals of the Demand Planning in the Styrene Plastics Division are the

- influence of the market behavior,
- using of sales departments' knowledge of customers' ordering behaviour,
- coordination of decentralized regional sales departments,
- ability to react quickly to short-term market fluctuations and
- creation of a midrange demand plan in order to calculate a production and procurement plan.

To achieve these goals, input from marketing department, sales department and logistics department is needed. The marketing department adjusts market quantities and influences prices directly or indirectly. The sales department is responsible for customer relationships and the “recording” of demands and placing the orders. Sales and logistics negotiate sales budgets to support “management by objectives.” The logistics department is responsible for the consolidation of the sales and marketing plans, the short-term adjustments and the creation of one single demand plan from the various plans.

The following plans are created:

- sales budget,
- rolling business forecast,
- marketing plan,
- sales plan and
- short-term adjustments.

Referring to the structure of Chap. 7 (p. 140) only the sales plan is supported by a statistical forecasting method. The sales budget is the result of a negotiation process; the marketing plan is created in the marketing department and entered manually; the rolling business forecast is also entered manually and based on the experience of the planner. Starting from statistical forecasting, the sales plan is revised and can be modified (revised judgmental forecast). From the different plans, one consensus-based demand plan is calculated.

The monthly planning cycle consists of the following steps:

1. The sales plan is statistically forecasted and revised by the regional sales departments. The plan contains demands for each stock keeping unit and customer and determines the portions of each characteristic combination.
2. Marketing decides on the total quantities that should be sold to the market, neither to decrease the price by shipping too much nor to loose market share by selling too little.
3. The total of the sales plan quantities has to be adjusted to reach the total marketing quantity without changing the portions of the sales plan.
4. While the first three steps are related to months one through three , the aggregate rolling business forecast is made up for in months four through 12.
5. To release demands for 12 months to SNP (see Fig. 22.2), the first three months from the adjusted sales plan (step three) are combined with months four through 12 of the rolling business forecast.
6. Demands from the released demand plan are assigned to delivery locations for an interval of 12 months.
7. Short-term adjustments can be done prior to the release to SNP.

An overview of the existing demand plans is shown in Fig. 22.2. In the following chart, the different plans are described in more detail.

Sales Budget

The sales budget is not a subject of the monthly planning cycle. The aim of the sales budgeting process is to negotiate the expected yearly quantities sold by a region or to a key account. The definition of the sales budget of the following calendar year takes place once a year in October and remains valid for 12 months. The result agreed upon is yearly quantities at the product group

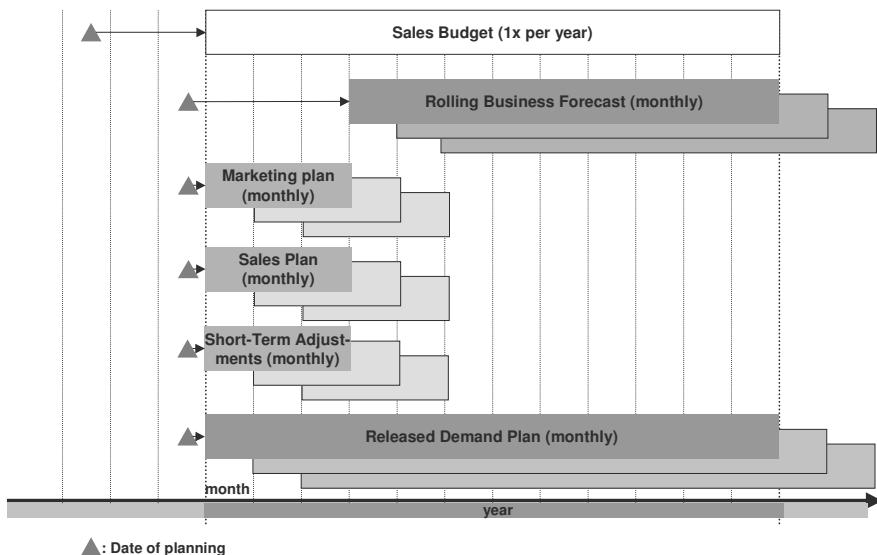


Fig. 22.6. Overview of the different demand plans

and region level. The negotiating partners are the logistics department of the business unit and the responsible regional managers of the sales department. The plan helps to keep track of the achievement of sales objectives and thus to shift priorities. It is also used to guide the promised product quantities – allocations – along the sales budgets. A special feature is an automated disaggregation procedure on the time structure that distributes the yearly quantity of the sales budget on monthly quantities. The distribution is based on the sales history of the previous year. This is done by the disaggregation rule of the sales budget based on the key figure “sales history.” Thus, seasonal effects are taken into account.

Rolling Business Forecast

In accordance to the midterm plans of the business unit, the logistics department enters the aggregate business forecast of the business unit manually. The definition or update of the business forecast for the next 12 months is performed each month in the so-called rolling business forecast. The result is a highly aggregated plan for the complete business unit on a monthly basis. The plan is needed to support 12 months production planning with SAP APO SNP. As the figures are highly aggregated on the product and customer structures, they have to be disaggregated for planning purposes. The sales budget is used as a reference key figure for the disaggregation.

Marketing Plan

The aim of the marketing plan is to influence prices in the oligopoly market by short-term adjustments of the expected sales quantities for the different product families. The plan is made up for three months on a monthly basis. Because of the global impact of the quantities sold to the market, a differentiation on regions or key accounts is not useful. The marketing department is responsible for the creation of a marketing plan. The quantities are used later in the adjustment process (see Short-term adjustments and plan consolidation).

Sales Plan

The sales plan is the central object of forecast collection and input of the sales department. It contains the forecast quantities for the next three months on a monthly basis for each stock keeping unit and each customer. The planning complexity is reduced by an ABC-classification of customers and stock keeping units; thus, only the most important items and customers are planned manually, all others are forecasted automatically. The forecast process can be classified as a revised judgmental forecast (see Sect. 7.3). The sales planner is supported by the historical sales quantities of the last two years as well as by forecasts calculated with a simple moving average method, whereby more than six months of historical sales quantities are averaged. Because of the absence of more complex regular patterns, the calculated forecast gives an impression of the expected quantities. The variations in the historical time series, e. g. reduced sales caused by vacation, are easily identified by the planner in a year-to-year comparison for each month and do not justify a more complex statistical model. Due to handling constraints, minimum delivery quantities have been taken into account for each combination of stock keeping unit and customer. Minimum quantities are automatically identified. If customer orders are already placed in a period, the forecasted quantities have to be at least as high as the ordered quantities, or else they are also identified.

Short-Term Adjustments and Plan Consolidation

Before the release of the demands to the finite capacity planning, various plans are consolidated and short-term adjustments are made.

One of the most important adjustments is matching the marketing plan with the sales plan. Here, the structure of the forecasted demand portions of the sales plan is combined with the quantities of the marketing plan. Consequently, the total quantity shipped to the market should not lead to an unwanted behavior, e. g. a decreasing price or loss of market share. Thus, quantities might change, but not portions. The calculation uses the disaggregation rule “based on a key figure,” where the reference key figure is the sales plan.

Another important adjustment is the adding of the delivery location. As shown above the forecast is made on stock keeping units and customers, whereas the sourcing location is missing. In the case of multiple production sites, customer demands could be produced at and shipped from different sites. Subject to transportation and production costs, the choice of the assigned location has an impact on the cost structure. Of course, transfers between production sites are possible, but they cause transportation costs. Thus, the assignment of demands depends on the product (single production site or multiple production sites) and customer site. If the planning region of the customer could be supplied from multiple locations, the ZIP-Code of the customer is used for assignment to the location. Those rules are also used for a first allocation within global ATP checks.

In between a planning cycle, demand figures may change. An instant reaction is supported by a direct communication process with the logistics department and a separate key figure for short-term adjustments to allow for the monitoring of changes.

After the adding of the location to the demand data, the release to Supply Network Planning is performed and production planning can be started.

22.5 Concluding Remarks

The above demand planning application has been used since the beginning of 2002. The planning process and the functionalities are highly accepted. Currently, there are about 70 users in different European countries working with the system.

The planners not only use the Demand Planning application, but also profit from the reporting functionality of the Business Information Warehouse. SAP's Business Explorer Analyzer (BEx), a Microsoft Excel front-end on the Business Information Warehouse, helps in analyzing forecast accuracy, printing formatted reports, checking master data and supporting online analytical processing.

Forecast accuracy is measured in different ways:

- Sales Budget fulfilment: The cumulated sales and the planned cumulated sales budget are compared on a monthly basis for each product group and region. If the sales go beyond the budget, the sales data are highlighted.
- Forecast development: Based on the historical sales data for each product group the mean value and the standard deviation sigma are calculated. Then the forecast of a product group is plotted in a graph with lines for the mean value and the mean value plus or minus one, two and three times sigma. If there is a demand forecast exceeding the one-sigma, two-sigma or three-sigma borders, a traffic-light function highlights the forecast and creates alerts of different severity.
- Sales – sales plan – marketing plan comparison: Because of the adjustments before the final release, the original forecast quantities can differ

from the adjusted quantities. Thus, the actual sales are compared with the original sales plan and with the quantities after the marketing adjustments.

The following benefits have helped to stabilize the revenues of the Styrene Plastics Division in spite of the weak economic situation in Europe:

- reduction of the planning cycle time, which led to a higher availability of the planners and to higher responsiveness,
- reduction of communication efforts by using a common data base with dedicated responsibilities for the different plans,
- increased forecast accuracy, through revised judgmental forecast and plausibility checks and
- better control of the decentralized sales regions by tracking the sales budget.

The lessons learned from the project are the following points:

- There is a vital impact of master data quality on the project efforts. The two main drivers are methods to identify master data inconsistencies, e. g. reports, and activities to repair them.
- The simplicity of building planning processes encourages a project team to create complex structures bearing the risk of increasing planning cycle times.

References

- Berry, M. J. A.; Linoff, G. (1997) *Data mining techniques: For marketing, sales, and customer support*, New York et al.
Reuter, B. (2004) *Datenanalyse und szenarioorientierte Kapazitätsplanung von Lagern*, Berlin

23 Semiconductor Manufacturing

Lorenz Schneeweiss¹ and Ulrich Wetterauer²

¹ Nestlé GLOBE Center Europe GmbH, Rohnestrassse 4, 60523 Frankfurt/Main, Germany

² j & m Management Consulting AG, Kaiserringforum, Willy-Brandt-Platz 5, 68161 Mannheim, Germany

The semiconductor industry has one of the longest and most complex manufacturing processes and supply chains. Yet the fast changing market and the continuous price decrease in the semiconductor industry imposes increasing pressure on the manufacturers to decrease production lead-time, while keeping costs low. As a result superior planning capabilities across the corporate-wide supply chain become more and more vital to meet the challenges of this industry.

This chapter provides some insight into the implementation of the *i2 Supply Chain Planner* (SCP) from i2 Technologies (see Chap. 18) for a large semiconductor manufacturer. The goal of this APS project was to unify the planning processes throughout the organization and to achieve a significant reduction of the planning cycle time.

After introducing the main characteristics of the semiconductor supply chain and the APS model typology a case study is described which investigates the possibility of using the software product *Factory Planner* (FP) from i2 Technologies instead of i2 SCP for the detailed short-term planning.

23.1 Case Description

23.1.1 Topology of the Semiconductor Supply Chain

The structure of the semiconductor supply chain is influenced by several characteristics. First, the ratio of value to volume of the product is very high, therefore transportation even over long distances e.g. by plane is not an issue. Second, the equipment necessary for production is expensive and difficult to transport and install. Furthermore, the production process can be split over several locations, e.g. the facilities wafer fabrication and wafer-test/sawing in Europe and the facilities die bonding, wire bonding, moulding and chip-test in Asia. The production sites for the different stages of the production process can therefore be placed all over the world to achieve goals like reduction of costs for real estate, equipment or wages.

Another characteristic is that the customers are typically resellers or large industry corporations with considerable influence on the manufacturers. It is for example not uncommon that a customer demands the fabrication of a logic component in a special production site or even on a special machine.

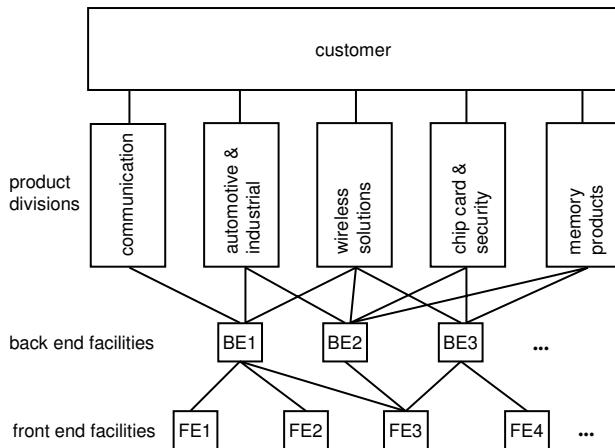


Fig. 23.1. The semiconductor industry supply chain

The semiconductor industry is at the beginning of a complex supply chain leading to a large demand variability with a corresponding uncertainty of demand forecasts. Usually a manufacturer offers a variety of products, mainly in the areas communication, automotive and industrial, wireless solutions, chip card & security and memory products. The organization of the manufacturer used for the case study is divided into product divisions according to these areas (see Fig. 23.1).

Table 23.1 shows the classification of the semiconductor industry according to the supply chain typology (see Chap. 3).

23.1.2 Production Process

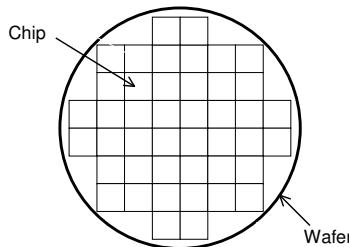
The production process of semiconductors consists of two stages each divided into two phases. The two stages are called *front end* and *back end* and are separated by a die bank. The front end consists of the *wafer production* and the *wafer test* phases whereas the back end is represented by the two phases *assembly* and *chip test* (see Fig. 23.3).

- **Wafer Production:** Starting point of every chip production is a round silicon disc. The so-called wafer does not represent a functional element of the future chip, but serves as the basis and carrier for the functional elements to be implemented (see Fig. 23.2). Layers of conductive and isolating material are spread on the wafer in several repetitive processes. After each layer a photo chemical layer is spread on the top layer and exposed with a mask carrying the circuit information of each layer. This film is then developed to etch the circuits in an acid bath. Afterwards the photo chemical layer are removed and the next conductive or isolating layer can follow.

Table 23.1. Supply chain typology for the semiconductor industry

Functional attributes	
Attributes	Contents
products procured	highly specific
sourcing type	single or double sourcing
organization of the production process	job shop and flow shop
repetition of operations	batch production
distribution structure	one and two stages
pattern of delivery	dynamic
deployment of transportation means	unlimited
availability of future demand	forecasted
products' life cycle	several months
degree of customization	highly specific
portion of service operations	low

Structural attributes	
Attributes	Contents
network structure	mixture
degree of globalization	global
location of decoupling point(s)	deliver-to-order and make-to-order
legal position	inter- and intra-organizational
direction of coordination	mixture
type of information exchanged	orders

**Fig. 23.2.** Structure of a silicon wafer

This process cycles six to twelve times per wafer depending on the complexity of the chip architecture. A number of several hundred small multi-layer circuits and transistors can be developed on a single wafer. All wafers are processed in lots of typically 50 pieces. The maximum lot-size is defined by technical restrictions of the racks used to carry the material. The lot-size can also be less than 50 pieces if necessary to reduce the number of intermediate parts in the inventory, but it is tried to maintain the maximum lot-size over the whole process.

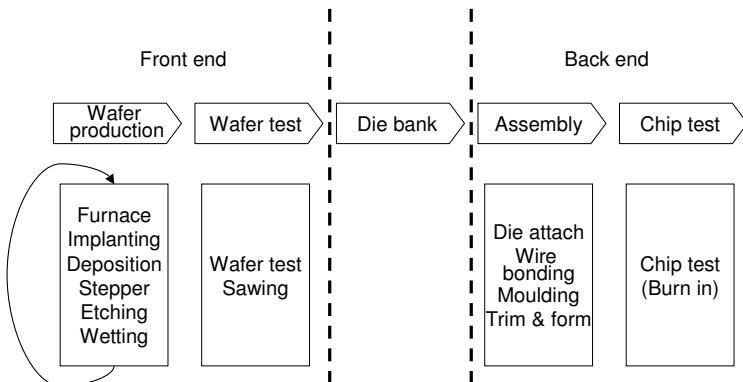


Fig. 23.3. The semiconductor production process

The wafers run through these repetitive process steps in four to six weeks. The constraint is a machine called "stepper", responsible for the positioning of the wafers and exposing the layers according to the mask. This process has to be conducted with great accuracy to avoid waste of material and capacity.

Subsequently the intermediate product can be used to produce different finished goods. The production lead-time of this process is again four to six weeks. At the end of this process the wafers are ready to be shipped to the wafer test.

- **Wafer Test:** The production process of the wafer test starts with the output of the preceding production step. This process is constrained by the availability of testing capacity. The testing capacity is usually a constraint to the complete semiconductor supply chain. After testing, the wafers are sawed to separate the single chips. The yield factor of the testing procedure depends on the quality of the production process and on the number of chips per wafer. Consecutively, the chips are shipped to the customer or stored in a finished goods stock.
- **Back End** The back end operates in a make-to-order mode. In the assembly the chips are supplied from the central ware house and are connected to the platforms, bonded and sealed in plastic. The production process is finished with the final quality control, where the finished element is aged artificially with high temperatures. The back end processes are not part of this case study.

23.1.3 The As-Is Situation

The rapid changing markets in the volatile semiconductor industry with its short development cycles complicates the forecasting process and typically leads to low forecasting reliability and obsolete, unsaleable inventory.

Planning was characterised by diverse, intransparent processes caused by the varying requirements of the different product divisions (see Fig. 23.1). The product families differ in production processes (make-to-stock, make-to-order) as well as in lead times and cycle times, resulting in significantly varying planning horizons (from days to months).

23.2 Objectives of Project

The goal of the APS implementation project was

- to decrease forecasting lead time and increase forecasting reliability,
- to gain visibility of product availability and potential production problems throughout the entire supply chain,
- to unify the planning processes and to reduce the planning cycle time,
- to decrease response times to customers and
- to reduce inventory.

This should be achieved by using an integrated planning system with a top-down approach for the product structure (from aggregated to detailed) and coordinated shared facility capacity through aggregated capacity groups (see below for details). The planning processes which had to be considered in this planning system were:

- demand/forecast planning,
- long-term production and distribution planning (corporate planning),
- mid-term master planning (divisional planning) and
- short-term production scheduling (facility planning).

23.2.1 The To-Be Vision

The solution is based on a three-tier architecture: corporate and demand planning, divisional planning, and short-term production scheduling. A graphical representation of the to-be planning tasks is given in Fig. 23.4.

- **Demand/forecast planning** is used to generate accurate forecast values and as a tool for marketing to manage volume and revenue planning. The results are used as input for corporate planning and master planning.
- **Corporate planning** tasks (see Fig. 23.4) comprise financial/budget planning, product aggregation to gather future demand and allocation of capacity to business units (divisions) with the focus on key equipment. This long-term planning is done once or twice a year with a time horizon of three years. The results of corporate planning are based on the interaction with the demand planning and master planning and include an

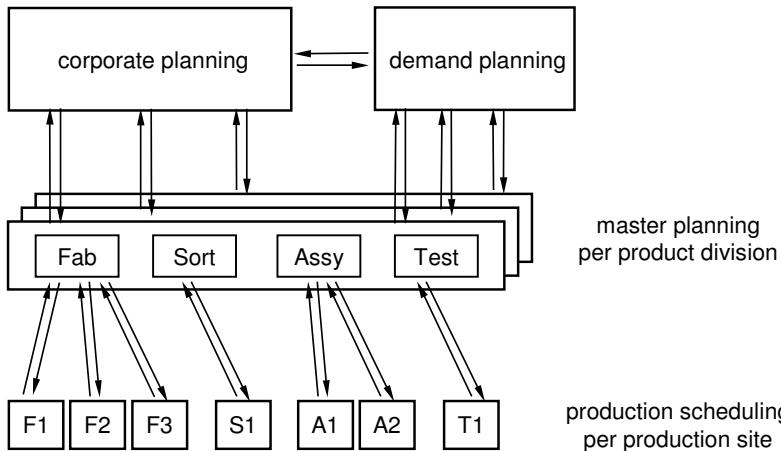


Fig. 23.4. Model architecture in the semiconductor APS implementation project

investment plan for additional capacity, capacity checked forecast (forwarded to demand planning) and capacity allocations per division, product group and production site (forwarded to master planning).

The interface between corporate planning and divisional planning (or master planning) are the aggregated capacity groups where the capacity per division and product group is modelled as a single resource.

- **Divisional planning** is done to quote and confirm orders, to determine the demand for the different production facilities (forwarded to production scheduling), to receive the feedback of the facilities and reconcile the master plan to the changes and finally to conduct early warnings in case of planning problems (late or short orders) within the planning horizon (12 months in case of the semiconductor company considered in this case study).
- **Short-term production scheduling** (or facility planning) is used to generate a dispatch list for the line control and to conduct a detailed scheduling on key resources. This information is based on the demand taken from the master plan and work in process information from the shop floor control system. If orders can only be planned late or short an early warning message should be generated. The planned quantities and dates are reported back to the divisional planning.

In the APS implementation project the three parts of the planning system corporate planning, divisional planning and production scheduling (facility planning) were carried out with the software tool Supply Chain Planner while demand planning was done using the Demand Planner from i2 Technologies, respectively. The integration with the Oracle database was done using the

software i2 RhythmLink. The actual scheduling process is supported by the shop floor control system Workstream by Consilium, a subsidiary of Applied Materials Inc. (see www.appliedmaterials.com).

23.3 Model Building with the i2 Factory Planner

As mentioned before the software product Supply Chain Planner from i2 Technologies was used for the detailed production planning in the actual APS implementation project. Although the to-be processes could be represented by the software, the efforts for modelling and customization were considerable as the SCP module, the master planning solution of i2, was not designed for the complexity of detailed scheduling tasks.

This case study investigates the possibility of using the software product Factory Planner instead of for the production planning in the semiconductor industry. For this purpose the two front-end production sites of the supply chain, the wafer fabrication and the wafer test, have been simulated in two separate models using the Factory Planner. Furthermore, a two-way communication between the two models was established (a detailed description of the two models will be given later).

- the wafer test site calculates material demand (requests) for the wafer fabrication based on orders, work in process and basic data like yield, cycle time and lot-sizes.
- the wafer fabrication calculates the material flow to the wafer test site based on these requests as well as work in process and basic data.

The goal of this case study was to understand the consequences of using Factory Planner for the short-term production planning in the semiconductor industries instead of the Supply Chain Planner and to investigate whether all relevant details of the production process could be included in the model.

23.3.1 The Modelling Concept of i2 Factory Planner

This section describes the procedure of creating a model of a production process using i2 RHYTHM Factory Planner (FP). First the basic data structure will be explained in order to get an impression of the modelling philosophy. In the following the most important elements of a model will be introduced, the basic model. Any additional modelling is case specific and will be explained subsequently in detail. Please note that RHYTHM FP, on which this case study is based, is not the latest version of the Factory Planner by i2. The concepts described in the following (data structure, modelling, ...) might differ from the concepts of later versions. For further details on the latest version of the i2 Technology SCM suite see Chap. 18).

Table 23.2. Definition of an example spec file

std_spec_file: (here: demand order record)	
demand_order_data	name of the data file to be read or written
mode: read	indicates the mode of the data file
Demand_Order_Record	name of the record, defined by the following fields:
order	unique identifier of the order
sales_due_date	desired due date
part_number	the end product to requested by this order
part_quantity	the ordered quantity of the end product
priority	priority
category	category
customer	customer

The Data Structure The whole data model of FP is based on flat files (ASCII) of two different types, spec files and data files. Spec files are used to structure and define the content of the data files. They provide the framework of the model whereas the data files build it.

Spec Files The *standard spec file*, provided with each release of FP, contains the standard definition for all possible ASCII data files that may be input to and output from FP. A *spec file* contains a subset of the definitions specified in the standard spec file. Each data file in the customer data directory must be defined either in the standard spec file or in the implementation specific spec file. The type and class of each field in the definition must match the corresponding field in the data file. Table 23.2 shows the field definitions of an example spec file.

Data Files The data files contain records matching the structure defined in the spec files. A record is one line in the data file. In a record the field information is separated by a delimiter. The following text is an excerpt from a larger data file that corresponds to the spec file definition of Table 23.2:

```
Demand_Order_Data
Order 1,03APR1999120000,Part A,20,,,Customer 1
Order 2,03APR1999120000,Part B,70,,,Customer 1
Order 3,03APR1999120000,Part B,40,,,Customer 3
```

In this example the fields *priority* and *category* are not populated and will be filled with default values.

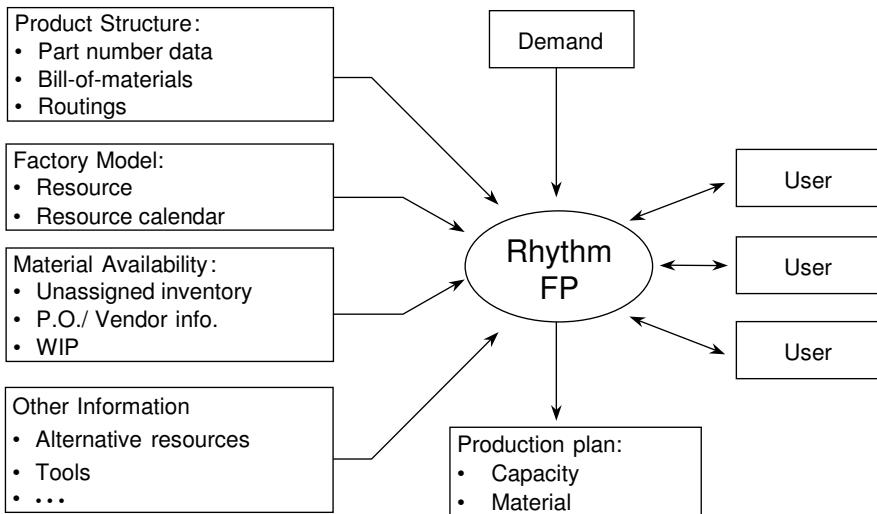


Fig. 23.5. Modelling data classification of i2 RHYTHM Factory Planner

Table 23.3. Description of basic FP model data

part_number_data	definition of all parts and products
bill_of_materials.data	the bill-of-materials for all products
routing_data	description of the operation plans
resource_data	defines all resources and their properties
resource_calender.data	availability calendar of the defined resources
unassigned_inventory_data	contains the unassigned inventory at any time
purchase_orders	planned and scheduled material receipts
supplier_part_data	vendor information
wip-data	work in process quantities

Building a New Data Set The I/O data model of the planning process with FP is displayed in Fig. 23.5 (see also Table 23.3 for explanation of input data). On the input side the basic model and the additional modelling features are shown. The basic model consists of data files describing the product structure, the factory model and the material availability. The dynamic part of every basic model is represented by the demand. FP in interaction with the users creates a material and capacity feasible plan as output (refer to Fig. 23.5).

There is a minimum of data files necessary to build a basic model in FP which can be extended using other modelling features (see Fig. 23.6 and Table 23.3).

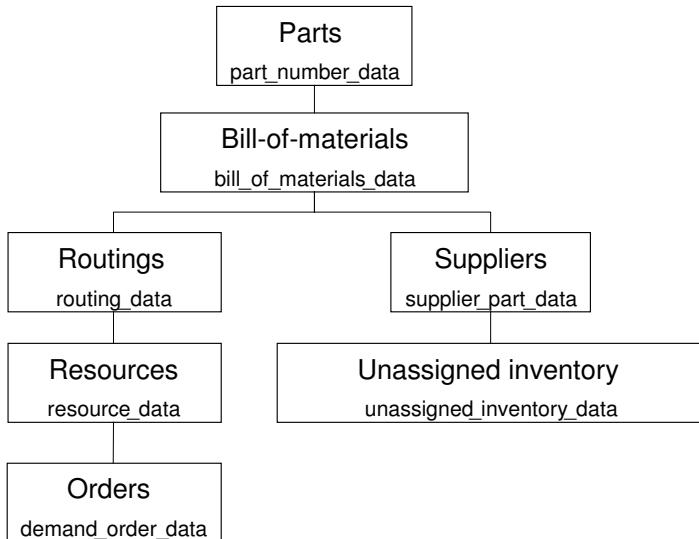


Fig. 23.6. The basic model used by i2 RHYTHM Factory Planner

23.3.2 Modelling

Wafer Production The wafer production model considers three products (wafer A, wafer B, wafer C, see Fig. 23.7). A relaxation will be used in the model that aggregates for each product all the process cycles mentioned before to one operation in order to keep the model simple. The machine is loaded from 2 hours (6 cycles) to 4 hours (12 cycles) per lot, depending on the production process of the wafer. This modelling is sufficient to create a rough capacity plan. The detailed scheduling of the process cycles will be done by a shop floor control system. The lead-time of the non-constraining processes can be modelled as a fixed set-up time, for example.

As regards to modelling the following requirements have been identified:

- lot-sizing
- safety stock on intermediate parts

After a common understanding of the requirements the aggregation of the modelling has to be defined. As a principal rule it can be said that only (potential) constraints have to be planned for. Therefore the aggregation level of the model concentrates on the two constraining operations.

Like in all implementations the modelling has to start with the basic model described above. So, all parts, bills of material, resources, routings, suppliers, inventory and demand have to be specified according to the model and spec file definitions. Then the model can be extended with additional features and functionalities.

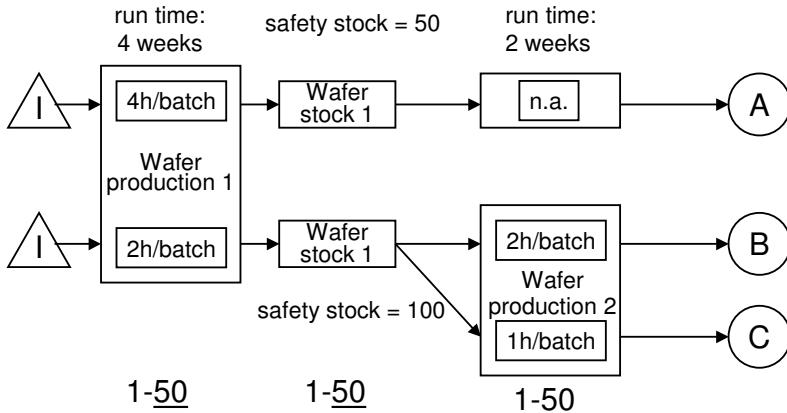


Fig. 23.7. The production process model of the phase *wafer production*

Lots can be defined using the following two data files populated accordingly. It would be sufficient to populate only the batch_size.data file if the lot-size definition was unique for one resource. But in case of different operations on the same resource it is necessary to identify a certain batch_type with every different operation. The batch_size_data has the following structure:

resource	type	max.	min	ideal	run_time	run_time uom
wafer production 1	Typ A	50	0	50	4	HOURS
wafer production 1	Typ BC	50	0	50	2	HOURS
wafer production 2	Typ B	50	0	50	2	HOURS
wafer production 2	Typ C	50	0	50	1	HOURS
wafer production A2		50	0	50	2	HOURS

The batch_type_data file is structured as follows:

routing	operation	primary
routing A1	preprocessing A	Typ A
routing B1	preprocessing B	Typ BC
routing B2	postprocessing B	Typ B
routing C2	postprocessing C	Typ C

The wafer production A2 represents a dummy resource with a machine load time 0 but a wait time greater than 0. Therefore the resource does not cause capacity problems in the capacity planning.

In contrast to time based safety stock a simple quantity based safety stock on intermediate parts is not supported explicitly by FP.

Wafer Test The planning model is represented in Fig. 23.8. In this model the rule can be extended that not only the potential constraints have to be modelled but also every process the planner needs to have detailed information about in order to be able to create reports or to take a decision. In this

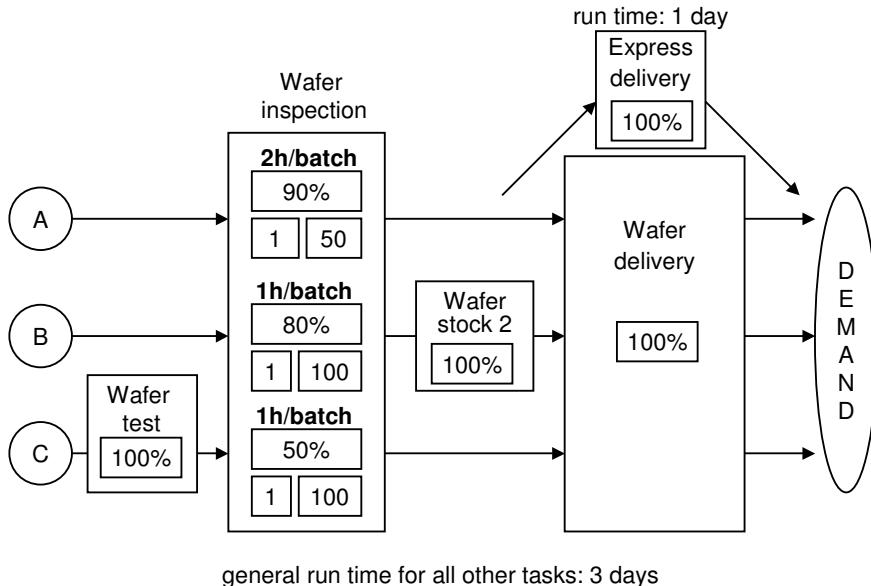


Fig. 23.8. The production process of the phase *wafer test*

case FP is to take a decision about the delivery service used thought it is not a constraint. A specialty of the semiconductor production are the low yield factors at the beginning of the use of a new production process (about 5%) which increase in the course of time to up to 95% with the learning curve effect. Consequently the following requirements have to be considered:

- ramp up/yield,
- lot-sizing and
- alternative resources.

The modelling starts again with the basic model.

In order to model ramp up and yield, time dependent bills of material are used to refer to different routings. The first record of the bill_of_materials_data is valid until 6/1/99 because then a new record is added (A) describing the same relationship and the former record is deleted (D) from the list of records. However, the new record references a different routing with a different base yield factor. Yield can therefore be modelled in a time dependent step function.

The following is an excerpt of the bill_of_materials_data:

produced_part	routing	consumed_part	ecn_code	ecn_date
produced_part_A	routing 1	consumed_part_A		
produced_part_A	routing 2	consumed_part_A	A	6/1/99
produced_part_A	routing 1	consumed_part_A	D	6/1/99

The routing_data file is structured as follows:

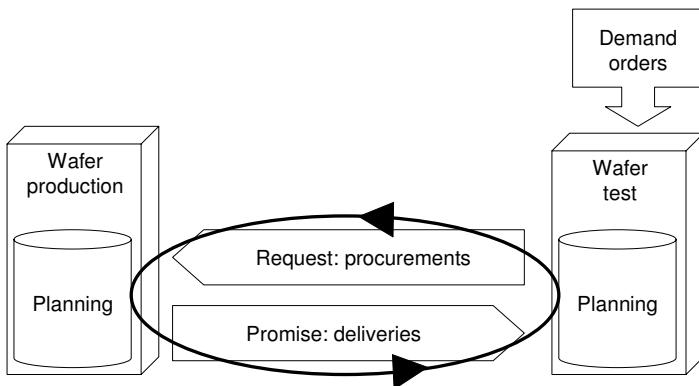


Fig. 23.9. Model communication

```

routing      operation      resource      run_time      uom      yield
routing 1    OP1           Res1          2            HOURS    0.8
routing 2    OP1           Res1          2            HOURS    0.9
  
```

Lot creation is modelled in the same principal way as presented before. Alternate resources can be specified in the `operation_resource_data`. By assigning a different resource to the same operation in a routing an alternate resource can be created. The time factors of the alternate resource reference the time factors of the primary resource.

The `operation_resources_data` file is structured as follows:

```

routing      operation      resource      runtime      cooldown_time
delivery    chip_delivery  fast delivery  1           0.3333
  
```

The decision to use an alternate resource is often linked with higher costs. Financial figures are not considered by the decision rules in FP though they can be displayed. Therefore the alternate resource can be chosen either automatically or manually if higher costs are involved in the planning decision.

23.3.3 Model Communication

A communication link between the two models was setup, leading to a collaborative planning process. The two models exchange data and enable the planners responsible for wafer production and wafer test to react to changes in the plans, based on data provided by the other model. One loop is shown in Fig. 23.9. This iterative planning process can be run as often as required until all severe planning problems are resolved.

The wafer test model creates a plan for the demand orders that were accepted. It therefore creates procurement recommendations anticipating the capacity and the behaviour of the supplier (wafer production model). The

wafer production model now plans for the demand imposed on it by the consumer (wafer test model) and generates deliveries for the requested material. The deliveries are imported by the wafer test model to generate a new plan based on the new supply information. Here the loop may stop or start again until an acceptable result is found.

23.4 Lessons Learned

FP models are relatively easy to develop and can be understood quickly and intuitively. The data required to create a FP model is sourced from flat files or from the i2 RHYTHM Active Data Warehouse. The flat file modelling is normally used to identify the data requirements in an early phase of an APS implementation project.

FP has been developed continuously and improved by i2 over more than 10 years, based on a large variety of real life implementation projects. The FP product embodies an enormous variety of modelling features and can be employed in many industry branches. But of course there are still a number of cases where the modelling functionality is not sufficient to properly reflect the real planning problem. In most cases work-arounds can be developed – extending the standard planning functionality of FP, at the cost of increased maintenance effort of the solution.

The FP on-line help function provides information about the modelling concepts and helps to solve most modelling problems. Still a lot of aspects are not completely described in the documentation though included properly in the program.

The user has access to a variety of useful information on the user interface (the FP client user interface is usually Windows based) and can customize reports according to the planning workflow. Two effects of the implementation of an intelligent planning tool have to be lined out here. First, automated decisions are taken by the tool according to a committed set of business rules so the planner can concentrate on relevant aspects and exception handling which increases the planner productivity. Second, the planner is forced by the tool to plan according to the predefined rules which guarantees a stability in decision-making (see also Chap. 17).

24 Scheduling of Synthetic Granulate

Marco Richter and Volker Stockrahm

axentiv AG, Uhlandstraße 9, 64297 Darmstadt, Germany

This case study deals with a project which has been finished in its first version in the process industry in 2000 with a quite early release of APO PP/DS and which has been further improved by new releases of SAP APO with their additional functions and the integration of more parts of the supply chain. It was the first APO PP/DS project that managed to keep up with the difficult scheduling requirements in the field of the chemical and process industries.

This case study is structured in the following sections: First, the general production process of the synthetic granulate in the featured plant is presented. This chapter focuses on the special planning problems which occurred in this example. Subsequently, the modelling of the production process in APO PP/DS is described in detail, and some more information about modelling production processes in APO PP/DS are provided in addition to the general information given in Chap. 10 as well as a short view to the planning process. At the end of this case study the results of this APO implementation are estimated briefly as they could be measured today and the lessons learned are presented.

24.1 Case Description

The production process dealt with in this case study is the production of synthetic granulate. In technical terms it is a four step hybrid-flow-shop production process. The granulate is widely used in many different industries, especially in the automobile and pharmaceutical industries. About 3000 different products make up the full product spectrum which grows and changes rapidly.

The basic principle of the process (see Fig. 24.1) is melting the undyed granulate in extruders, adding colour substances and perhaps other additives and extruding the coloured granulate again. Depending on the product type, a mixer is used afterwards to create homogenous batches. If the granulate is shipped in bags, an automatic bag filling machine may be used. Otherwise the filling of the granulate in different types of containers is done directly at the extruders or mixer. Depending on the production sequence, transport containers may be needed during a part of the production process. At the end of the production process some more days are needed for the necessary quality checks.

The selection of resources for the production process depends on the product type. For the extrusion process several individual extruders can be used.

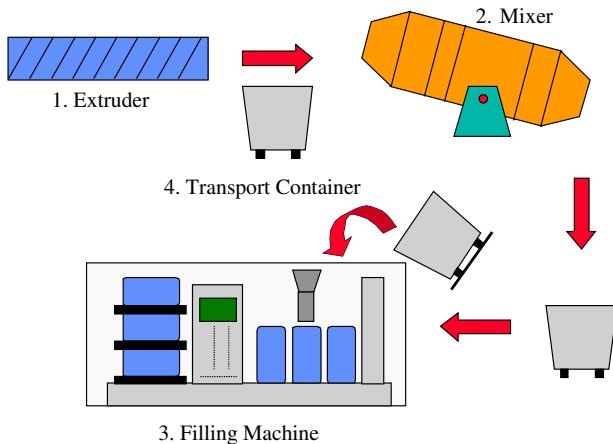


Fig. 24.1. Production process

These resources differ with respect to speed, types of colour that can be added and types of undyed granulate that can be processed. The actual usage of the individual extruder depends on the product type and the lot-size of the production order. Generally speaking, for each product there are several extruders with different priorities that can be used for the extrusion process. As there is a high variety of products with very different chemical and physical characteristics, the scheduling of orders on the extruders is very critical. Depending on the sequence of the production orders, setup times for cleaning the extruders vary between nearly no setup time and up to five hours.

Products with special quality requirements by the customer have to be mixed afterwards to create batches with homogeneous characteristics. For this part of the production process several mixers with different capacities are available. The selection of the mixer is lot-size dependent. Also the setup times on the mixers depend on the sequence of production orders, but the scheduling is less critical, as the setup times are shorter and the mixer is usually not a bottleneck. Granulates which are shipped in bags can be packed in two different ways. The first alternative is packing directly at the extruder or mixer which requires no additional resources. This procedure is chosen for production orders with small lot-sizes. Large production orders are packed with a special automatic filling machine which has to be planned separately. The setup times are sequence dependent as well, but less problematic than setup times on extruders. A further resource group, the transport containers, is needed, if a product needs the mixer or the filling machine or both of them. As the automatic filling in bags does not take place directly at the extruder, the transport containers are used to transport the loose granulate from extruder to mixer and further on to the filling machine. Since the number of available transport containers is limited, they must be considered as a relevant resource. The last resource group is the personnel required to operate

the machines. Several different qualification groups can be distinguished, all of them have to be considered for production planning. If several workers are not present, use of certain machines might not be possible and production must be rescheduled.

24.2 Objectives

As mentioned before, the sequencing of production orders is the critical task in the planning process to avoid setup times and costs. An ideal sequence of production orders regarding the setup times would be a sequence starting with a very bright colour (e.g. yellow) and ending with a very dark colour like brown or black. This sequence leads to no setup times between the orders (as there is no cleaning necessary when changing from bright to dark) and in a long setup / cleaning activity at the end of this "campaign". This setup optimization task and the fact that many different resource combinations for the products must be considered makes it hard for the production planner to generate feasible and economical plans in a short period of time. Especially the setup problems usually have been solved by building large standardized campaigns of similar products. Moreover, the plans once generated could not be changed easily when machines broke down or special short-term orders had to be fulfilled. This situation has been tackled by allowing buffer times in the production schedule. If this buffer time was not needed, the production capacities were not exploited to their maximum. So there has been the clear demand for an intelligent production planning and scheduling solution. This and as the integration with SAP R/3 should be as seamless as possible led to the decision to implement SAP's *Advanced Planner and Optimizer (APO)*.

24.3 Modelling the Production Process in APO PP/DS

Within this section of the case study not only the actual modelling of the production process in APO is described, but also some general principles of modelling processes in APO for production planning and scheduling in addition to the general description in Chap. 10. When this project was introduced for the first time, APO release 2.0 was the most current release. Although we are talking about mySAP SCM release 4.0 today, the basic principles of modelling supply chains in APO are still the same. Also the presented case has already undergone some release changes but still works in the original configuration and model. Of course, not all options of modelling with APO can be presented here.

24.3.1 General

In Chap. 10, the groups of data needed for planning have been defined. Especially the necessary master data will be explained in this chapter.

To use the PP/DS module of APO for planning and scheduling in industry, basically the following groups of master data must be maintained in the system:

- locations,
- products or parts,
- resources,
- Production Process Models (PPMs),
- setup matrices and
- supply chain models.

Additionally, transactional data (e. g. sales orders, planned orders, inventories and setup states of resources) will be needed for planning. As the APO is using a standard R/3 basis system to maintain the system functionality, it uses a relational database of its own to maintain master and transactional data. Therefore, data are not handed over using flat ASCII files which are read by the system on start unlike most other advanced planning systems. Considering this, some information about filling the system with data will be provided:

Usually, a special interface provided by SAP will be used to connect the APO system to an R/3 system (APO Core Interface, a part of the R/3 PlugIn). This interface generates the master data by an initial upload and communicates the transactional data as soon as they are changed by one of the systems. This guarantees the fastest and most recent data transmission. Nevertheless, other interfaces to non-R/3-systems may be used as well.

24.3.2 Locations

The *location* is the first step, when creating a model. As APO is an integrated supply chain planning tool, it is important that all the subsequent data can be assigned to individual locations. Although for the *supply network planning* there are several kinds of locations (supplier, production plant, distribution centre, customer, transportation zone, MRP area, transportation service provider and terminal), for the PP/DS only the production plants are relevant. This makes up one location – a production plant – for this production process. This data object corresponds with the organisational R/3 data object "plant" and is transferred using the standard Core Interface.

24.3.3 Products

For every product which is to be planned in APO (final product or raw material) a set of *product master data* has to be generated. The APO philosophy for the selection of the "relevant" products suggests that only critical materials should be planned in APO. So, one will usually plan the final products and some of their critical components in the APO system. Thus, only a portion of the materials contained in the complete bill-of-materials (BOM) is

transferred to APO. In this particular case, the final products – the coloured granulate – and the undyed granulate are planned in APO, although the BOMs in R/3 contain many additional components like some additives. But these are no critical components and are not planned in APO. The complete BOM is exploded in R/3 and the additional components are generated as secondary demands, when a generated or changed order is retransmitted to the ERP system.

A lot of the settings and product properties are not relevant for PP/DS planning and are not presented here. The important values for PP/DS are

- basic unit of measurement,
- alternative unit of measurement,
- lot-size calculation,
- planning method and
- procurement method.

The *units of measurement* are taken over automatically from the R/3 system, depending on the product. It is usually “unit” or “kg” for this production process. Regarding the *lot-size calculation* APO offers the following options: fixed lot-size, lot for lot with a maximum/minimum lot-size and lot-for-lot without maximum/minimum lot-size. For the lot-for-lot calculation a rounding value can be defined. In our case, all the products use lot-for-lot with maximum/minimum lot-size. All these individual values are taken from the R/3 system.

The *planning method* describes how APO will react, when a demand is transferred. If the planning method is “automatic planning”, the system checks the availability and – if the check is negative – creates a planned order or a purchase proposal (depending on the procurement method). If “manual planning with check” is selected, the system checks the availability and creates an alert in the *Alert Monitor*, if the check is negative. But it creates no orders of any kind. The third alternative “manual planning without check” always assumes that there is enough material to fulfil the demand. The *procurement method* determines, what APO will do if a demand cannot be fulfilled using stored materials. The procurement method offers the settings “in-house production”, “external procurement” or “in-house production or external procurement”. The last option is “direct procurement from other plants”.

When “in-house production” is selected, the system creates a planned order for the product, considering resource capacity availability and material availability simultaneously. When “external procurement” is selected, the system creates a purchase proposal. In case of “in-house production or external procurement” the cheaper alternative is selected, so costs for production and external procurement have to be maintained.

The planning method “automatic planning” and the procurement method “in-house production” have been selected for coloured granulates (the final products) in the first version of this case. So, automatically planned orders

were created and sent back to R/3 immediately, if a sales order cannot be fulfilled using stored materials or work-in-process. The undyed granulate status was set to “manual planning with check”. So, if there was not enough material for production, a warning in the Alert Monitor has been created. The reason for this was that the undyed granulate production which takes place in another plant of the same company was not yet integrated in the PP/DS planning process. As soon as this integration was completed, it was possible to check the availability of the corresponding undyed granulate and create a planned order for it immediately in the other plant together with a transportation order to the plant, where the coloured granulate is produced. As soon as this second plant was integrated into SAP APO a common planning process was established using a APO PP/DS planning heuristic working with finite capacities (“finite MRP”) to substitute the conventional MRP run in R/3. The automatic planning functionality proved to cause too many changes in the production schedule. Figure 24.2 shows the actual planning process and communication between APO and the ERP system.

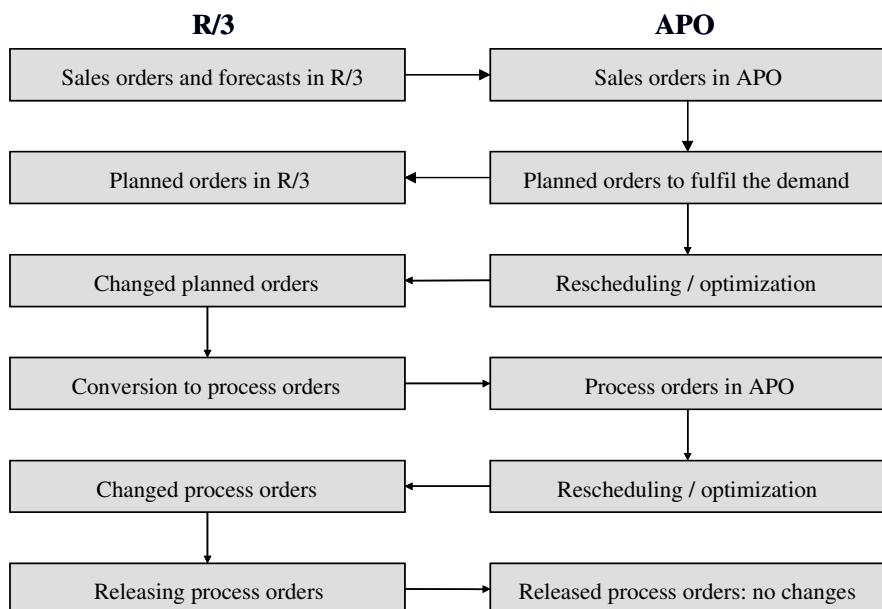


Fig. 24.2. Planning process using APO and R/3

24.3.4 Resources

APO uses several different types of *resources* for different planning requirements. For PP/DS *single* or *multi-activity resources* are used. Planning on

these resources is not based on periods. They use a continuous time stream, and orders are scheduled using seconds as time units.

Single-activity resources always have the capacity of 1 without any unit of measurement. They represent a machine that can only process one order at the same time. Multi-activity resources are used to model either groups of identical machines which lead to a capacity of more than 1 without a dimension or single machines which can process more than one order at the same time and every order requires a certain amount of capacity. For example, an oven can have a capacity of 10m^3 , and every order processed in the oven at the same time requires some volume of the oven. Several orders can be processed in the oven at the same time as long as the sum of their individual capacity requirements does not exceed 10m^3 .

Not only different capacity types can be distinguished in APO, also the *usage* of a resource is indicated. The resource types “production”, “transportation”, “storage” and “handling” are possible, but only the production resource is relevant for PP/DS.

Capacities can be defined in multiple variants in APO. In this way one can model different capacities for e.g. different shifts or reduced capacities for breakdown times. Besides the amount of capacity there is the possibility to indicate, when the capacity of a resource can be used. While the *factory calendar* describes on which days the resource can be used or not because of e.g. weekends, holidays etc., the *resource calendar* describes the working times for the working days. So, for the working days, the start time and end time of resource availability are specified. Additionally the resource usage can be defined to allow buffer times or reserve capacity for some reasons. This resource usage is measured in percent. Further settings concerning properties of resources can allow some overlap of orders without creating an alert. For each resource a flag can be set whether the actual capacity load should be considered during scheduling (“*finite planning*”) while another flag indicates whether the resource is a *bottleneck*. If the bottleneck flag is set, the system schedules an order first on the bottleneck resource and then the other activities of this order on the other (non-bottleneck) resources. To model sequence dependent setup times, a *setup matrix* must be created and assigned to the resources where these setup times occur. The setup times automatically reduce the resources’ capacity.

The setup times and costs are the only relevant factors to build lots on the resource and activity level as it is done in PP/DS. When using the optimizing algorithms to reschedule an initial plan, activities are planned on the resources in an order which creates the fewest losses by setup times and costs while concerning lateness, production costs and makespan simultaneously.

The classic lot-sizing which regards the trade-off between setup, transportation and storage costs has to be done in the mid-term planning, using the *Supply Network Planning*.

For the granulate production process, the following resources have been created in the APO system:

- a single-activity resource for each extruder,
- three multi-activity resources for the three personnel groups,
- one single-activity resource for the filling machine,
- one single-activity resource for each mixer and
- one multi-activity resource for the transport containers.

The extruders have been marked as bottleneck resources. So the system first schedules the extruders, as there is the biggest planning problem. All the resources, except the quality testing, are available for 24h on work days, the quality testing department works for ten hours. Although the quality testing is in fact just a dummy resource – there is no finite planning – the exact working times are necessary to model that the quality testing takes three days. Quality testing always starts at the beginning of a shift (6 a.m.) and ends always at the end of a shift (4 p.m.). The use of a dummy resource is necessary in APO, as waiting times without a resource cannot be modelled. If a resource shows a variable capacity, an additional capacity can be defined for each shift to represent the actual number of available workers or transport containers. If no specific capacity is given for a shift, the standard capacity for that resource will be used.

The following Table 24.1 shows the individual resource properties, i. e. the detailed definition of the resources. In fact there are some more fields which can be used in APO, but only the essential ones are described here.

Table 24.1. Modelling the resources

Name	Type	Start	End	Usage	Matrix	Bottleneck	Finite	Capacity	Unit
PERS_1	Multi	00:00:00	24:00:00	90%			X	0-2	
...
PERS_N	Multi	00:00:00	24:00:00	90%			X	0-12	
FILLING	Single	00:00:00	24:00:00	90%	SHORT		X	1	
MIXER_1	Single	00:00:00	24:00:00	90%	SHORT		X	1	
...
MIXER_M	Single	00:00:00	24:00:00	90%	SHORT		X	1	
EXTRUDER_1	Single	00:00:00	24:00:00	90%	LONG	X	X	1	
...
EXTRUDER_O	Single	00:00:00	24:00:00	90%	LONG	X	X	1	
QUALITY	Single	06:00:00	16:00:00	100%				1	
TRANSPORT	Multi	00:00:00	24:00:00	90%			X	30-40	

For the synthetic granulate process two setup matrices (SHORT and LONG) have been defined. They both contain the same setup keys, but different setup times for the individual product combination. The matrix with the longer setup times is assigned to the extruders, the one with the shorter setup times to the mixers and the filling machine. The matrix with the shorter entries can be regarded as a copy of the first matrix with all entries divided by a constant factor. The huge number of products can be reduced for the setup

matrices as a product in different shipment containers is represented by individual product numbers. The actual setup matrices contain about 2,000,000 entries each. The generation of the setup matrices could not be handled manually, so a special ABAP/4 programme in APO generates the setup times, using physical and chemical characteristics of the products (a non-standard functionality).

24.3.5 Production Process Models

The *Production Process Model (PPM)* is the most essential element of an APO model. As presented in Chap. 10, it represents both the routing and the BOMs. So, here it is determined which resources are used for what time and which components enter or leave the production process. This is indicated at activity level. So the production step, when a component is needed or ready, is described precisely. Also the temporal relations between single production steps are defined. APO is the first APS which actually uses the complete PPM concept, while most other systems work with separated routings and BOMs.

According to Chap. 10, a PPM is a hierarchical structure of elements which together form the production process. The elements of a PPM are:

- *Operations* which describe a group of production steps which take place on the same resource without interruption by other production orders;
- *Activities* as the single steps of an operation, e. g. setup, production, wait, tear down;
- *Activity relationships* which determine the sequence of the activities and their relative position in time;
- *Modes* which describe the resource or the alternative resources an activity can use and their duration;
- *Capacity requirements* for the primary and secondary resources of each mode;
- *Logical components* which serve as containers for groups of physical products (inputs or outputs) and are attached to activities;
- *Physical components* which describe the groups of real products represented by the logical components;
- The list of products which can be produced using this PPM (which may be all or just a part of the output components) and the lot-size ranges for which the PPM is valid.

As mentioned at the beginning in the description of the production process, many different routings through the process exist depending on the product. Here is presented the “maximum” PPM for a product which uses the extruder, the mixer, the filling machine, the transport containers and several personnel resources. First, the general modelling possibilities and principles for the elements of a PPM are presented, immediately after each element. A

practical example is given by showing how the synthetic granulate process was modelled in this step.

For every production step which takes place on another resource an *operation* is defined, as long as this resource is not only needed as a *secondary resource*, parallel to the primary resource (e.g. a worker who is needed to operate the machine). If sequence dependent setup times shall be used, the *setup key* which identifies the manufactured product in the setup matrix is specified in the operation. As there can only be one setup activity per operation the setup key is not specified in the setup activity. Regarding the granulate production process, there is a maximum of five operations. The transport containers cannot be modelled as secondary resources for reasons to be explained later. Therefore they require an operation of its own. The naming of the location is necessary because the location is needed to identify the setup key. Using a graphical representation, the PPM structure is described in Fig. 24.3.

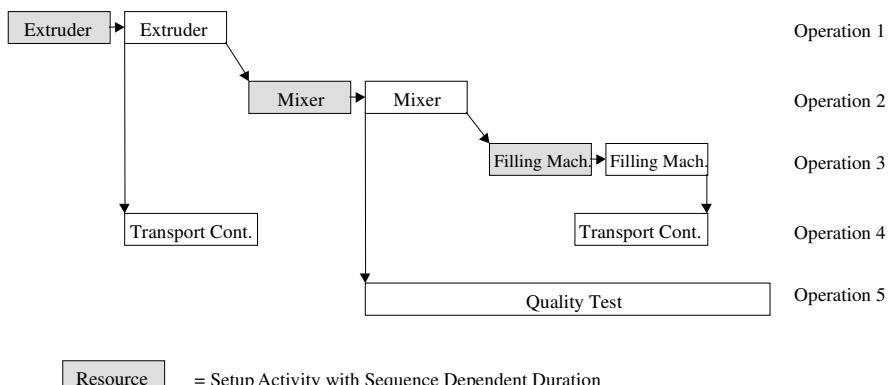


Fig. 24.3. PPM structure of the production process

The complete list of operations for this example is shown in Table 24.2.

Table 24.2. Modelling the operations

Operation key	Description	Setup key	Location
0010	Extruder	SETUP_KEY	GRANULATE_PLANT
0030	Mixer	SETUP_KEY	GRANULATE_PLANT
0050	Filling machine	SETUP_KEY	GRANULATE_PLANT
0070	Transport container		
0090	Quality test		

Every operation possesses at least one *activity*, usually the production activity. In many cases more than one activity will be defined in one operation, to model a single production step. The standard types of activities are “setup”, “production”, “wait” and “tear down” which all can be used only once in one operation. The activities are marked by an “S”, “P”, “W” or “T” accordingly. Only the setup activity can have a sequence dependent duration which says the duration is looked up in the setup matrix. If this feature is to be used, the flag for sequence dependent setup must be set. This flag makes it impossible to enter a duration for the setup activity later on in the modes. Another field in every activity determines the percentage of scrap which occurs during this activity. This scrap percentage is used to determine the order size which is necessary to produce the quantity ordered. All of the first three operations in this example have the same activity structure (see Table 24.3).

Table 24.3. Modelling the activities

Activity key	Description	Type	Sequ. dep. setup flag
0010	Cleaning	S	X
0020	Producing	P	

The transport container and quality test operations do not require a setup activity. So the quality test operation has only one activity type “production”, the transport operation has two activity types: “production” and “wait”. These activities represent the start and the end of the container usage. APO always lays a so-called cover chain around the activities which belong to the same operation. This guarantees that no activity of other production orders which are also processed on this particular resource is scheduled in between these activities.

The activities for the transport and the quality test operations can be described as shown in Tables 24.4 and 24.5.

Table 24.4. Transport activities

Activity key	Description	Type	Sequ. dep. setup flag
0070	Container start	P	
0080	Container end	W	

Every activity must have at least one *relationship* to another activity. APO does not automatically connect the orders and activities in the way

Table 24.5. Quality test activity

Activity key	Description	Type	Sequ. dep.	setup flag
0090	Quality test	P		

they are numbered or sorted in the tables. This provides a high amount of flexibility in modelling production processes. There are several types of activity relationships already introduced in Chap. 10 which are also used by APO: end-start, start-start, end-end and start-end. For every relation between activities a minimum and maximum time difference can be maintained. The *resource connection flag* at each activity relationship can be used to force APO to use the same primary resource, even if the activities belong to different operations. If the activities belong to the same operation, APO uses the same primary resource anyway, as there is not much sense in doing the setup on one machine and the production on another. The *material flow flag* must be set for every activity relationship which represents an actual flow of material. APO uses this path from input activity to output activity to calculate the total percentage of scrap which occurs during the whole production process.

Generally speaking, all the activities in this example which belong to extruder, mixer or filling machine have simply been connected with end-start relationships and no minimum or maximum time constraints. So they are processed one after the other, and the material can be stored for an infinite time. In fact the system will not let the waiting times between the activities of one operation become too long, as the order would consume transport container capacity while waiting. This would lead to a longer makespan, as other orders have to wait for the containers. This way the optimizer keeps the gaps short and no maximum time constraint is required. Activity relationships are shown in Table 24.6.

Table 24.6. Activity relationship

From	To	Type	Resource flag	Min. deviation	Max. deviation
Produce extr.	Setup mixer	end-start		0	0

The activities of the transport container operation have been connected to the other activities in a slightly different way. The “container start” activity has a start-start relationship to the activity “produce with extruder” with a minimum and maximum deviation of zero. So, the containers are occupied once production begins. The “container end” activity has an “end-end”

relationship also with no time deviation allowed with the last production activity (mixer or filling machine). The quality test activity has simply been connected to the last production activity as well.

Every activity must have one or more *modes*. A mode represents an alternative primary resource and may also contain one or more secondary resources which are used simultaneously. Every mode can be given a *priority* from A (first selection) to Z (only manually selectable). This priority influences the resource selection during incremental scheduling and optimization. Actually, the priorities represent penalty costs. If no priorities are used, the system always tries to use the fastest machine first. The mode also contains the information about the activity duration – depending on the resource the mode represents. Selecting the mode with the fastest machine therefore leads to the shortest activity duration. So, in APO the production speed and power of a resource can differ, dependent on the PPM which uses the resource. The resource speed is not maintained with the resource, but with the actual product/resource combination. The activity duration in the mode can be defined with a *fixed* and a *variable* part which grows with the order size. If the activity is a sequence dependent setup activity, the activity duration is taken from the matrix, and the fields in the mode are ignored.

For every mode there are the *capacity requirements* of the resources defined in a separate table. Here also the names of the secondary resources for the specific mode are given. A secondary resource is always covered as long as the primary resource with the same start and end times. The primary resource which is also given in the mode definition is always the so-called *calendar resource*. This says that the times of availability of this resource affect the availability of all the secondary resources in this mode. As mentioned in Sect. 24.3.4 an activity on a single activity resource always has the capacity requirement of 1. On multi-activity resources, capacity requirements other than 1 will occur and have to be defined here. Like the activity duration the resource consumption can be defined using a *fixed* and a *variable* part. In the Tables 24.7 and 24.8 an example is given for the modelling of modes and capacity requirements.

Table 24.7. Modes

Activity	Mode	Resource	Dur. fixed	Dur. variable	Break	In shift	Priority
Produce extr.	1	Extruder 1	0	1.5 h			A
Produce extr.	2	Extruder 5	0	2.75 h	X		C

As the last part of the PPM the *components* (inputs and outputs) are maintained. The entries are made for the activity, where the input or output occurs. It is defined whether the material enters or leaves the production

Table 24.8. Capacity requirements

Activity	Mode	Resource	Cap. req. fixed	Cap. req. variable	Calendar
Produce extr.	1	Extruder 1	1	0	X
Produce extr.	1	Worker	0.5	0	

process at the begin of the activity, at the end of the activity or continuously. Continuous input and output was introduced in the model when in a second step the production of undyed granulate started to be planned with APO as well. Contrary to the production of dyed granulate, which is a batch production process, the production of undyed granulate is a rather continuous production.

24.3.6 Supply Chain Model

Finally all the elements described above must be added to a *model*. A model allows actual planning with the system. Without creating a model the locations, products, resources and production process models cannot be used yet. Using the model philosophy one can create completely separated planning environments in the same system with the data in the same storage device (*liveCache*). Every model can have several *planning versions*. This says that several copies of the transaction data of the model are used to simulate different scenarios and to answer what-if questions. Only one planning version – the active version – is relevant for transferring the planning results back to the connected OLTP system and for receiving new planning data.

24.4 Planning Process

The planning process and the tools involved in this process are presented briefly in this section. The integration with other tools within APO and R/3 is described as well.

The demands used for the production planning and scheduling process are derived from R/3 sales orders (short term) and from APO Demand Planning (long term). At this company there is meanwhile also a tool for the mid term planning (master planning) in use, the APO Supply Network Planning. This tool is on the one hand used to perform a rough cut production capacity check in a horizon of the next 12 months across several business units and all production stages to provide a feedback to the demand planning (the sales people). On the other hand the production amounts created by the Supply Network Planning are passed on to the PP/DS modules (where available) or directly to the R/3 system.

The integration between the mid term planning (SNP) and the short term production planning (PP/DS) is important to perform a consistent hierarchical planning process. As the PP/DS plans only a rather short period of time (in this case three months) seasonal changes in the demand structure cannot be taken into account. The PP/DS would not trigger enough production in time respecting the limited resource capacity as the actual future demand is still outside its horizon. The SNP must provide this information to the PP/DS by handing over its own planned production and procurement amounts. In the short term production planning tool changes to this mid term plan by additional sales orders or more detailed resource availability information are made and the production schedule is planned in greater detail as the PP/DS uses more detailed master and transactional data.

In a nightly cycle the PP/DS production planning heuristic performs a planning run which is quite similar to the traditional MRP but which also takes resource and material availability into account simultaneously. The result of this planning run is a production schedule which is already feasible but not yet optimized.

After this production planning run the PP/DS optimizer is used each night to create an optimized production schedule in respect to setup times, machinery costs, delays and total lead time. For the optimization part the standard PP/DS optimizer based on a genetic algorithm is used. This optimized production plan is transferred automatically to the connected R/3 system.

In their daily work the planners use the graphical planning board to check the suggested production plan. Changes which may be necessary are performed using drag and drop functionality in a Gantt chart. Another important tool is the alert monitor which visualizes exceptions in the planning schedule and allows the planners to react directly. All changes which are made to the production schedule are transferred back to the R/3 system online without any delay. The execution of the production plan is still performed in the R/3 system e.g. the release and confirmation of process orders, as these are no planning tasks.

24.5 Results and Lessons Learned

The implementation of APO PP/DS proved to be very successful. As the APO detailed scheduling solution has now been in active use for more than three years in the first edition which covers only the dyed granulate production it can be stated that expectations are fully met.

Results show that the quality of the generated plans is as good as that of the plans created by experts. The difference between the plans created manually and those created by APO is the speed and the flexibility in planning: previously, it took several planners more than one day to create a production plan which was fixed for several weeks. Now, APO plans the same number of

orders in an optimizer run which takes approximately one hour. The frozen horizon could be reduced from about one week to one or two days. Only one of the planners is involved in the detailed production scheduling using the graphical planning board and the genetic algorithm optimizer, which runs every night. Two other planners can now concentrate on manufacturing execution and other important aspects of their daily work.

A new plan can be created with APO immediately. This is especially important in a critical situation, such as a machine breaking down. The integration with the R/3 system is seamless: no additional steps are necessary to transfer planning results and new orders between the systems.

The quality of the created plans can be measured in terms of production time consumed for a certain number of orders. In this case, the production plan generated by APO usually has a makespan less than the manually created plan. This results mainly from the excessive reserved buffer times which are no longer needed.

A very important factor is the acceptance by the user. This new planning tool has been fully accepted by the production planners who see that they have a powerful system to help them in their daily routine work and make the production more flexible and profitable.

24.5.1 Lessons Learned

Several lessons were learned in this project. One of the most important is that master data quality has to improve significantly in the connected R/3 system. As the R/3 is not only used to perform ERP functionality as before but also has to serve as master data source for an APS the quality of data must be improved. An APS reacts much more sensitive to master data inconsistencies than an ERP system because these master data are used for a very detailed planning process. Master data which are created and maintained for ERP functions will not be good enough.

Furthermore the integration of the planners in the project proved to be crucial. As they have all the knowledge which is necessary to create a good production schedule they must be part of the project from the beginning to ensure success. An APS project can never be brought to life without the planners who are supposed to work with that tool afterwards.

Another big issue is the integration aspect between APS and ERP. Although in this case a very good standard interface has been used, developed by the same software manufacturer of both the APS and the ERP, there was a lot of effort in testing until a smooth integration process could be guaranteed. Without this standard interface much more work, money and time would have been invested in this interface.

24.5.2 Outlook - Further APO Implementations Within this Company

The successful completion of the first APO project has led to further APO implementations at the company. Three more production plants are currently using PP/DS. With the help of experiences gained in the first PP/DS project, these projects were running smoothly and in schedule. In addition to these PP/DS implementations, the Supply Network Planning (SNP) module has been chosen to provide mid-term production planning across all three business units. This supply chain planning process has been in active use since the end of 2002. On the demand planning side, two of the three business units are using the APO Demand Planning module to provide the SNP module with the necessary demand forecast and support the sales people with detailed forecasts. Considering the short term sales and distribution side, the Global Available to Promise module (gATP) will complete the advanced planning functionalities of APO within this company in one of the business units.

Part V

Conclusions and Outlook

25 Conclusions and Outlook

Hartmut Stadtler

Darmstadt University of Technology, Department of Production & Supply Chain Management, Hochschulstraße 1, 64289 Darmstadt, Germany

The preceding chapters have shown the different steps of introducing an APS in industry, starting with an analysis of a given supply chain, its redesign and subsequently modelling the supply chain from long-term to short-term decision levels. The integration of all planning tasks relating to the order fulfilment process will result in a new era of enterprise wide and supply chain wide planning.

Thereby, an APS not only will yield improvements on the three crucial factors of competitiveness, namely costs, quality and time, but it will also allow for the

- making of processes more transparent,
- improvement of flexibility and
- revealing of system constraints.

Widely available information from all over the supply chain results in a *transparent* order fulfilment process. It enables companies and supply chains to provide customers with accurate information about the order status and provides alerts in the case that an unexpected event causes the delayed delivery of an order. However, before this happens a decision-maker can find and check alternative ways to fulfil the customer's order, either by a shipment from another warehouse or another production site or by offering parts of the next higher grade. Additionally, transparent processes will reduce waste along the supply chain, because waste, e. g. resulting from excessive inventories or resources with low utilization rates, will be recognized quickly and measures for its improvement may be introduced. More importantly, due to its optimization capabilities, an APS will keep waste to a minimum, right from the beginning.

With markets and customer expectations changing quickly, supply chains not only have to respond but to anticipate new trends. In some cases, this may be achieved by integrating key customers in the supply chain. On the other hand, *flexibility* comes into play, which can be discussed along two dimensions: One is to be able to cope with changes in actual demands given the current inventory position, equipment and personnel, the other is to be able to adapt to changing markets over time (sometimes called agility, see Pfohl and Mayer (1999)). An APS supports both dimensions. As an example, the ATP module can show ways of using existing inventories in the most effective manner. Also, Production Planning and Scheduling allows for the

re-optimization of a new mix of orders quickly. Flexibility is further enhanced by an APS, due to a significant reduction of the frozen (firmed) horizon (as an example see Chap. 24). Finally, midterm Master Planning should not only coordinate the decentralized decision units, but also plan for a reasonable degree of flexibility over time.

In order to improve competitiveness, the *revelation of system constraints* is a crucial part of a continuous improvement process (see also Goldratt (1999)). System constraints may be detected at different levels of the planning hierarchy. For example, midterm Master Planning will not only provide an optimal solution for a given situation, it also shows which constraints are binding, i. e. preventing a higher level of our objectives. Looking for ways to lift the system's constraints, e. g. by a more flexible employment of the workforce, will further improve competitiveness. This will give rise to defining several scenarios from which to choose. Compared with former times, defining a scenario and getting an answer is now a matter of hours, not weeks. Therefore, management and planning staff can work together more closely and effectively than before.

Some of the above statements may still be regarded as visions. But as our case studies have shown, there are already implementations of APS in industry that are showing impressive improvements. In order to extend these success stories to a wider range of companies and supply chains, four main topics have to be addressed carefully:

- improving modelling and solution capabilities of APS,
- linking APS to controlling,
- extending the applicability of APS to polycentric supply chains and
- providing special training for managers, employees and consultants.

Although APS are around for several years, additional features are expected to be introduced in the near future. However, the standard architecture of modules should remain stable. Experiences with some modules have shown that some restrictions still may exist in modelling a given (production) process adequately. Given that supply chains have to adapt to new market trends quickly, modelling should be easy to learn and fast to implement. Likewise, one should expect a similar modelling language for all modules provided by an APS vendor (unfortunately this is not always the case).

Furthermore, we experienced that not all models generated have been solvable within reasonable time limits or have not shown a satisfactory solution quality. However, minor changes in the model have improved solvability significantly. Hence, enhanced modelling capabilities and more robust solution procedures solving large problem instances are still looked for.

Models and plans are only of value if they have an impact on a supply chain's decision making and operations. This obvious statement is often ignored in practice. Once a plan has been accepted, it is stored in a safe place and business is done as usual. In order to improve the situation, some companies have introduced so-called SC managers to act as "liaison officers".

However, often their responsibilities are unclear and their power to intervene remains limited.

Today, management is more inclined to make use of tools from (the) controlling (department), like budgets or the balanced scorecard, than to rely on APS. Hence, greater emphasis should be placed on *linking APS modules with controlling*, either by linking decision models to key performance indicators or – even better – to incorporate APS into the toolset of controlling.

So far, APS are best suited for supply chains with centralized control, i. e. mainly for intra-organizational supply chains. Although information exchange, in principle, is no problem for APS implemented in an inter-organizational supply chain, the willingness to operate on the basis of “open books” (e. g. regarding costs and available capacities) cannot always be assumed. Although collaborative planning (Chap. 14) has been introduced, the knowledge of how to adapt plans generated in different planning domains is still in its infancy.

In order to use APS effectively, managers and employees must have *special training*, enabling them to interpret solutions, recognize interactions with other parts of the supply chain, set up scenarios and react to alerts appropriately. In addition to project management, the mastery of change management and the basics of computer science, consultants now must have knowledge and experience in generating adequate models of the supply chain for the different modules of an APS. These models, neither being too detailed nor too rough, have to support decision-making and must be solvable with reasonable computational efforts. Inadequate models may even deteriorate the position of the supply chain instead of improving it.

Introducing an APS is not just adding another software package to those already existing in a company. On the contrary, it will replace many individual software solutions formerly “owned” by individual employees. Also, some types of decisions, which formerly required several employees, like creating a detailed schedule for the shop floor, will now be made automatically. Consequently, some of the employees will have to change to other positions, which may result in some resistance to change. On the other hand, optimization capabilities of APS will yield better plans than before with the additional option of checking alternatives interactively, thus giving those involved a greater satisfaction.

Last but not least, one should bear in mind that introducing an APS changes the way an organization or supply chain works. The definition of processes fulfilling the needs of different market segments will have to be reflected within the organizational structure. For legally separated firms or profit centres within a single firm covering only a portion of a given process, an effective reward system has to be installed in order to achieve the best solution for a supply chain as a whole and not get trapped in isolated sub-optima (see e. g. Fleischmann (1999)).

Recent reports on APS implementation projects have shown that the above recommendations should be taken seriously. Some APS users may have observed a discrepancy between their expectations, the vendors' promises and the capabilities of consultants as well as the APS software. In order that all parties involved get a realistic view, prototyping seems to be a good choice. Also, great visions should not be approached in one step, instead a stepwise introduction of SCM ideas and software support seems more appropriate.

Great efforts are currently undertaken by APS vendors to match planning issues facing industry with the capabilities of an APS, e. g. the issue of safety stocks in a multi-level supply chain or the issue of incorporating lot-sizing (rules) at different (hierarchical) planning levels. An additional strategy towards a better fit is to devise specific modules focusing on the specific needs of certain lines of businesses.

As stated in the introduction, SCM and APS are closely related to new developments in information and communication technology. Since business via the Internet is growing rapidly, new challenges with respect to the order fulfilment process arise. Topics like "supply chain execution" or "customer relationship management" have been reconsidered in this respect. Some software vendors now even proclaim that an APS is only part of a much larger software suite. However, this should not take one's mind off the fact that integrating modules and data flows within an APS in a consistent manner still remains a formidable task.

In recent years, several empirical studies have been conducted aiming at the revelation of key success factors for an effective, superior supply chain. As one may expect, a great number of factors have been proposed and tested. Two noteworthy factors, which have been found to be significant, are "... structural elements such as having an integrated information system, and behavioral relationship building elements such as trust and commitment ..." (Jayaram et al., 2004). These findings are in accordance with our recommendations throughout this book.

With regard to future developments of SCM as a whole, one can expect that it will not only concentrate on the order fulfilment process alone, but will incorporate neighbouring processes like the product design or recovery process.

References

- Fleischmann, B. (1999) *Kooperation von Herstellern in der Konsumgüterdistribution*, in: Engelhard, E.; Sinz, E. J. (Eds.) Kooperation im Wettbewerb. Neue Formen und Gestaltungskonzepte im Zeichen der Globalisierung und Informationstechnologie, Wiesbaden, 68–196
- Goldratt, E. M. (1999) *Theory of constraints*, Croton-on-Hudson, N.Y.
- Jayaram, J.; Kannan, V. R.; Tan, K. C. (2004) *The influence of initiators on supply chain value creation*, accepted for publication in: Int. J. Prod. Res.
- Pfohl, H.-Ch.; Mayer, S. (1999) *Wettbewerbsvorteile durch exzellentes Logistikmanagement*, Logistik Management, Vol. 1., No. 4, 275–281

Part VI

Supplement

26 Forecast Methods

Herbert Meyr

University of Augsburg, Department of Production and Logistics,
Universitätsstraße 16, 86135 Augsburg, Germany

In Chap. 26 we will show how demand planning can be done when seasonality and trend are given. For a comprehensive and extensive introduction to forecasting in general the reader is referred to Hanke et al. (2001) or Waters (1992).

26.1 Forecasting for Seasonality and Trend

This section introduces *Winters' method* which is appropriate for multiplicative seasonal models (see Chap. 7). In Sect. 26.2 the parameters of Winters' method are initialized. This incorporates the introduction of *linear regression*, too. A working example illustrates the explanations.

26.1.1 Working Example

Figure 26.1 shows the sales volume of a supplementary product of a large German shoe retailer. The data are aggregated over the whole sales region and comprise a time horizon of four weeks. In our working example we use the first three weeks (days -20, ..., 0) as input and – starting with day 1 – try to estimate day by day the sales of the fourth week.

Two observations are striking when analyzing the data:

- There seems to be a common sales pattern with weekly repetition. Saturdays usually show the highest, Sundays the lowest sales volume of a week. So weekly seasonality can be assumed with a cycle length of $T = 7$ days.
- Sales per week appear to be continuously increasing. This is obvious when all four weeks are considered. But even within the first three weeks a (weaker) trend of growing sales is visible.

Since the amplitude of seasonality is increasing, too, a seasonal multiplicative forecast model seems justified. All subsequent explanations will be demonstrated by use of this working example.

26.1.2 Modelling Seasonality and Trend

As already shown in Sect. 7.2 a multiplicative seasonal model is characterized by the parameters a and b describing the trend and the seasonal coefficients

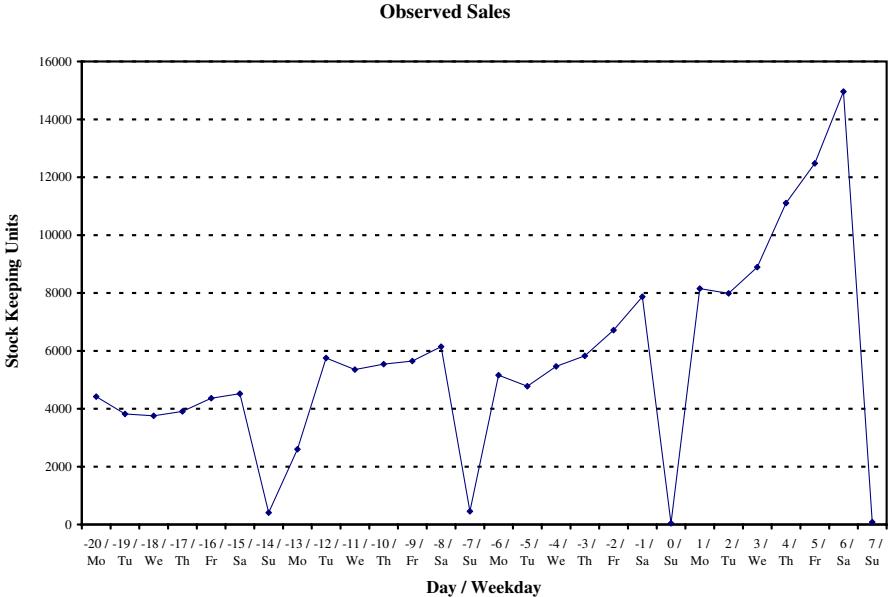


Fig. 26.1. Sales volume of a supplementary product of a German shoe retailer

c_t modelling the seasonality of period t . Figure 26.2 makes clear that the trend is expressed by the linear function $a + b \cdot t$ with t denoting periods of time (e.g. days in our working example).

A sales volume x_t observed in period t is modelled by

$$x_t = (a + b \cdot t) \cdot c_t + u_t \quad (26.1)$$

with the seasonal coefficients c_t increasing or decreasing the trend. Please note, if all seasonal coefficients are equal to 1, seasonality disappears and the model reduces to a simple trend model (see e.g. Silver et al. (1998, pp. 93)). The erratic noise u_t makes things difficult. Because of the randomness that is represented by u_t the other parameters cannot be measured exactly, but have to be predicted. In the following the superscript $\hat{\cdot}$ is used to distinguish between an observation (no superscript) that has been measured and its forecast being estimated without this knowledge.

Let \hat{a}_t , \hat{b}_t and $\hat{c}_{t-T+1}, \dots, \hat{c}_t$ denote the forecasts of a , b and the seasonal coefficients c that are valid in period t . Then (26.1) can be engaged to estimate the sales volume \hat{x}_{t+s}^t of all subsequent periods $t+s$ ($s = 1, 2, \dots$). For example, the sales volume of the next seasonal cycle is predicted in period t by

$$\hat{x}_{t+s}^t = (\hat{a}_t + \hat{b}_t \cdot s) \cdot \hat{c}_{t+s-T} \quad (s = 1, \dots, T). \quad (26.2)$$

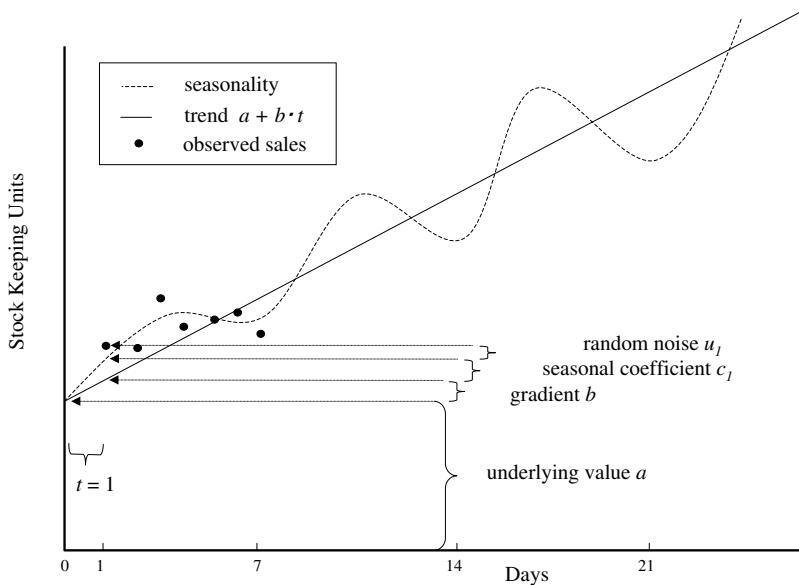


Fig. 26.2. Modelling seasonality and trend

The method of Winters described in the next subsection iteratively computes the sales estimation of only the subsequent period $t+1$. For this reason we can use the simpler notation \hat{x}_{t+1} instead of \hat{x}_{t+1}^t .

26.1.3 Winters' Method

The method of Winters (1960) basically builds on (26.2) and the principle of exponential smoothing which has been introduced in Chap. 7. Since sales are predicted indirectly via \hat{a}_t , \hat{b}_t and \hat{c}_t in (26.2), these three types of parameters have to be estimated by means of exponential smoothing instead of the sales volume itself (as it has been done by (7.5) for models without trend and seasonality). Remember the generic principle of exponential smoothing:

$$\text{new forecast} = sc \cdot \text{latest observation} + (1 - sc) \cdot \text{last forecast}. \quad (26.3)$$

The *new forecast* of the current period estimating the subsequent period(s) can be calculated by smoothing the *latest observation*, i. e. the observation in the current period and the *last forecast* that has been made to predict the current period's observation. The smoothing constant $sc \in (0; 1)$ determines the weight the new observation has. The higher the smoothing constant the more importance is given to the latest observation. Table 26.1 summarizes how Winters applies exponential smoothing in period $t+1$ to estimate the parameters \hat{a}_{t+1} , \hat{b}_{t+1} and \hat{c}_{t+1} determining the sales forecast \hat{x}_{t+2} of the subsequent period (26.2).

Table 26.1. Exponential smoothing applied in Winters' method

new forecast	smoothing constant sc	latest observation	last forecast
\hat{a}_{t+1}	α	x_{t+1}/\hat{c}_{t+1-T}	$\hat{a}_t + \hat{b}_t \cdot 1$
\hat{b}_{t+1}	β	$\hat{a}_{t+1} - \hat{a}_t$	\hat{b}_t
\hat{c}_{t+1}	γ	x_{t+1}/\hat{a}_{t+1}	\hat{c}_{t+1-T}

These three types of equations become clear when looking at our working example. We start our computation at the end of day $t = 0$. Table 26.2 further illustrates this proceeding:

1. Initialization:

In order to get things work initial values \hat{a}_0 , \hat{b}_0 and $\hat{c}_{-6}, \dots, \hat{c}_0$ (seasonal coefficients for each weekday) have to be given. As examples Subsections 26.2.2 (for $\hat{c}_.$) and 26.2.3 (for \hat{a}_0, \hat{b}_0) show how these values can be computed from the sales observations of the first three weeks (day -20, $\dots, 0$). For the moment we will accept in blank the values $\hat{a}_0 = 5849.0$, $\hat{b}_0 = 123.3$ and $\hat{c}_{-6} = 1.245693$ that are used in Table 26.2.¹

2. Estimating the sales volume of period $t + 1$:

Applying (26.2) we can estimate the sales volume \hat{x}_1 of period 1:

$$\hat{x}_1 = (\hat{a}_0 + \hat{b}_0 \cdot 1) \cdot \hat{c}_{-6} = (5849.0 + 123.3) \cdot 1.245693 = 7440$$

The linear trend $(\hat{a}_0 + \hat{b}_0 \cdot 1)$ does not consider any seasonal influences and will therefore be called “deseasonalized”. Since sales on Mondays are (estimated to be) about 25 % higher than average weekly sales ($c_{-6} = 1.245693$), the trend has to be increased accordingly. Please note that at the end of day 0 sales of day 2 (Tuesday) could roughly be estimated to amount to $(\hat{a}_0 + \hat{b}_0 \cdot 2) \cdot \hat{c}_{-5} = 6798$. However, a more accurate forecast of \hat{x}_2 can be given at the end of day 1 because the sales observation x_1 of day 1 offers further information.

3. Observation in period $t + 1$:

In day 1 sales x_1 of 8152 stock keeping units (SKU) are observed.

4. Using the latest observation to update trend and seasonal coefficients:

The latest observation x_1 improves the forecast of the trend and Monday's seasonal coefficient. So the smoothing constants $\alpha = 0.8$, $\beta = 0.8$ and $\gamma = 0.3$ are applied to the three exponential smoothing equations defined in Table 26.1:

¹ Note that the initial seasonal coefficients $\hat{c}_{-6}, \dots, \hat{c}_0$ are printed with two additional digits in order to indicate that high precision floating point arithmetic – commonly used in APS, programming languages, and spreadsheets – is applied throughout the working example.

Table 26.2. Forecasting the fourth week using Winters' method

t	weekday	x_t	\hat{x}_t	\hat{a}_t	\hat{b}_t	\hat{c}_t
-6	Monday				1.245693	
-5	Tuesday				1.115265	
-4	Wednesday				1.088853	
-3	Thursday		Initialization		1.135378	
-2	Friday				1.178552	
-1	Saturday				1.229739	
0	Sunday		5849.0	123.3	0.006520	
1	Monday	8152	7440	6429.8	489.3	1.2523...
2	Tuesday	7986	7717	7112.3	643.9	1.1175...
3	Wednesday	8891	8445	8083.6	905.8	1.0922...
4	Thursday	11107	10206	9624.0	1413.5	1.1410...
5	Friday	12478	13008	10677.6	1125.5	1.1756...
6	Saturday	14960	14515	12092.8	1357.3	1.2319...
7	Sunday	81	88	12628.5	700.0	0.0065...

- (a) The underlying value \hat{a} . of the trend is updated as follows:

$$\hat{a}_1 = \alpha \frac{x_1}{\hat{c}_{-6}} + (1 - \alpha)(\hat{a}_0 + \hat{b}_0) = 0.8 \cdot \frac{8152}{1.245693} + 0.2 \cdot 5972.3 = 6429.8$$

Thereby, the “deseasonalized” sales volume $\frac{x_1}{\hat{c}_{-6}}$ of day 1 serves as a new observation for the underlying value, while $(\hat{a}_0 + \hat{b}_0 \cdot 1)$ was the forecast of the deseasonalized sales of day 1 which has been obtained in period 0 (26.1).

- (b) Using \hat{a}_1 , the new gradient \hat{b}_1 can be calculated:

$$\hat{b}_1 = \beta(\hat{a}_1 - \hat{a}_0) + (1 - \beta)\hat{b}_0 = 0.8(6429.8 - 5849) + 0.2 \cdot 123.3 = 489.3$$

Between day 0 and day 1 the underlying value a . has been increased from \hat{a}_0 to \hat{a}_1 . Since \hat{a}_1 is based on the latest sales observation x_1 , this is interpreted as the “new observation” of the gradient b . which again has to be exponentially smoothed.

- (c) The same procedure is applied to the seasonal coefficient \hat{c}_1 :

$$\hat{c}_1 = \gamma \frac{x_1}{\hat{a}_1} + (1 - \gamma)\hat{c}_{-6} = 0.3 \frac{8152}{6429.8} + 0.7 \cdot 1.245693 = 1.2523$$

\hat{c}_{-6} was the last forecast of the Monday’s seasonal coefficient. The new observation of the seasonal influence of a Monday, however, is

achieved by dividing the observed sales volume x_1 (including seasonal influences) by \hat{a}_1 (deseasonalized).

5. Stepping forward in time:

Now we can go one day ahead (increasing t by 1) and repeat the steps (2) to (5). At the end of day 1 the sales volume \hat{x}_2 of day 2 is estimated by

$$\hat{x}_2 = (\hat{a}_1 + \hat{b}_1 \cdot 1) \cdot \hat{c}_{-5} = (6429.8 + 489.3 \cdot 1) \cdot 1.115265 = 7717$$

and so on...

Table 26.2 shows the results of Winters' method when applied to the days 2 to 7.

Figure 26.3 illustrates the consequences of a variation of the smoothing constants α and β of the trend. Generally, smoothing constants out of the intervals $\alpha \in [0.02; 0.51]$, $\beta \in [0.005; 0.176]$ and $\gamma \in [0.05; 0.5]$ are recommended (see Silver et al. (1998, p. 108)). In our working example, however, $\alpha = \beta = 0.8$ perform best, i.e. the few latest observations get a very high weight and smoothing is only weak. Thus, the forecast is able to react quickly to the progressively rising sales of the fourth week.

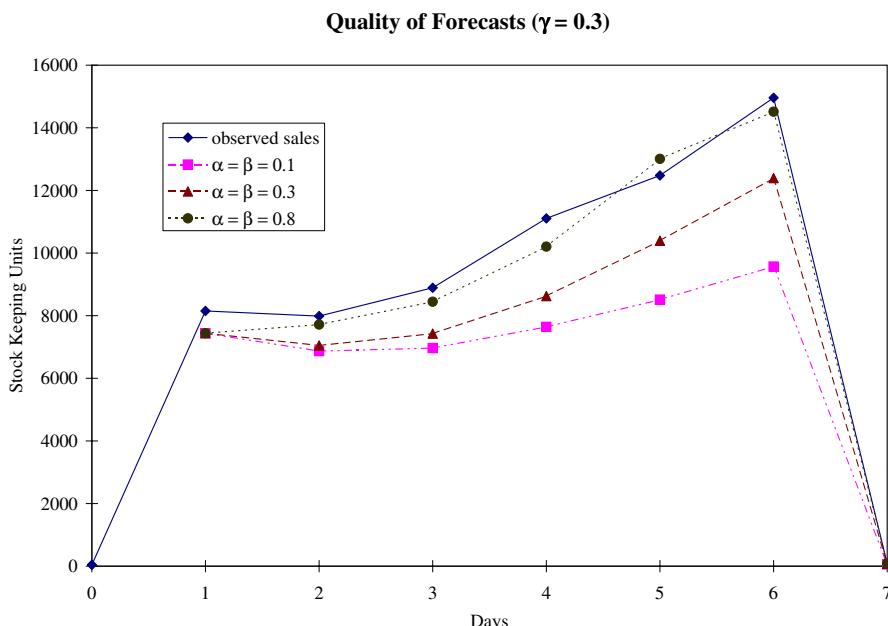


Fig. 26.3. Variation of the smoothing constants α and β

26.2 Initialization of Trend and Seasonal Coefficients

Until now we have not shown how the trend and seasonal coefficients can be initialized using the information that is given by the sales volume of the first three weeks. The next subsection demonstrates how the data basis can be improved if additional information is considered. Sections 26.2.2 and 26.2.3 finally present the initialization of the seasonal coefficients \hat{c}_s and the trend parameters \hat{a}_t and \hat{b}_t .

26.2.1 Consideration of Further Information

When looking at the data of the first three weeks (see Fig. 26.1) two phenomena seem to be contradictory to the assumption of a linear trend with seasonality:

1. Sales on Monday -13 are unexpectedly low. In weeks 1 and 3 sales on Mondays are clearly higher than sales on Tuesdays.
2. While the trend of weekly increasing sales is obvious, sales on Sunday 0 are much lower than sales on the respective Sundays of the first two weeks (days -14 and -7).

We want to know whether these inconsistencies are purely random or due to an identifiable actuator and get the following information:

1. In some parts of Germany Monday -13 was a holiday. Therefore, 58 % of the stores of the shoe retailer were closed this day.
2. Usually, shoe stores have to be closed on Sundays in Germany. Some few cities, however, granted a special authorization for sale. Starting with the third week $93\frac{1}{3}\%$ of these cities do not grant such an authorization any more.

We can now improve our data basis by exploiting this information about special influences in our further investigations. Therefore, the sales volume of day -13 is increased by 138.1 % ($x_{-13} = 2600 \cdot \frac{100}{100-58} = 6190.4761$) and sales on Sundays -14 (410 SKU) and -7 (457 SKU) are decreased by $93\frac{1}{3}\%$ so that $x_{-14} = 27.\bar{3}$ and $x_{-7} = 30.4\bar{6}$. In the next two subsections original sales are replaced by these corrected sales.

26.2.2 Determination of Seasonal Coefficients by the Ratio-to-Moving Averages Decomposition

The ratio-to-moving averages decomposition (see e. g. Makridakis et al. (1998, pp. 109)) is used as an example to determine the initial seasonal coefficients of Winters' method. In Sect. 26.1.3 we already applied the equation:

$$\text{observed sales in } t = (\text{deseasonalized sales in } t) \cdot (\text{seasonal coefficient of } t).$$

In other words, if we want to isolate seasonal coefficients, we have to compute

$$\text{seasonal coefficient of period } t = \frac{\text{observed sales in } t}{\text{deseasonalized sales in } t} \quad (26.4)$$

where the *deseasonalized sale in period t* is a sales volume that does not contain any seasonal influences. But how to determine such a value?

Considering our working example, the sales volume of a full week is apparently not influenced by daily sales peaks. So the most intuitive way to obtain sales data without seasonal influences is to compute daily sales averaged over a full week. This leads to average daily sales $\frac{4419+\dots+27.3}{7} = 3544.6, 4951.6$ and 5122.4 SKU for the weeks 1 to 3 (see Table 26.3). Thereby, the Thursday is settled in the middle of each week.

But we can employ the same procedure for each other time period of seven days, e.g. day -19, ..., -13, and assign the average daily sales 3797.7 to the medium Friday -16. By doing so we compute moving averages over a full seasonal cycle of 7 days for each day -17, ..., -3 which represent deseasonalized daily sales volumes. Table 26.3 illustrates the whole procedure.

In a next step we apply (26.4), thus setting the observed sales x_t in *ratio to* the deseasonalized *moving averages* (remember the name of the algorithm). The result are multiple observations of seasonal coefficients $o_{\text{weekday}}^{\text{week}}(t)$ for each day of the week (three for a Thursday and two for each other weekday) which still contain the random noise u_t .

In order to reduce this randomness, now we compute the average seasonal coefficients $o_{\text{weekday}}^{\text{aver}}$ of each weekday (Table 26.4). For example, for the Thursday we get

$$\begin{aligned} o_{\text{Thursday}}^{\text{aver}} &= \frac{o_{\text{Thursday}}^{\text{week } 1}(-17) + o_{\text{Thursday}}^{\text{week } 2}(-10) + o_{\text{Thursday}}^{\text{week } 3}(-3)}{\text{number of weeks}} = \\ &= \frac{1.1031 + 1.1188 + 1.1377}{3} = 1.1199 \end{aligned}$$

If a pure trend without any seasonal influence is given, one would expect all seasonal coefficients to equal 1 (see Sect. 26.1.2), thus summing up to 7 for a weekly seasonal cycle. As we can see in Table 26.4, the sum of our average seasonal coefficients $o^{\text{total}} = \sum_{\text{day}=\text{Monday}}^{\text{Sunday}} o_{\text{day}}^{\text{aver}} = 6.9045$ falls short of 7. To reflect the trend correctly, we have to normalize our o^{aver} by multiplying them with the constant $7/o^{\text{total}}$. The resulting final seasonal coefficients for *Monday* ... *Sunday* are already known as $\hat{c}_{-6}, \dots, \hat{c}_0$ from Table 26.2.

26.2.3 Determining the Trend by Linear Regression

Finally it will be shown how the trend parameters a and b can be determined. When “deseasonalizing” the observed sales by dividing through c_t one can see from (26.5) that the trend $a + b \cdot t$, distorted by some random noise $\frac{u_t}{c_t}$, results:

$$d_t = \frac{x_t}{c_t} = \frac{(a + b \cdot t) \cdot c_t + u_t}{c_t} = a + b \cdot t + \frac{u_t}{c_t}. \quad (26.5)$$

Table 26.3. Ratio-to-moving averages decomposition

week	day	weekday	(corr.) x_t	moving aver. (ma_t)	$o_{weekday}^{week}(t) = \frac{x_t}{ma_t}$
1	-20	Monday	4419		
1	-19	Tuesday	3821		
1	-18	Wednesday	3754		
1	-17	Thursday	3910	3544.6	1.1031
1	-16	Friday	4363	3797.7	1.1489
1	-15	Saturday	4518	4074.0	1.1090
1	-14	Sunday	(27.3333̄)	4302.3	0.0064
2	-13	Monday	(6190.4761)	4535.1	1.3650
2	-12	Tuesday	5755	4719.0	1.2195
2	-11	Wednesday	5352	4951.1	1.0810
2	-10	Thursday	5540	4951.6	1.1188
2	-9	Friday	5650	4804.1	1.1761
2	-8	Saturday	6143	4664.6	1.3169
2	-7	Sunday	(30.4666̄)	4680.6	0.0065
3	-6	Monday	5158	4721.8	1.0924
3	-5	Tuesday	4779	4873.8	0.9806
3	-4	Wednesday	5464	5120.8	1.0670
3	-3	Thursday	5828	5122.4	1.1377
3	-2	Friday	6714		
3	-1	Saturday	7872		
3	0	Sunday	42		

Table 26.4. Reducing randomness of seasonal coefficients

week	Mo	Tu	We	Th	Fr	Sa	Su	\sum
1				1.1031	1.1489	1.1090	0.0064	
2	1.3650	1.2195	1.0810	1.1188	1.1761	1.3169	0.0065	
3	1.0924	0.9806	1.0670	1.1377				$o^{total}:$
$o_{weekday}^{aver}$	1.2287	1.1000	1.0740	1.1199	1.1625	1.2130	0.0064	6.9045
$\hat{c}_.$	1.2457	1.1153	1.0889	1.1354	1.1786	1.2297	0.0065	7.00

The parameters a and b can be estimated by means of *linear regression* (see Wood and Field (1976, pp. 76)). As Fig. 26.4 shows, appropriate estimators \hat{a} and \hat{b} are computed by minimizing the (squared) vertical distances between the deseasonalized sales $d_t = \frac{x_t}{\hat{c}_t}$ and the trend line $\hat{a} + \hat{b} \cdot t$. This useful way of eliminating the random noise is also applied in causal forecasts and has already been introduced in Sect. 7.2.2.

Linear Regression

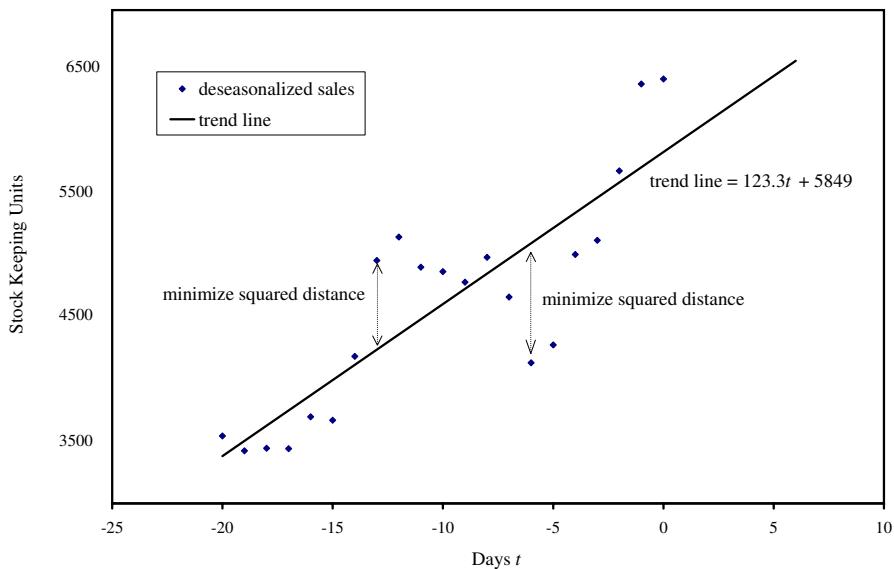


Fig. 26.4. Visualization of linear regression

Table 26.5 and Equations (26.6) and (26.7) illustrate how the trend parameters \hat{a}_0 and \hat{b}_0 have been calculated by linear regression to initialize Winters' method in Sect. 26.1.3:

$$\hat{b}_0 = \frac{\sum_t (t - \bar{t})(d_t - \bar{d})}{\sum_t (t - \bar{t})^2} = \frac{94943}{770} = 123.3 \quad (26.6)$$

$$\hat{a}_0 = \bar{d} - \hat{b}_0 \cdot \bar{t} = 4616 - 123.3 \cdot (-10) = 5849. \quad (26.7)$$

Here $\bar{t} = \frac{1}{21} \cdot \sum_t t = \frac{-210}{21} = -10$ and $\bar{d} = \frac{1}{21} \cdot \sum_t d_t = \frac{96936}{21} = 4616$ represent the average values of t and d_t over the first weeks of our working example.

Please note that similar deseasonalized sales have been obtained by the moving averages computation in the last subsection. These could also be used to estimate \hat{a} and \hat{b} by linear regression. In this case, however, only 15

Table 26.5. Calculation of linear regression

week	day	(corr.) x_t	\hat{c}_t	$d_t = \frac{x_t}{\hat{c}_t}$	$(t - \bar{t})^2$	$(t - \bar{t})(d_t - \bar{d}_t)$
1	-20	4419	1.2457	3547	100	10686
1	-19	3821	1.1153	3426	81	10709
1	-18	3754	1.0889	3448	64	9347
1	-17	3910	1.1354	3444	49	8206
1	-16	4363	1.1786	3702	36	5484
1	-15	4518	1.2297	3674	25	4710
1	-14	(27.3333̄)	0.0065	4192	16	1695
2	-13	(6190.4761)	1.2457	4970	9	-1060
2	-12	5755	1.1153	5160	4	-1088
2	-11	5352	1.0889	4915	1	-299
2	-10	5540	1.1354	4879	0	0
2	-9	5650	1.1786	4794	1	178
2	-8	6143	1.2297	4995	4	759
2	-7	(30.4666̄)	0.0065	4673	9	170
3	-6	5158	1.2457	4141	16	-1901
3	-5	4779	1.1153	4285	25	-1655
3	-4	5464	1.0889	5018	36	2413
3	-3	5828	1.1354	5133	49	3620
3	-2	6714	1.1786	5697	64	8646
3	-1	7872	1.2297	6401	81	16068
3	0	42	0.0065	6442	100	18256
\sum		-210		96936	770	94943

instead of 21 observations of deseasonalized sales would have been available, thus preparing a noticeably smaller sample to overcome randomness.

References

- Hanke, J. E.; Wichern, D. E.; Reitsch, A. G. (2001) *Business forecasting*, 7th ed., New Jersey
- Makridakis, S.; Wheelwright, S. C.; Hyndman, R. J. (1998) *Forecasting: Methods and applications*, 3rd ed., New York et al.
- Silver, E. A.; Pyke, D. F.; Peterson, R. (1998) *Inventory management and production planning and scheduling*, 3rd ed., New York et al.

- Waters, C. D. J. (1992) *Inventory control and management*, Chichester et al.
- Winters, P. (1960) *Forecasting sales by exponentially weighted moving averages*,
Management Science, Vol. 6, No. 3, 324–342
- Wood, D.; Field, R. (1976) *Forecasting for business: Methods and applications*,
London et al.

27 Linear and Mixed Integer Programming

Hartmut Stadtler

Darmstadt University of Technology, Department of Production & Supply Chain Management, Hochschulstraße 1, 64289 Darmstadt, Germany

Linear Programming (LP) is one of the most famous optimization techniques introduced independently by Kantarowitsch in 1939 and by Dantzig in 1949 (Krekó, 1973). LP is applicable in decision situations where quantities (variables) can take any real values only restricted by linear (in-) equalities, e. g. for representing capacity constraints. Still, LP has turned out to be very useful for many companies so far. LP is used in APS e. g. in Master Planning as well as in Distribution and Transport Planning. Very powerful solution algorithms have been developed (named solvers), solving LP models with thousands of variables and constraints within a few minutes on a personal computer.

In case some decisions can only be expressed by integer values, e. g. the number of additional shifts for a given week, LP usually will not provide a feasible solution. Similarly, logical implications might be modelled by binary variables. As an example consider the decision whether to setup a flow line for a certain product or not: A value of “0” will be attributed to a decision “no” and a value of “1” to “yes”. Still, the corresponding model may be described by linear (in-) equalities. In case the model solely consists of integer variables, it is called a pure *Integer Programming* (IP) model. If the model contains both real and integer variables a *Mixed Integer Programming* (MIP) model is given.

Thus, both LP and MIP comprise special model types and associated solution algorithms. Numerous articles and textbooks have been written on LP and MIP (e. g. Martin (1999), Winston (1994) and Wolsey (1998)) representing a high level of knowledge which cannot be reviewed here. In order to give an understanding of LP and MIP, only the basic ideas will be provided in the following by means of an example.

First, an LP model is presented and solved graphically (Sect. 27.1). This model is then converted into an IP model and solved by Branch and Bound (Sect. 27.2), where for each submodel a LP model is solved graphically. Finally, a few remarks and recommendations regarding the effective use of LP and MIP complements this chapter (Sect. 27.3).

27.1 Linear Programming

A hypothetical production planning problem is considered here, where two products A and B can be produced within the next month. The associated production amounts are represented by (real) variables x_1 and x_2 measured

in ten tons. Both products have to pass through the same production process. The available capacity is 20 days (on a two shift basis). The production of ten tons, or one unit, of product A lasts 5 days, while the respective coefficient for product B is 4 days. This situation is represented by inequality (27.2). LP model:

$$\text{Max!} \quad 19x_1 + 16x_2 \quad (27.1)$$

subject to

$$(1) \quad 5x_1 + 4x_2 \leq 20 \quad (27.2)$$

$$(2) \quad -x_1 + 2x_2 \leq 5 \quad (27.3)$$

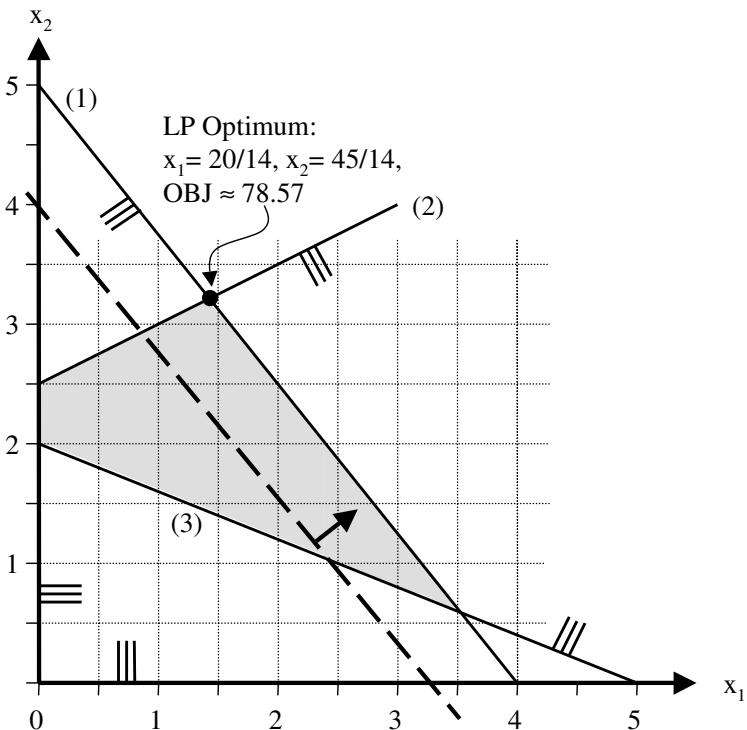
$$(3) \quad 2x_1 + 5x_2 \geq 10 \quad (27.4)$$

$$(\text{NNC}) \quad x_1 \geq 0, x_2 \geq 0 \quad (27.5)$$

Inequality (27.3) represents the demand constraints, stating that only sales of product B are limited. However, we might increase sales if we also offer product A: For every two units of product A we can extend sales of product B by one unit (the reason may be that one has to offer a complete product range to some customer groups in order to sell product B). Although we aim at maximizing our revenue (27.1), we also want to make sure that a contribution margin of at least ten thousand \\$ is reached within the next month (27.4). Note, the dimension “one thousand” is scaled down to “one” for the contribution margin constraint. Obviously, one cannot produce negative amounts which is reflected by the non-negativity constraints (NNC, see (27.5)).

This small LP model can be solved algebraically by the Simplex algorithm (or one of its variants, see Martin (1999)). However, we will resort to a graphical representation (Fig. 27.1). Variables x_1 and x_2 depict the two dimensions. Inequalities restrict the combination of feasible values of variables. The limits of the corresponding set of feasible solutions are illustrated by a line (see Fig. 27.1). Whether the set of feasible solutions lies below or above a line is depicted by three adjacent strokes being part of the set of feasible solutions.

The intersection of all the (in-) equalities of a model defines the set of feasible solutions (shaded area in Fig. 27.1). For a given objective function value the objective function itself is an equation (see dashed line in Fig. 27.1, corresponding to a value of 76 [\\$ 000]). Since we do not know the optimal value of the objective function we can try out several objective function values. An arrow shows the direction in which the objective function value can be increased. Actually, we can move the dashed line further to the right. The maximum is reached once it cannot be moved any further without leaving the set of feasible solutions. This is the case for $x_1 = 20/14$ and $x_2 = 45/14$ resulting in a revenue of 78.57 [\\$ 000]. The optimal solution has been reached at the intersection of inequalities (1) and (2). It can be shown that it suffices to look for an optimal solution only at the intersections of (in-) equalities



Explanations:

- region of feasible LP solutions

- — constraint relating to a given level of the objective function
- OBJ maximum objective function value
- grid (as a guidance for recognizing solutions)
- LP Optimum

Fig. 27.1. Graphical representation of an LP model

limiting the set of feasible solutions or, graphically speaking, at the “corners” of the shaded area.

The Simplex algorithm (and its variants) carries out the search for an optimal solution in two phases, namely

- creating an initial feasible solution and
- finding an optimal solution.

In our example a first feasible solution may be $x_1 = 0$ and $x_2 = 2$ with a revenue of $2 \cdot 16 = 32$ [\$'000]. Now, the second phase is started, probably generating an improved second solution, e.g. $x_1 = 0$ and $x_2 = 2.5$ with a

revenue of 40 [\$ 000]. In the next iteration variable x_1 will be introduced, resulting in the optimal LP solution.

However, an initial feasible solution may not always exist. As an example, assume that a minimum contribution margin of 22 [\$ 000] is required (see inequality (3') in Fig. 27.2). The set of feasible solutions is empty and thus no feasible solution exists.

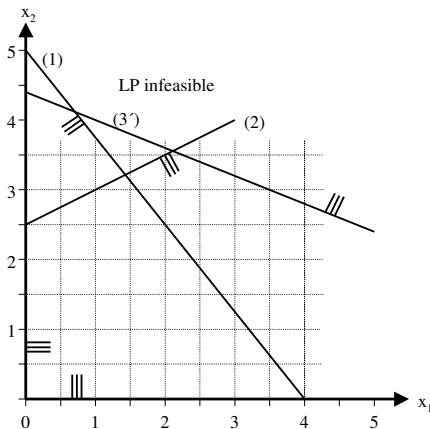


Fig. 27.2. An infeasible LP model

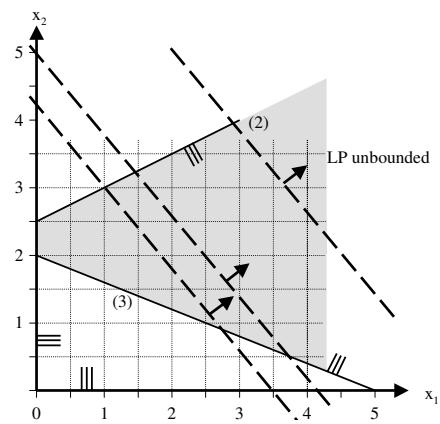


Fig. 27.3. An unbounded LP model

Now consider the situation where there is no production constraint (i.e. eliminating inequality (1)), resulting in an unrestricted shaded area (Fig. 27.3) and an unbounded objective function value. This case will also be detected in the first phase. Actually, an unbounded solution indicates that the model or the data have not been created correctly.

We would like to point out that an LP solution does not only provide optimal values for the decision variables. It also shows the *dual values* associated with the (in-) equalities of an LP model. As an example consider the production capacity (27.2). If we were able to increase the number of working days from 20 to 21, the optimal objective function value would rise from 1100/14 to 1154/14. Thus an additional capacity unit has a dual value of 3.86 [\$ 000]. Management now may look for options to extend capacity which are worth further revenues of 3.86 [\$ 000] per working day. Note that only inequalities which are binding in the optimal solution may have a positive dual value. Although dual values have to be interpreted with caution, they are a fruitful source for finding ways to improve the current decision situation.

As has already been stated at the beginning of this chapter, very powerful solution algorithms and respective standard software exist for solving LP models (e.g. CPLEX (ILOG CPLEX Division, 2002) and XPRESS-MP (DASH Associates, 2002)). However, users of an APS do not have to deal with these solvers directly. Instead, special modelling features have been selected

within APS modules for building correct models. Still, care should be taken regarding the numbers entering the model. If possible, appropriate scaling should be introduced first, such that the coefficients of variables are in the range from 0.01 to 100 to avoid numerical problems.

27.2 Pure Integer and Mixed Integer Programming

Now let us assume that a product can only be produced in integer multiples of ten [tons], since this is the size of a tub which has to be filled completely for producing either product A or B. Then the above model (27.1)–(27.4) has to be complemented by the additional constraints

$$x_1 \in \mathbb{N}_0, \quad x_2 \in \mathbb{N}_0 \quad (27.6)$$

The set of feasible solutions reduces drastically (see the five integer solutions in Fig. 27.4). Still, in practice the number of solutions to consider before an integer solution has been proven to be optimal may be enormous.

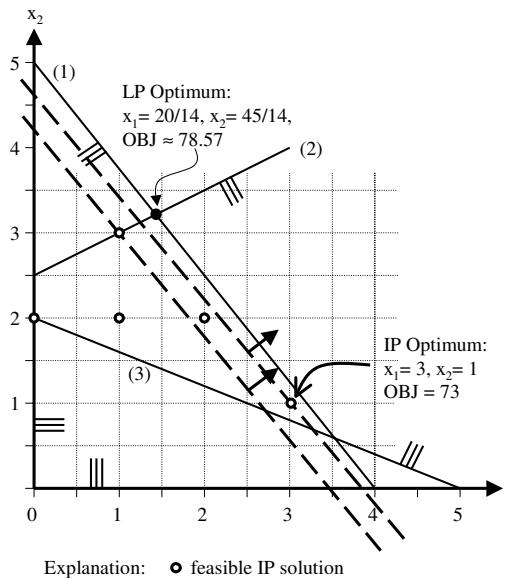


Fig. 27.4. A graphical representation of an IP model

As can be seen from Fig. 27.4 a straightforward idea, namely rounding the optimal LP solution to the next feasible integer values ($x_1 = 1, x_2 = 3$ with a revenue of 67 [\$000]), does not result in an optimal integer solution (which is $x_1 = 3, x_2 = 1$ and with a revenue of 73 [\$000]).

Anyway, an intelligent rounding heuristic might be appropriate for some applications. Hence, some APS incorporate rounding heuristics which usually

require much less computational efforts than Branch and Bound which is explained next.

Four building blocks have to be considered describing a *Branch and Bound* algorithm, namely

- relaxation,
- separation rules,
- search strategy,
- fathoming rules.

The two building blocks *separation rules* and *search strategy* relate to “branch” while *relaxation* and *fathoming rules* concern “bound”. These building blocks will now be explained by solving our example.

Although solving the associated LP model directly usually does not yield an optimal integer solution, we can conclude that the set of feasible integer solutions is a subset of the set of feasible LP solutions. So, if we were able to cut off some parts of the non-integer solution space, then we would finally arrive at an integer solution.

Consequently, we first *relax* the integer requirements (27.6) in favour of the non-negativity constraints (27.5). The resultant model is called an *LP relaxation*. If we solve an LP relaxation of a maximization problem, the optimal objective function will be an *upper bound* for all integer solutions contained in the associated set of feasible (integer) solutions. Hence, if the solution of an LP relaxation fulfills the integer requirements (27.6), it will be an optimal integer solution for this (sub-) model.

Next, submodels are created by introducing additional constraints, such that a portion of the real-valued non-integer solution space is eliminated (see Fig. 27.5). Here, the constraint $x_1 \leq 1$ is added resulting in submodel SM¹, while constraint $x_1 \geq 2$ yields submodel SM². Now, we have to solve two submodels with a reduced set of feasible solutions. Note that the union of the set of feasible *integer* solutions of both submodels matches the initial set of feasible integer solutions, i.e. no integer solution is lost by *separation*.

Submodel SM¹ results in a first *integer* solution ($x_1 = 1$, $x_2 = 3$ with a revenue of 67 [\$000] representing the local upper bound of SM¹). Subsequently, we will only be interested in solutions with a revenue of more than 67 [\$000]. Thus, we set the global lower bound to 67 [\$000] (OBJ = 67). The term “global” is used in order to refer to our original IP model. Since submodel SM¹ has resulted in an *integer solution* (and cannot yield a better solution) it will be discarded from our list of open submodels, i.e. it is *fathomed*.

The second submodel has a local upper bound of 78 [\$000] which is clearly better than our current global lower bound, but its solution is non-integer valued ($x_2 = 2.5$).

The search for an optimal solution can be represented by a *search tree* (see right hand side of Fig. 27.5). Each node corresponds to an LP (sub-)model.

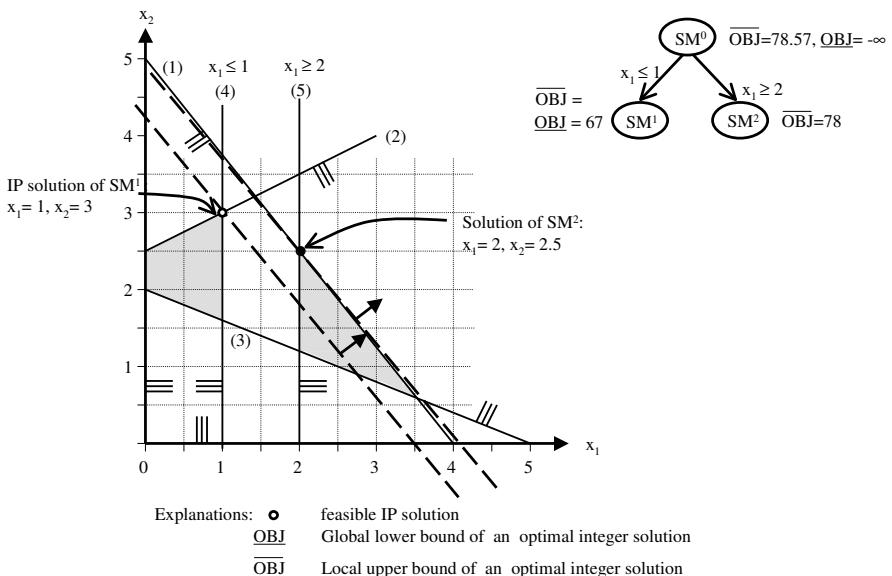


Fig. 27.5. A graphical representation of the first and second submodel

Now an unfathomed submodel has to be chosen for further investigations. However, only submodel SM^2 is unfathomed here. Subsequently, one has to decide on the non-integer-valued variable to branch on. These two choices make up the *search strategy* and may have a great impact on the number of submodels to solve and hence the computational effort.

The only variable which is non-integer valued in the optimal solution for submodel SM^2 is x_2 . Two new submodels are created, submodel SM^3 with the additional constraint $x_2 \leq 2$ and submodel SM^4 with the additional constraint $x_2 \geq 3$. Note that all additional constraints that have been generated on the path from the origin (SM^0) to a given submodel in the search tree have to be taken into account (here $x_1 \geq 2$).

Since, there is no feasible (real valued) solution for submodel SM^4 (see Fig. 27.6) it may be fathomed. For submodel SM^3 a non-integer valued solution with an upper bound of 77.6 [\$000] is calculated. Since this local upper bound exceeds the global lower bound (i. e. the best objective function value known) submodel SM^3 must not be fathomed.

It now takes three further separations until we reach submodel SM^9 (Fig. 27.7), where the LP relaxation yields an integer solution with an objective function value of 73 [\$000].

Usually, there will be some unfathomed submodels which have been generated in the course of the search. An unfathomed submodel has to be selected for a further separation until all submodels are fathomed. Then the best feasible integer solution found will be the optimal one for the initial IP model.

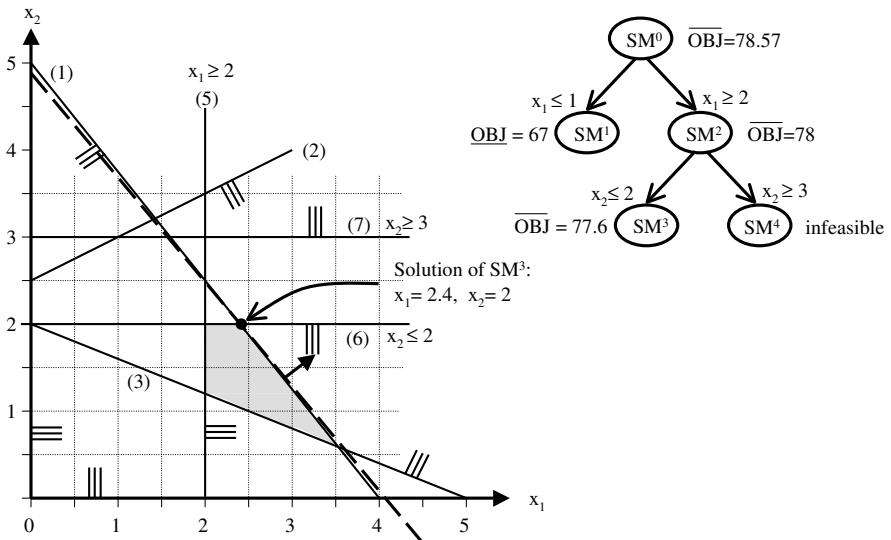


Fig. 27.6. A graphical representation of the third and fourth submodel

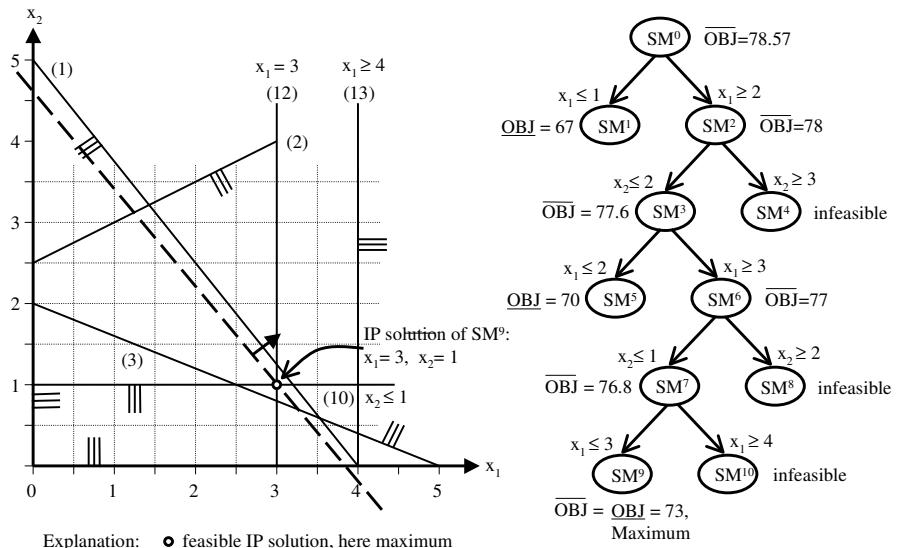


Fig. 27.7. A graphical representation of the optimal integer solution and the complete search tree

In our example, the search ends once it has been found out that submodel SM^{10} has no feasible solution. Now we have proven that the solution to submodel SM^9 is optimal.

Finally, we would like to add that the Branch and Bound scheme is almost the same for MIP models. As an example consider that only x_2 has to take integer values. Then we would start separating on variable x_2 (i.e. $x_2 \leq 3$ and $x_2 \geq 4$). Only constraint $x_2 \leq 3$ results in a feasible solution for the LP relaxation. Since it is also feasible with respect to the mixed integer constraints it is the optimal solution, too.

27.3 Remarks and Recommendations

Although the examples presented are rather simple, they have illustrated the differences in solving an LP model and a MIP model. Generating an optimal solution for an LP model requires “some” Simplex iterations leading from one “corner” of the feasible solution space to the next and finally to the optimal one. However, solving a MIP model by Branch and Bound incurs solving an LP (sub-) model for each node of the search tree – and there may be several thousand nodes to explore until an optimal solution has been proven.

One way to reduce the number of submodels to investigate is to *truncate the search effort*. For example, the user may either set a certain time limit for the search or indicate that the search has to be stopped once the k-th feasible integer solution has been found. However, the problem with truncation is that one does not know in advance at which point in time a feasible or good solution will be found.

Another option to limit the computational effort of Branch and Bound is to specify in advance that the search for an improved solution should be stopped, once we are sure that there is no feasible integer solution which is at least $\delta\%$ better than our current best solution. This allows us to calculate an aspiration level in the course of Branch and Bound, simply by multiplying the objective function value of the current best solution by $(1 + \delta\%)$. The question whether there exists a feasible integer solution with an objective function value no less than the aspiration level is known from the maximum upper bound of *all* unfathomed submodels. If the maximum is less than our aspiration level the search is stopped.

In our example (see the search tree in Fig. 27.7) we now assume $\delta = 10$. Having generated the first integer solution (OBJ = 67) the aspiration level is 73.7 [\$000]. Since the maximum of the upper bounds of unfathomed submodels is 78 [\$000] (submodel 2) the search will continue. Having reached the second integer solution with an objective function value of 70 [\$000], an aspiration level of 77 [\$000] is calculated. In this example the search stops once the maximum upper bound of all unfathomed submodels falls below 77 [\$000] which is true after having generating submodel 8.

The number of submodels to solve largely depends on the relative difference between the objective function value of the LP relaxation and the optimal integer solution, named *integrality gap*. For our example the integrality gap is rather modest (e.g. $(78.57 - 73)/73 = 0.076$ or 7.6%). The smaller

the integrality gap is, the greater is the chance to fathom submodels and thus to keep the search tree small. Today much effort is invested in deriving additional *valid inequalities (cuts)* to yield small integrality gaps for each submodel generated within Branch and Bound (see Wolsey (1998)).

A further option applied by advanced MIP solvers to reduce the search effort is *preprocessing*. Here one investigates the interactions of the model's constraints in order to restrict or even fix the values of some integer variables before starting Branch and Bound. For our example, one might conclude that the set of feasible integer values for x_1 will be restricted to $\{0,1,2,3\}$ and to $\{1, 2, 3\}$ for x_2 . Preprocessing is very similar to the ideas of Constraint Programming (Chap. 24).

A frequently asked question is: "Will our Master Planning model be solvable within reasonable CPU-times?" Before answering this question one has to differentiate whether the Master Planning model is to be solved by an LP or a MIP solver or a simple heuristic.

As already stated, purely linear models are much easier to solve than MIP models. Actually, solution capabilities of state-of-the-art LP solvers should be sufficient for solving almost all reasonable real word applications. However, if elapsed time plays a role a few experiments at an early stage of a project should clarify matters: The idea is to generate an LP model with only a subset J^r of all products J and/or with a reduced number of time periods T^r compared with T periods in the final model ($T^r \ll T$), but representing the same model structure, i.e. containing all types of constraints of the final model. Assuming that the reduced model requires a CPU-time CPU^r , a rule-of-thumb for calculating the CPU-time (CPU) of the final model is:

$$CPU \sim \left(\frac{T}{T^r} \cdot \frac{|J|}{|J^r|} \right)^3 \cdot CPU^r \quad (27.7)$$

This rule-of-thumb is derived from the observation that the computational time required for solving an LP by a Simplex method tends to be roughly proportional to the cube of the number of explicit constraints, so that doubling this number may multiply the computational time by a factor of approximately 8 (Hillier and Liebermann, 2001, p. 161).

For MIP models an optimal solution usually cannot be expected within reasonable CPU times. Hence the search for an optimal solution is truncated (see above). Also, remember that always a (relaxed) LP model has to be solved before the search for a MIP solution starts. In any case the CPU time limit must be sufficient for at least generating a first feasible MIP solution.

Again preliminary experiments with the MIP model can provide valuable insights. One approach is to start with relaxing all integer requirements, resulting in a purely linear model. If this model is solved easily, then the most important integer requirements can be introduced and the associated computational effort observed. Then further variables may be declared integer

until a good compromise between the solution effort, solution quality and model adequacy is reached.

A second related approach is to specify a complete MIP model including all desired integer requirements but to relax some integer requirements for variables in later periods in the planning horizon, where e.g. only a rough capacity check suffices.

A third approach supported by some software vendors is to use time or stage oriented decomposition. If this option is chosen, the overall model is partitioned into smaller submodels (automatically) which are then solved successively (e.g. a MIP model covering 13 periods is partitioned into 4 MIP models with 4 periods each, while there is one overlapping period). In the end the user will get a complete solution for the original decision problem.

In any case the user should use integer or binary variables carefully – a MIP model incorporating (only) one hundred integer variables may already turn out to require excessive computational efforts.

References

- DASH Associates (2002) *Homepage*, URL: <http://www.dash.co.uk>, Date: July 19, 2004
- Hillier, F. S.; Liebermann, G. J. (2001) *Introduction to Operations Research*, 7th ed., New York et al.
- ILOG CPLEX Division (2002) *Homepage*, URL: <http://www.ilog.com/products/cplex>, Date: July 19, 2004
- Krekó, B. (1973) *Lehrbuch der Linearen Optimierung*, Berlin
- Martin, R. K. (1999) *Large scale linear and integer optimization: A unified approach*, Boston et al.
- Winston, W. L. (1994) *Operations Research: Applications and algorithms*, 3rd ed., Belmont, California
- Wolsey, L. A. (1998) *Integer Programming*, New York et al.

28 Genetic Algorithms

Robert Klein

Darmstadt University of Technology, Department of Operations Research,
Hochschulstraße 1, 64289 Darmstadt, Germany

28.1 General Idea

Many optimization problems of the type arising in scheduling and routing (see Chaps. 10 and 12) are of combinatorial nature, i. e. solutions are obtained by combining and sequencing solution elements. When solving such problems to optimality, the number of solutions to be examined exponentially grows with the problem size. For example, for n solution elements $n!$ different sequences exist.

Recently, *genetic algorithms* (GA) have become increasingly popular as a means for solving such optimization problems heuristically, i.e. for determining near-optimal solutions within reasonable time. One of the main reasons for this popularity is the relative ease of programming at least a simple genetic algorithm. Furthermore, many researchers have observed empirically that already basic versions of GA will give very acceptable results without excessively fine-tuning them for the problem on hand. Finally, since GA work on a representation (coding) of a problem (see Sect. 28.2), it is possible to adapt existing procedures to modified problem versions quite easily or to write one general computer programme for solving many different problems. GA were initially developed by Holland and his associates at the University of Michigan and the first systematic but rather technical treatment was published in Holland (1975). For comprehensive descriptions from a more practical point of view, we refer to Goldberg (1989), Reeves (1993) and Michalewicz (2000). Surveys on successful applications of GA for solving combinatorial optimization problems are, for example, given in Dowsland (1996) and Reeves (1997).

According to the biological evolution, GA work with *populations of individuals* which represent feasible solutions for the problem considered. The populations are constructed iteratively through a number of *generations*. Following the idea of Darwinism (“survival of the fittest”), each individual of the current generation “contributes” to the subsequent one according to its quality which is measured by a *fitness value*. This is achieved by *selecting* individuals randomly with the probability of choosing a certain individual depending on its fitness value. In order to obtain the next generation from the individuals selected, two basic operations exist. Using a *crossover*, the features of two (parent) individuals are *recombined* to one or more new (child) ones. By *mutation*, some features of an individual are modified randomly. A template for a single iteration of a genetic algorithm is depicted in Fig. 28.1.

Usually, GA are executed until a prespecified stopping criterion is fulfilled, e.g. a certain number of generations has been evaluated or a time limit is reached.

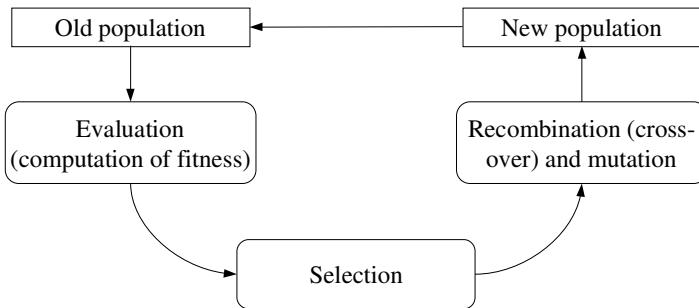


Fig. 28.1. Template for a single iteration of a genetic algorithm

In the following, we discuss the different aspects of GA in more detail. To ease presentation, the following *production scheduling problem* is considered. A number n of jobs has to be processed on a single machine (with simultaneous execution being impossible). Each job $j = 1, \dots, n$ has a fixed processing time (duration) of d_j periods and preemption is not allowed. Furthermore, job j cannot be started before its release date rd_j and should be terminated until a due date dd_j . In case it is finished later than dd_j , a penalty cost c_j for each time unit of tardiness arises. Hence, the problem consists of finding a schedule, i.e. a starting time s_j for each job, such that the total tardiness costs are minimized. The data of an example with $n = 8$ jobs are given in Table 28.1.

Table 28.1. Data of an example problem

j	1	2	3	4	5	6	7	8
d_j	5	4	7	8	3	2	6	4
rd_j	0	2	4	16	18	28	25	28
dd_j	17	10	13	28	22	31	36	36
c_j	3	2	5	3	4	1	3	4

28.2 Populations and Individuals

As stated before, a population consists of a set of individuals. Each individual is *represented* by a vector (*string*) of fixed length in which the corresponding solution is coded by assigning specific values to the vector elements (*string*

positions). In order to obtain the solution associated with an individual, the respective string has to be *decoded*. Both the dimension (length) of the string as well as the domains (sets of feasible values) of the string positions depend on which representation is chosen for coding the solution.

In our example, a solution can be represented by a sequence S of jobs. That is, the string consists of n positions and each position can take one of the values $1, \dots, n$ (with all positions having different values). For decoding the string, we proceed as follows. The jobs are considered in accordance to the sequence S . The job j in turn is started at the smallest possible point in time $s_j \geq rd_j$ at which its execution does not overlap with a job already scheduled. After having scheduled all jobs, the total tardiness costs can be computed. Consider the string $S = \langle 1, 2, 3, 4, 5, 6, 7, 8 \rangle$ for our example. By decoding this sequence, the solution shown in the gantt-chart of Fig. 28.2 is obtained. The numbers within the bars denote the job numbers and the tardiness of the jobs, respectively. The lengths of the bars correspond to the processing times.

Job 1 can be started at the earliest. Scheduling job 2 results in $s_2 = 5$ due to the processing of job 1 which does not allow for a smaller starting time. After terminating job 2, job 3 can begin at $s_3 = 9$, hence, finishing three periods after its due date $dd_3 = 13$. The jobs 4 and 5 are scheduled subsequently. Job 6 cannot be launched earlier than $s_6 = rd_6 = 28$. Finally, the jobs 7 and 8 are considered with the latter terminating after 40 periods. The total tardiness costs are $3 \cdot 5 + 5 \cdot 4 + 4 \cdot 4 = 51$.

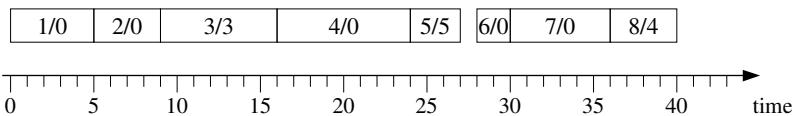


Fig. 28.2. Gantt-chart for $S = \langle 1, 2, 3, 4, 5, 6, 7, 8 \rangle$

Note that we have chosen the above representation, because it is well suited for a large number of scheduling and routing problems and, hence, is used by a large number of GA for such problems (Reeves, 1997). Most commonly, GA are described using a representation where problems are coded in a bitwise fashion, i. e. each string position can take either the value 0 or 1. However, for most combinatorial optimization problems this representation is not appropriate.

Choosing an efficient representation for coding solutions is important for the performance of the genetic algorithm to be developed. The representation should be designed such that both the decoding does not require too much computational effort as well as the operations crossover and mutation can be performed efficiently. Within our example, an alternative representation of solutions could consist of using a string with n elements, where each position $j = 1, \dots, n$ defines a priority value for job j . This representation can be

decoded in two steps. First of all, a job sequence S is obtained by sorting the jobs according to, e. g. non-decreasing priority values. Subsequently, a feasible schedule can be constructed as described above for the sequence based representation. Obviously, the sorting step results in an additional effort which is not justified unless the representation has some other advantages, e. g. it would allow for more efficient crossover and mutation operators. When using a representation where each string position denotes the starting time s_j of a job j , the implementation of the crossover and mutation operations becomes difficult. As already stated earlier, a mutation consists of modifying a string randomly (see Sect. 28.4 for details). Simply changing the starting time of a single job randomly may result in an infeasible schedule due to processing jobs in parallel. The problem of infeasibility is even more difficult to resolve when recombining individuals.

In any application of GA, an important question consists of choosing an appropriate *population size* P , i. e. the number of individuals considered in each iteration. If the population size is too small, the search space of feasible solutions may only be evaluated partially, because just a few existing individuals are recombined in each iteration and these individuals increasingly resemble each other with each additional generation. Otherwise, in case of a too large population size, also rather poor individuals may be considered for recombination. This is in particular disadvantageous in case that for each new string large parts of the corresponding solution have to be reconstructed as in our example, which requires a considerable computational effort. Hence, with an increasing problem size, most of the computational time will be spent for constructing solutions rather than examining the search space and the search will only proceed slowly towards high-quality solutions. In the literature, most successful applications of GA propose an even-numbered population size of $P \in [50, 100]$ (Reeves, 1997).

Finally, an *initial population* has to be determined before starting a genetic algorithm. Most commonly, the corresponding individuals are obtained by randomly assigning values to the string positions. In our example, sequences of jobs may be constructed randomly. Alternatively, simple heuristics, such as randomized priority-rule based approaches (e. g. Drexel (1991)), may be applied in order to start the search with promising solutions.

28.3 Evaluation and Selection of Individuals

As stated previously, individuals contribute to the next generation with a probability depending on their fitness value. For this purpose, a *gene pool* consisting of P copies of individuals is constructed. For those individuals with a high fitness value, several copies are included in the pool, i. e. the individuals are selected several times, whereas for those with low values no copy may be contained at all. This reflects the analogy to biological evolution. The best individuals should contribute to the next generation the most often,

i. e. their positive features are reproduced in many of the new individuals. By way of contrast, the worst ones with a low selection probability should be discarded and, hence, “die off”.

In the most simple form, determining the fitness values v_i for the individuals $i = 1, \dots, P$ consists in computing the objective function values f_i of the corresponding solutions.

For maximization problems, the selection process often used within GA can be subdivided in the following two steps. In the first step, a roulette wheel with $i = 1, \dots, P$ slots sized according to the fitness values $v_i = f_i$ is constructed. For this purpose, the total fitness of the population is computed by $T = \sum_{i=1}^P v_i$. Subsequently, each individual i is assigned a selection probability of $p_i = v_i/T$ as well as a cumulative probability $q_i = \sum_{h=1}^i p_i$. In the second step, the roulette wheel is spun P times. In each iteration, a single individual is selected, i. e. a copy is included in the gene pool, as follows. After generating a random float number $\beta \in [0, 1]$, the individual $i = 1$ is chosen in case of $\beta \leq q_1$. Otherwise, the i -th individual with $q_{i-1} < \beta \leq q_i$ is picked.

The above selection process bears the difficulty that if the objective is minimization instead of maximization as in our example, a transformation of the objective function values has to be performed. One simple transformation consists of defining an upper bound F which exceeds all possible objective function values and subsequently using the fitness value $v_i = F - f_i$. Another difficulty is that the scale on which the values are measured may not be considered appropriately. For example, values of 1,020 and 1,040 are less distinctive than values of 20 and 40.

Therefore, two possible alternatives for designing the selection process have been proposed in the literature. When using a *ranking* approach, the individuals are ordered according to non-deteriorating fitness values with r_i denoting the rank of individual i . Subsequently, a selection probability is computed by, e.g. $p_i = 2r_i/(P \cdot (P + 1))$. In this case, the best individual with $r_i = P$ has the chance of $p_i = 2/(P+1)$ of being selected. This is roughly twice of that of the median whose chance is $p_i = 1/P$. With the values p_i on hand, the selection can be performed by spinning the roulette wheel as described above.

The other possibility is the *tournament selection*. In this approach, a list of individuals is obtained by randomly permuting their index numbers $i = 1, \dots, P$. Afterwards, successive groups of L individuals are taken from the list. Among these individuals, the one with the best objective function value is chosen for reproduction and a copy is added to the gene pool. Then, the process is continued with the next L individuals until the list is exhausted or the gene pool contains P copies, whatever comes first. In the first case, the tournament process is continued to determine the missing members of the gene pool after determining a new list randomly.

Except for the tournament selection, the above approaches have in common that there is no guarantee that the best of all individuals is selected

for reproduction. From the optimization perspective this may not be efficient. Therefore, the concept of *elitism* has been introduced which consists of putting a copy of the best individual into the gene pool by default and applying the roulette wheel and ranking approaches only $P - 1$ times. In generalized versions, a larger number of individuals is chosen by default.

28.4 Recombination and Mutation

For the *recombination* process, a pair of individuals is chosen from the gene pool either randomly or systematically. A *crossover* is carried out with a certain probability γ , i. e. the pair is recombined into two new individuals. In case that no recombination is performed, the original individuals become part of the new population with a probability of $1 - \gamma$. This process is repeated until P individuals have been considered and, hence, a new population with P individuals has been obtained. In the literature, different values for γ have been proposed with values of $\gamma < 0.6$ not being efficient (Reeves, 1997).

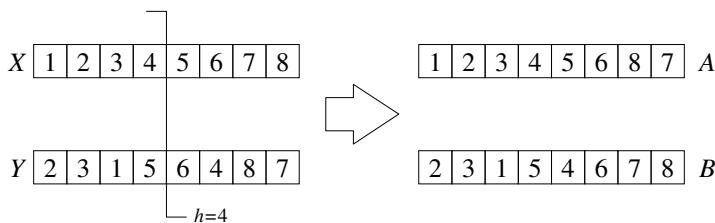


Fig. 28.3. 1-Point crossover for sequence representations

In the following, we describe the simple *1-point crossover* which is the one most commonly used. In general, it is defined for strings with length n as follows. For each pair of parent individuals X and Y , a *crossover point* $h \in [1, n - 1]$ is determined randomly. Afterwards, a first individual is obtained by concatenating the first h string positions of X with the $n - h$ last positions of Y . The second individual is obtained just the other way round. Unfortunately, this definition does not work for every possible representation of solutions. For our example problem, such a crossover results in individuals with feasible solutions when the representation based on priority values is applied but fails for the representation relying on sequences. In the latter case, it yields individuals with some jobs occurring twice and others being discarded.

Therefore, a different approach is used for sequence based representations, the principle of which is depicted for two possible strings of our example problem (Fig. 28.3). After selecting a crossover point $h \in [1, n - 1]$ randomly, the first h string positions of the parent individual X are copied into the child one A . Subsequently, the remaining $n - h$ positions are filled up with those

elements which have not been considered yet in the order in which they are contained in individual Y . The second child B is constructed accordingly now starting with the first h positions of individual Y . Note that in our example both individuals obtained by the crossover yield better objective function values than their parent ones. The total tardiness costs of the schedules represented by X and Y are 51 and 93, whereas for A and B we obtain 47 and 29, respectively.

In addition to recombination, *mutation* is applied for some of the new individuals to diversify the search, i. e. to avoid that the same set of solutions is examined repeatedly through a number of consecutive generations. For this purpose, a probability δ with which each individual is mutated has to be specified. According to the selection process, the decision whether to mutate an individual or not can be made by randomly generating a number from the interval $[0, 1]$. The usual approaches for determining δ are either to choose a very small value, e. g. $\delta = 0.01$ or to use a value $\delta = 1/n$, because there is some theoretical and practical evidence that this is a reasonable value for many problems (Reeves, 1997). In general, mutation consists of randomly altering the value at a random string position. In order to preserve the feasibility for sequence based representations, two more versatile mutation possibilities are distinguished. Within an *exchange* mutation, two string positions are randomly selected and the corresponding elements are interchanged. In our example, the positions three and six may be selected for individual A resulting in the mutated sequence $A' = \langle 1, 2, 6, 4, 5, 3, 8, 7 \rangle$. A *shift* mutation consists of randomly choosing a single string position and moving the corresponding element by a random number of positions to the left or right. After selecting position six of individual B and left shifting the element by three positions, we yield $B' = \langle 2, 3, 6, 1, 5, 4, 7, 8 \rangle$.

28.5 Conclusions

The previous expositions aim at reviewing the basic ideas of GA in the context of solving combinatorial optimization problems. They also show that a large variety of design possibilities exist when implementing GA for particular problems. This includes choosing a representation of solutions, a selection mechanism as well as efficient recombination and mutation strategies. Fortunately, as stated at the beginning, already basic versions of GA are robust in the sense that they are able to yield satisfying results for many problems. However, most computational experiments presented in the literature so far consider only small and medium sized problem instances. For example in the context of scheduling, the largest instances evaluated usually do not consist of more than one hundred jobs. Since for n jobs up to $n!$ job sequences may exist, the search space explodes with an increasing job number and, hence, a very large number of generations may have to be evaluated in order to obtain near optimal solutions.

Furthermore, depending on the problem to be solved, it may be difficult to consider constraints appropriately, i. e., to avoid that infeasible solutions are obtained throughout the solution process. Within production scheduling, such constraints may be due to generalized precedence relationships among jobs or to time windows for their execution. In order to overcome such difficulties, several concepts have been developed. The most common one is to modify the objective function by a penalty term such that infeasible solutions are assigned a low fitness value. For more comprehensive discussions we refer to Reeves (1997).

References

- Dowsland, K. A. (1996) *Genetic algorithms – A tool for OR?*, Journal of the Operational Research Society, Vol. 47, 550–561
- Drexel, A. (1991) *Scheduling of project networks by job assignment*, Management Science, Vol. 37, 1590–1602
- Goldberg, D. E. (1989) *Genetic algorithms in search, optimization, and machine learning*, Reading/Massachusetts et al.
- Holland, J. H. (1975) *Adaptation in natural and artificial intelligence*, Ann Arbor
- Michalewicz, Z. (2000) *Genetic algorithms + data structures = Evolution programs*, 3rd ed., Berlin et al.
- Reeves, C. R. (Ed.) (1993) *Modern heuristic techniques for combinatorial problems*, Oxford et al.
- Reeves, C. R. (1997) *Genetic algorithms for the operations researcher*, INFORMS Journal on Computing, Vol. 9, 231–250

29 Constraint Programming

Robert Klein

Darmstadt University of Technology, Department of Operations Research,
Hochschulstraße 1, 64289 Darmstadt, Germany

29.1 Overview and General Idea

Constraint programming (CP) represents a relatively new technique for computing feasible (and optimal) solutions to combinatorial decision problems like those typically arising in scheduling and routing (see Chaps. 10 and 12). In the mid eighties, it was developed as a computer science technique by combining concepts of *Artificial Intelligence* with new programming languages. In the meantime, it has received considerable attention in practice as well as in the *Operations Research* (OR) community, in particular, since it has successfully been included into commercial software systems (e.g. ILOG OPL Studio). The basic idea of CP consists of providing an integrated framework for formulating and solving decision problems based on a single programming language. For the latter purpose, generalized solution procedures are included within CP systems, the application of which can be controlled by the user. Hence, in contrast to classical OR techniques such as mixed integer programming, the user of CP does not only specify the decision problem to be solved but also determines how the search for corresponding feasible solutions should be performed.

For solving decision problems, CP does not rely on mathematical optimization problems but on constraint satisfaction problems which, basically, consist of variables, domains as well as constraints (Sect. 29.2). For each variable, an associated *domain* defines a set of feasible values which need not necessarily be real or integral. The variables are related to each other by *constraints* describing restrictions that have to be observed by feasible solutions. In general, constraints need not be simple expressions, in particular, they need not be linear inequalities or equations as common in mixed integer programming. Possible examples involving two variables x_1 and x_2 are $x_1 \neq x_2$, $x_1 \cdot x_2 < 10$ or $x_1 > 3 \Rightarrow x_2 > 7$. The corresponding *constraint satisfaction problem* (CSP) is to assign a value to each variable of its domain such that each constraint is satisfied. Obviously, a CSP differs from a classical optimization problem by not considering an objective function. That is, solving a CSP aims at finding feasible solutions to a real-world problem rather than an optimal one. However, a possible objective function can be represented within CSPs by particular constraints and optimized by solving several CSPs consecutively.

In order to compute feasible solutions to a CSP by CP, constraint propagation techniques are usually combined with special purpose search algorithms.

Constraint propagation provides an effective mechanism to systematically reduce the domains of variables by carefully analyzing the constraints for the problem data on hand and resolving inconsistencies (Sect. 29.3). The *search algorithms* applied within CP are based on systematically enumerating all feasible solutions of a CSP (with reduced domains) using a backtracking approach (Sect. 29.4).

Among the major advantages of CP are the ease of application and the flexibility to add new constraints to existing problems. This is due to the rich set of possible constraint types and to the search algorithms employed being rather general. A disadvantage may be the rather poor performance with respect to solution quality and computation time (Brailsford et al., 1999).

In the following, we describe the different components of CP in more detail. To ease the presentation, we use the example and the notation of the production scheduling problem introduced in Chap. 28. For recent introductions and surveys on constraint programming see, e.g., Brailsford et al. (1999) as well as Lustig and Puget (2001).

29.2 Constraint Satisfaction Problems

As stated in the previous section, a CSP considers a set of n *variables* x_1, x_2, \dots, x_n . Associated with each variable $x_j (j = 1, \dots, n)$ is a finite *domain* (set) D_j of possible values. In our production scheduling example (see Chap. 28), the variables x_j may denote the start times of the jobs $j = 1, \dots, n$. Assuming that all jobs have to be terminated until a common deadline of T , i.e. T periods after the beginning of the planning horizon, the domain of a job j is $D_j = \{rd_j, \dots, T - d_j\}$, because it cannot begin before its release date rd_j and has to start d_j periods before T at the latest. Note that in general the values of D_j need not be a set of consecutive integers. Furthermore, they even need not be numeric, e.g. they can correspond to elements of some general set. In case domains are not finite like in Linear Programming problems, the solution techniques described in the following sections have to be modified.

The variables are related by a set of *constraints*. Formally, a constraint C_{ij} between two variables x_i and x_j corresponds to a feasible subset of all possible combinations of the values of x_i and x_j , i.e. $C_{ij} \subseteq D_i \times D_j$. If $(x_i, x_j) \in C_{ij}$, the constraint is said to be *satisfied*. For example, if $D_1 = \{1, 2\}$ and $D_2 = \{3, 4\}$, then the constraint $x_1 + 2 = x_2$ is equivalent to the subset $\{(1, 3), (2, 4)\}$ of the set of all possible combinations $\{(1, 3), (1, 4), (2, 3), (2, 4)\}$. For constraints referring to a larger number of variables, the definition can easily be extended.

In practice, the programming languages of CP systems provide more efficient approaches for representing constraints. For our production scheduling problem we obtain the following CP formulation when omitting the objective function:

$$x_i + d_i \leq x_j \text{ or } x_j + d_j \leq x_i \quad \text{for } i = 1, \dots, n, j = 1, \dots, n,$$

$$\text{and } i < j \quad (29.1)$$

$$D_j = \{rd_j, \dots, T - d_j\} \text{ for } j = 1, \dots, n \quad (29.2)$$

The first set of constraints is commonly called *disjunctive* constraints. It says that for two jobs i and j , either job i must finish before job j starts or job j must terminate before job i begins. This type of constraints plays an important role in many production scheduling problems, where two jobs are not allowed to be processed simultaneously on a single machine, as this is e.g. the case in our example or in a flow-shop or job-shop environment. Note that such a straightforward formulation of disjunctive constraints as in (29.1) is not possible in a mixed integer programme where binary variables have to be introduced for this purpose. This is also true for a large number of further types of constraints (Williams and Wilson, 1998). Another typical example of a constraint which can easily be defined within CP but is difficult to express in a mixed integer programme is the following. Each of the five variables x_1, x_2, \dots, x_5 is to be assigned a different value from the interval $[1, \dots, 5]$.

Obviously, a *feasible solution* to a CSP is an assignment of a value to each variable from its domain such that all constraints are satisfied. Basically, we may be interested in computing just one or all feasible solutions of a CSP with no preference as to which one. In case an optimal (e.g. a minimal) or at least a good solution for some objective function has to be determined, several CSPs have to be solved consecutively. For this purpose, an objective variable is additionally defined which corresponds to the objective function value. After finding a first feasible solution, a modified CSP is obtained by introducing a new (objective) constraint which specifies that the value of the objective variable has to be smaller than in the initial solution. That is, an upper bound on the objective function value is established such that only solutions with smaller values are considered feasible when solving the modified CSP. This process is continued by tightening the upper bound each time a new feasible solution has been determined until a CSP is obtained for which no feasible solution exists. Then, the last solution found represents a minimal one. In case that the process is terminated prematurely, e.g. due to limited computation time, only a heuristic solution is determined.

For our example, we define $y = \sum_{j=1}^n c_j \cdot \max\{x_j + d_j - dd_j, 0\}$ as objective variable. When the solution depicted in Fig. 28.2 with total tardiness cost of 51 is to be improved, a CSP consisting of the constraints (29.1), (29.2) and $y < 51$ has to be solved.

29.3 Constraint Propagation

The basic idea of *constraint propagation* is to “propagate” the effects of modifying a variable’s domain to any constraint that interacts with that variable. By analyzing each of these constraints, possible inconsistencies resulting from the modification are discovered and subsequently resolved by removing in-

consistent values from the domains of the remaining variables participating in the affected constraint. This step is usually referred to as *domain reduction*.

In the following, we describe the principle of domain reduction for constraints which only concern two variables, such as the disjunctive constraints discussed in the previous section. In this case, the variables and the constraints can be depicted in a constraint graph with the nodes representing variables. Arcs are introduced between two nodes, if a constraint C_{ij} is defined between the corresponding variables x_i and x_j . Furthermore, the arc (x_i, x_j) is called (*arc*) *consistent* if for every value $a \in D_i$, there is a value $b \in D_j$ such that the assignments $x_i = a$ and $x_j = b$ do not violate the constraint C_{ij} . Any value $a \in D_i$ for which this is not true, i. e. no corresponding value b exists can be removed from D_i , because it cannot be contained in any feasible solution. By treating all such values from D_i accordingly, consistency for the arc (x_i, x_j) is obtained. This is best illustrated by an example. Consider two variables x_1 and x_2 with domains $D_1 = \{1, \dots, 5\}$ and $D_2 = \{1, \dots, 5\}$. The constraint to be observed is $x_1 < x_2 - 2$. By examining the variable x_1 , we see that due to $x_2 \leq 5$ the constraint can not be satisfied for the values $x_1 \in \{3, 4, 5\}$ and, hence, the values can be removed from D_1 resulting in $D'_1 = \{1, 2\}$. Subsequently performing the same check for x_2 yields $D'_2 = \{4, 5\}$ due to $x_1 \geq 1$. In general, a constraint considering more than two variables is called consistent when for each possible value from the domain of a variable affected an assignment to all other variables from their domains can be made such that the constraint is satisfied. Furthermore, a CSP is consistent when this is true for all its constraints.

Within a constraint programming system, constraint propagation is usually applied iteratively to make the domains of each variable as small as possible, while making the entire CSP consistent. For this purpose, a number of algorithms have been developed among which the predominant one is called AC-5 (Van Hentenryck et al., 1992).

In the above example, evaluating the constraint $2x_1 = x_2$ for the initial domains $D_1 = \{1, \dots, 5\}$ and $D_2 = \{1, \dots, 5\}$ leads to the reduced domains $D''_1 = \{1, 2\}$ and $D''_2 = \{2, 4\}$ which guarantees consistency of the corresponding arcs. Now, if the results for the domain reduction applied to the constraint $x_1 < x_2 - 2$ have been propagated, the reduced domains $D'_1 = \{1, 2\}$ and $D'_2 = \{4, 5\}$ can be used in the evaluation. Then, we yield $D''_1 = \{2\}$ and $D''_2 = \{4\}$ which represents the only feasible solution to the CSP on hand.

Though for the small example constraint propagation seems to be rather simple, it may be much more complicated in case of more complex constraints. Therefore, a typical constraint programming system allows the user to define new propagation and domain reduction algorithms. Fortunately, state-of-the-art systems such as OPL provide large libraries of predefined constraints, including disjunctive ones as in our example. Therefore, it is often not necessary to create new constraints and to develop specialized propagation algorithms.

29.4 Search Algorithms

In general, algorithms for solving CSPs systematically enumerate all possible assignments of values to variables. By verifying for each combination of values whether it corresponds to a feasible solution or not, the algorithms are guaranteed to either determine a feasible solution, if one exists, or to prove that the problem is unsatisfiable. The most simple and common approach applied for this purpose is backtracking. To increase its performance, several extensions have been proposed among which forward checking and maintaining (arc) consistency seem to be most promising (Brailsford et al., 1999).

By *backtracking*, a multi-level enumeration (search) tree is systematically constructed. Each node of the tree corresponds to a partial solution in which values have been determined for a subset of variables. In each node on the current level of the tree, a yet unconsidered variable is selected. Subsequently, it is assigned a value from its domain thereby defining a node on the next level of the tree. If for this value, any of the constraints between this variable and those already considered is violated, a *dead end* is detected. In this case, the assignment is abandoned and a new neighbouring node is obtained by examining the next value of the variable's domain. Otherwise, if the assignment is feasible, the next variable for which no value has been determined yet is chosen and treated in the same fashion. As soon as all values of a variable have been examined, the search backtracks to the previous level and assigns a new value to the corresponding variable. For a CSP, the search can stop when a complete consistent solution has been obtained, i. e. a value has been determined to each variable such that all constraints are satisfied. If no feasible solution exists, the search is terminated after examining all possibilities of assigning values to variables.

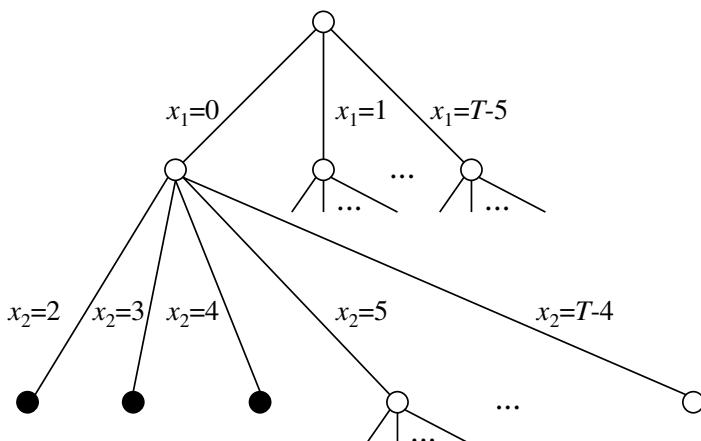


Fig. 29.1. Partial search tree for the example problem

Figure 29.1 shows a part of the search tree obtained for our production scheduling problem. On the root level of the tree, variable x_1 is selected. After assigning the value $x_1 = 0$, the variable x_2 is considered on the subsequent level of the tree. However, the values $x_2 \in \{2, \dots, 4\}$ are not feasible due to the constraints (29.1) which prevent that jobs are executed in parallel and, hence, lead to dead ends (black nodes). For $x_2 = 5$, the search may continue with selecting e.g. variable x_3 . After having examined all possible values of $x_2 \in D_2$ for $x_1 = 0$ as well as the possible assignments for the remaining variables on the subsequent levels of the tree, the search backtracks to the root node and the next value for x_1 is examined. Then, for $x_1 = 1$ all values $x_2 \in D_2$ have to be evaluated again etc.

Backtracking as described above only verifies the constraints between the current variable on a level of a tree and the variables considered on the previous levels. Within *forward checking*, after assigning a value to the current variable, all constraints affecting this variable are examined and values in the domains of yet unconsidered variables conflicting with this assignment are temporarily removed. If for one of these variables, the corresponding domain becomes empty, no feasible solution can be obtained by completing the current partial solution and, hence, the current variable value is infeasible and backtracking is performed. In case of *maintaining (arc) consistency*, additionally all available constraint propagation techniques are applied each time a variable has been assigned a value to temporarily reduce domains of the unfixed variables. That is, also inconsistencies in constraints which do not contain the variable just fixed itself but are affected indirectly are detected.

Obviously, the order in which variables are selected has a considerable influence on the size of the search tree. In particular, this is true if fixing the values of some variables also allows for reducing the domains of others. In this case, to keep the search tree as small as possible, those variables should be selected first the fixing of which lead to the largest domain reductions. Therefore, CP systems usually offer the possibility to determine the order in which variables are chosen (Van Hentenryck, 1999, Chap. 7).

29.5 Concluding Remarks

The previous sections show that CP for which a number of modern and easy-to-use software packages exist compares favourably with classical OR techniques in terms of modelling combinatorial decision problems. This is mainly due to the ease of defining logical constraints such as e.g. disjunctive constraints which often arise in production scheduling problems (Sect. 29.2). The formulation of such constraints is difficult within mixed integer programming (for further examples see Williams and Wilson (1998)). However, the performance of CP systems still seems to be rather poor with respect to solution quality and computation time (Brailsford et al., 1999). For a CP approach to be competitive to modern OR methods such as highly developed

branch and bound procedures or meta-heuristics (e.g. genetic algorithms), the constraints of the problem to be solved should be rather restrictive, like this is e.g. true in production scheduling in case of tight due dates. Then, after fixing a single variable, constraint propagation allows a large number of reductions in the domains of other variables. Unfortunately, the disadvantage remains that objective functions are considered indirectly in form of constraints. As a consequence, the search process has to be restarted each time after finding an improved solution such that certain parts of the search tree may be constructed repeatedly. This unnecessary computational effort for reconstruction considerably restricts the application of CP in particular to large problem instances. One way to limit the search effort is to provide an aspiration level for the objective (function) value instead of its optimization. However, one should bear in mind that a user specified aspiration level can be far from “optimal” or even result in finding no feasible solution.

In general, when CP is applied to solve problems optimally, it may benefit from OR by including bounding techniques as well as more versatile techniques to evaluate the search tree (see e.g. Klein and Scholl (1999) and Dorndorf (2002, Chap. 5)). If CP is used for determining heuristic solutions, it is promising to incorporate ideas from local search. Such approaches are, e.g., discussed in Nuijten and Aarts (1996) and Dorndorf (2002, Chap. 7.2).

References

- Brailsford, S. C.; Potts, C. N.; Smith, B. M. (1999) *Constraint satisfaction problems: Algorithms and applications*, European Journal of Operational Research, Vol. 119, 557–581
- Dorndorf, U. (2002) *Project scheduling with time windows – From theory to application*, Heidelberg
- Klein, R.; Scholl, A. (1999) *Scattered branch and bound – An adaptive search strategy applied to resource-constrained project scheduling*, Central European Journal of Operations Research, Vol. 7, 177–202
- Lustig, I. J.; Puget, J.-F. (2001) *Program does not equal program: Constraint programming and its relationship to mathematical programming*, INFORMS Journal on Computing, Vol. 31, 29–53
- Nuijten, W. P. M.; Aarts, E. H. L. (1996) *A computational study of constraint satisfaction for multiple capacitated job shop scheduling*, European Journal of Operational Research, Vol. 90, 269–284
- Van Hentenryck, P. (1999) *The OPL optimization programming language*, Cambridge/Massachusetts et al.
- Van Hentenryck, P.; Deville, Y.; Teng, T. M. (1992) *A generic arc-consistency algorithm and its specializations*, Artificial Intelligence, Vol. 57, 291–321
- Williams, H. P.; Wilson, J. M. (1998) *Connections between integer linear programming and constraint logic programming – An overview and introduction to the cluster of articles*, INFORMS Journal on Computing, Vol. 10, 261–264

Index

- AATP (allocated ATP), 186, 402
- aggregation, 85, 170
- alert monitor, 248, 350, 352
- all-units discount, 223
- allocation, 288
- allocation rule
 - fixed split, 190
 - per committed, 190
 - rank based, 189
- anticipation, 160
- APO PP/DS (Production Planning and Detailed Scheduling), 435
 - modelling philosophy, 438
 - modelling technique, 437
 - planning method, 439
 - product master data, 438
- APS (Advanced Planning System), 24, 84
 - assessment, 310
 - enabler, 298, 299, 396
 - implementation, 311, *see APS implementation project*
 - industry focus, 304
 - integration, 248, 312, 403
 - license fees, 308
 - number of installations, 306
 - post-implementation, 315
 - prototype, 311
 - release change, 315
 - selection, 303
 - super-user, 315
 - support, 315
 - system administration, 315
 - vendor, 304
- APS implementation project
 - definition phase, 309
 - implementation costs, 308
 - implementation time, 308
 - requirements, 309
 - selection phase, 303
- APS module
 - Collaborative Planning, 113, 259–277
 - coordination of, 245–248
- Demand Fulfilment and ATP, 110, 179, 195
- Demand Planning, 110, 139–157
- Distribution Planning, 110, 229–243
- interaction of, 246
- Master Planning, 110, 159–177
- Material Requirements Planning and Purchasing, 111
- Production Planning, 110, 197–213
- Scheduling, 110, 197–213
- Strategic Network Planning, 110, 117–136
- structure, 109
- Transport Planning, 110, 229–243
- ARIMA, 147, 347
- assemble-to-order, 184
- ATP
 - multi-level, 184
 - rule-based, 194
- ATP (available-to-promise), 91, 180, 181, 234, 401
 - allocation planning, 190
 - consumer goods industry, 184
 - consumption, 194
 - customer hierarchy, 188
 - dimension, 193
 - granularity, 182
 - product dimension, 182
 - retained, 190
 - search rule, 192, 193
 - time buckets, 186
- B2B (business-to-business), 17
- B2C (business-to-consumer), 18
- backorders, 151
- batch production, 372
- beer distribution game, 28
- best practice, 46
- BOC (bill of capacities), 217
- BOM (bill of materials), 202, 217
- BOM explosion, 110, 217
- bottleneck, 90, 160, 163
- Box-Jenkins-method, 147
- Branch and Bound, 478

- bullwhip effect, 27–28, 92
- business case, 283, 300
- business performance, 292
- capacity, 297
- CDCM (Collaborative Development Chain Management), 260
- change management, 329
- channel research, 25
- collaboration, 113, 261
 - capacity, 266
 - consumer-driven, 263
 - demand, 263
 - inventory, 265
 - materials and services, 267
 - multi-tier, 269
 - process, 270
 - procurement, 266
 - single-tier, 269
 - supplier-driven, 263
 - transport, 267
- collaborative forecast, 398
- Collaborative Planning, 259–277
 - software support, 276–277
- computer assembly, 99
 - capacity planning, 99
 - coordination, 103
 - distribution planning, 99
 - integration, 103
 - lot-sizing, 102
 - machine scheduling, 102
 - master production scheduling, 99
 - sales planning, 100, 101
 - transport planning, 103
- computer industry, 142, 389
 - assembly process, 392
 - business performance, 393
 - component planning, 398
 - product life cycle, 397
 - product structure, 391
- configure-to-order, 184, 391
- consensus-based forecasting, 263
- constraint propagation, 494, 495
- consumer goods industry, 93, 347, 371
 - capacity planning, 93
 - coordination, 96
 - distribution planning, 93, 95
 - forecasting, 95
 - integration, 96
- lot-sizing, 95
- machine scheduling, 95
- master planning, 93
- master production scheduling, 93
- sales planning, 94
- seasonality, 93
- sequence dependent setup time, 95
- continuous improvement, 300
- continuous review system, 156
- cooperation, 26
- coordination, 11, 26
- coordination of APS modules
 - Demand Fulfilment and ATP, 247
 - Demand Planning, 247
 - Distribution and Transport Planning, 248
 - Master Planning, 247
 - Production Planning and Scheduling, 247
 - Purchasing and Material Requirements Planning, 248
 - Strategic Network Planning, 246
- CP (constraint programming), 482, 493
- CPFR (Collaborative Planning, Forecasting and Replenishment), 260
- CPLEX, *see* ILOG
- CSP (constraint satisfaction problem), 493
- CTP (capable-to-promise), 91, 185
- customer service, 12, 371
- customer service level, 288
- cycle stock, 58, 132, 231, 238, 385
- data mining, 253
- decoupling point, 13, 182, 284
- demand
 - fluctuations, 295
- Demand Fulfilment and ATP, 179, 195
- Demand Planning, 139–157, 286
 - causal forecasting, 144
 - collaborative forecasting, 140, 142, 143
 - geographic dimension, 188
 - judgmental forecast, 140, 148
 - multi-dimensional forecast, 141
 - process, 415–419
 - rule-based forecasting, 150
 - seasonality, 144, 146
 - simulation, 140

- statistical forecasting, 140, 143
- trend, 143, 146
- what-if-analysis, 140
- demand supply matching, 402
- deployment, 239
- disaggregation, 170
- discount
 - all-units, 223
 - incremental, 223
- Distribution and Transport Planning, 229–243
- downstream, 294
- dual instruction, 172
- dual value, 175, 476
- DW (Data Warehouse), 249, 276
 - APS integration, 253
- EAI (Enterprise Application Integration), 254
 - integration adapter, 254
 - integration facade, 255
 - integration mediator, 255
 - integration messenger, 255
 - process automator, 256
- EBITA (earnings before interest, taxes and amortization), 122
- EDI (electronic data interchange), 17
- enabler, 298
- enabler-KPI-value network, 298
- ERP (enterprise resource planning), 84, 110, 133, 215, 290, 294
- every-day-low-price, 89
- exchange rate, 118
- exponential smoothing, 145, 347, 463
- external variability, 294
- financial performance indicators, 292
- flexibility, 129
- flexibility funnel, 288
- flow lines, 210
- flow shop, 393
- focal company, 16
- food and beverages industry, 371
 - demand planning, 375
 - distribution planning, 375
 - master planning, 376
 - production planning, 375
 - production scheduling, 377
 - sequence dependent setup time, 372
- setup time, 372, 379
- food industry, 143
- forecast
 - accuracy, 54, 152, 288, 289
 - error, 141, 152
 - netting, 400
- forecasting, *see* Demand Planning
 - seasonal coefficient, 461
 - seasonality and trend, 461
 - Winters' method, *see* Winters' method Planner)
 - BOM (bill-of-materials), 432
- FP (Factory Planner), 427
 - alternate resource, 433
 - data files, 428
 - modelling concept, 427–429
 - routing data, 433
 - spec files, 428
- frozen horizon, 163, 208, 213
- frozen period, 83, 377, 387
- GA (genetic algorithm), 343, 351, 485
- gantt-chart, 206
- heuristic, 83, 386, 485
- hierarchical (production) planning, 31
- hierarchical planning system, 84
- House of SCM, 11
- hybrid-flow-shop, 435
- i2 Technologies, 341
 - Active Data Warehouse, 394, 404
 - adapter, 405
 - Collaboration Planner, 394
 - Constraint Anchored Optimization, 343
 - Content, 344
 - Demand Fulfillment, 394, 401, 405
 - Demand Manager, 425
 - Demand Planner, 342, 393, 396, 427
 - Factory Planner, 341, 343, 394, 402, 421, 427
 - Factory Planner modelling, 430
 - Fulfillment Optimization, 345
 - High Availability, 405
 - PRO (Product Relationship Object), 393, 399
 - Production Scheduler, 343

- Revenue and Profit Optimization, 345
- RhythmLink, 394, 404, 427
- ROI (Optimization Interface), 405
- ROI (Optimization Interface), 394
- SDP (Strategy Driven Planning), 343
- Six.One, 341–346
- Supplier Relationship Management, 345
- Supply Chain Management, 342
- Supply Chain Planner, 343, 394, 401, 421, 427
- Supply Chain Strategist, 342
- TradeMatrixLink, 313
- Transportation Modeler, Optimizer and Manager, 344
- ILOG, 351, 376
- ILOG Cartridge, 409
- implementation
 - business case, 318
 - business plan, 327
 - close phase, 332
 - communication plan, 326
 - contractors, 320
 - execution and deployment phase, 330
 - phases, 317
 - project control, 320
 - project definition phase, 318
 - project management, 324
 - project sponsor, 320
 - scope, 319
 - solution design phase, 322
 - solution details phase, 327
- incremental discount, 223
- incremental planning, 207, 313
- indicator, *see* KPI (key performance indicator)
 - functions of, 49
- Integer Programming, 473
- integration, 10
 - mode, 313
 - of APS, 248
- inventory, 296
- inventory analysis, 56
- inventory in transit, 59
- inventory level, 289
- J. D. Edwards, 346
- judgmental forecast, *see* Demand Planning
- just-in-time, 89, 372
- KDD (knowledge discovery in databases), 253
- knowledge management system, 331
- KPI (key performance indicator), 13, 53–387
 - inventory age, 56
 - service level, 54
 - as-is value, 299
 - asset turns, 55
 - assets, 55
 - cash flow, 293
 - costs, 56
 - delivery performance, 53, 182
 - forecast accuracy, 54, 393, 396
 - inventory turns, 55, 182, 393
 - on time delivery, 54, 179, 181, 191, 288, 289, 393
 - order fill rate, 54, 288
 - order lead-time, 54, 179, 182, 288
 - order promising lead-time, 191
 - planning cycle time, 55
 - production lead-time, 289
 - profile, 299, 300
 - ROA (return on assets), 293
 - ROCE (return on capital employed), 292
 - supply chain flexibility, 55
 - supply chain responsiveness, 55
 - to-be value, 299
 - value-added employee productivity, 56
 - warranty costs, 56
- lead-time, 218
- lead-time offset, 212, 216
- life-cycle-management, 88, 154
- linear regression, 146, 470
- logistics, 23
- lost sales, 151
- lot-sizing, 164, 205
- lot-sizing stock, *see* cycle stock
- low-level code, 217
- LP (Linear Programming), 83, 166, 239, 343, 347, 351, 379, 386, 473
- LSP (logistics service provider), 229

- make-to-order, 185, 373, 391
- make-to-stock, 91, 183, 373, 391
- management-by-exception, 248
- Manugistics NetWORKS Demand, 387
- marketing, 287
- Master Planning, 159–177, 180, 181, 232
 - data, 165
 - decisions, 163
 - model adjustment, 176
 - model building, 167
 - objectives, 164
 - planning horizon, 162
 - planning-profile, 169
 - results, 166
- master production scheduling, 90
- mean absolute deviation, 152
- mean absolute percentage error, 152
- mean squared error, 152
- metrics, 46, 49
- MIP (Mixed Integer Programming), 118, 166, 347, 351, 379, 380, 473, 477
- model, 81
 - forecasting, 81
 - optimization, 81
 - simulation, 81
- moving average, 145
- MRO (maintenance, repair and operations), 223
- MRP (material requirements planning), 65, 84, 90, 215–222
- multi-level ATP, 184
- multi-objective decision problem, 82
- NCF (net cash flow), 120
- network flow model, 239
- NIBT (net income before taxes), 125
- NPV (net present value), 117
- objective function, 81
- OLAP (online analytical processing), 253, 342, 351
- OLTP (online transaction processing), 249
- operations research, 24, 83
- OPT, 303
- optimization
 - problem, 485
- Oracle, 376, 386, 427
- order
 - initial promise, 288
 - life cycle, 288
 - management, 288
 - matching, 373
 - promising, 179, 191, 393
 - promising (MRP logic), 181
 - quoting, 179
- order picking inventory, 62
- Pareto-dominated, 122
- pegging, 220
- penalty costs, 173, 385
- Peoplesoft, 290, 294, 346
 - Advanced Planning Agent, 348, 376, 386
 - commodity, 378
 - Demand Consensus, 347
 - Demand Forecasting, 347
 - Demand Management, 347
 - Distributed Object Messaging Architecture, 349
 - model building, 377
 - Order Promising, 348
 - Production & Distribution Planning, 347
 - Production Scheduling, 347, 376, 377, 387
 - Production Scheduling Discrete, 348
 - Production Scheduling Process, 347
 - Strategic Network Optimization, 347, 371, 376, 377, 381, 386, 387
 - Supply Chain Planning, 346–350, 371
 - Vehicle Loading, 347
- performance measures, 49
- periodic review system, 156
- pharmaceutical industry
 - APO PP/DS model, 361
 - APO SNP model, 365
 - master data model, 364, 368
 - planning processes, 361
 - project objectives, 358
 - supply chain characteristics, 355
- phase-in/phase-out, 154
- planning, 81
 - event-driven, 83
 - horizon, 81, 109, 162
 - interval, 197
 - module, 84

- operational, 82
- scenario, 294
- strategic, 82
- tactical, 82
- planning domain, 259
- planning tasks
 - long-term, 88–89
 - mid-term, 89–91
 - short-term, 91–92
- POS (point-of-sale), 30
- postponement, 25, 287
- PPM (Production Process Model), 202, 443
- pre-built stock, *see* seasonal stock
- primal instruction, 172
- priority rules, 205
- process organization, 209
- procurement, 287
- product
 - life cycle, 287, 295
 - management, 287
 - quality, 289
 - substitution, 194
- production lot-sizing stock, *see* cycle stock
- Production Planning and Scheduling, 197–213
 - activity, 203
 - location, 202
 - model, 201
 - objectives, 205
 - operating instruction, 202
 - operation, 203
 - part, 202
 - resource group, 211
 - routing, 202, 205
- production schedule, 197
- production scheduling problem, 486
- project
 - control and reporting, 324
 - coordinator, 321
 - functional and process team, 321
 - management, 321
 - members, 328
 - phase, 317
 - plan, 325
 - roadmap, 297
 - ROI (return on invest), 300
 - team, 321, 328
 - team leaders, 321
- promotion, 295
- quarantine stock, 385
- ratio-to-moving averages decomposition, 467
- regression analysis, 146
- return on assets (ROA), 293, 297
- return on capital employed (ROCE), 292
- rolling horizon, 83
- rule-based ATP, 194
- runtime objects, 204
- safety stock, 61, 89, 90, 96, 133, 139, 155, 165, 238, 377, 430
- safety stock management, 400
- sales
 - ATP consumption, 190
 - central, 286
 - forecast, 287
 - regions, 286
- SAP, 350
 - liveCache, 352, 410
 - Alert Monitor, 352, 439, 440
 - Optimizer), 312
 - APO (Advanced Planner and Optimizer), 350–353
 - BAPI (Business Application Programming Interface), 352
 - Business Information Warehouse, 350, 353, 361
 - CIF (Core Interface), 352
 - Collaborative Planning, 353
 - Data Warehouse, 350, 410
 - Demand Planning, 350, 361, 408, 410–415
 - Deployment and Transport Load Builder, 351
 - Global ATP, 351, 409
 - InfoCube, 411
 - mySAP SCM, 350
 - Plan Monitor, 352
 - Production Planning and Detailed Scheduling, 351, 361–365, 409, 435
 - Purchasing Workbench, 352
 - R/3, 290, 294, 312, 357, 361, 387, 403, 409

- SCEM (Supply Chain Event Management), 352
- Supply Chain Cockpit, 350, 352
- Supply Chain Engineer, 352
- Supply Network Planning, 351, 361, 365–369, 408, 441
- Transportation Planning and Vehicle Scheduling, 352
- SCC (Supply-Chain Council), 41
- SCEM (Supply Chain Event Management), 256
- SCM (Supply Chain Management), 24
 - building blocks, 11–24
 - definition, 11
 - gains, 1
 - project, 281
- SCM strategy, 19
- SCOR-model (Supply Chain Operations Reference-model), 41
 - application, 47
 - best practice, 46
 - deliver, 43
 - levels, 42
 - make, 43
 - metrics, 46
 - plan, 42
 - process category, 43
 - process element, 44
 - process type, 42
 - return, 43
 - source, 42
 - standard terminology, 41
- SCP-Matrix (supply chain planning matrix), 87, 109
- seasonal stock, 59–61, 93, 132, 160, 166, 387
- selection of an APS, 304
- semiconductor industry, 421
 - corporate planning, 426
 - demand planning, 425
 - divisional planning, 426
 - facility planning, 426
 - process, 422
 - production scheduling, 426
 - production, 422
 - wafer test, 424, 431
- sequence dependent setup time, 211
- service level, 54, 156
- software
 - component, 109, 111, 112
 - module, 87, 109–112
- sporadic demand, 151
- spreadsheet, 290
- star scheme, 411
- steering committee, 16, 321
- stochastic optimization, 118, 120
- stock components, 57
- stockout, 294
- Strategic Network Planning, 117–136, 232
 - constraints, 123–125
 - functions of APS modules, 134
 - objective, 121
 - planning horizon, 117
 - planning problem, 120
 - planning process, 129
- strategy, 19
- styrene plastics
 - Demand Plannig, 407
 - planning architecture, 408
 - supply chain characteristics, 408
- supplier
 - contract, 288
 - flexibility, 288
 - forecast, 287
 - lead-time, 287
- supply chain, 9
 - demand constrained, 186
 - material constrained, 389
 - potential analysis, 291, 292
 - process, 87, 109
 - supply constrained, 186
 - type, 111, 373
- supply chain design, 117, 128–133
- supply chain evaluation, 284
- Supply Chain Performance Benchmarking Study, 46
- supply chain visibility, 257
- synthetic granulate industry, 435
 - APO PP/DS model, 437
 - extrusion process, 435
 - production process, 435
- theory of constraints, 303
- time bucket, 163, 210, 376, 378
- time-series-analysis, 143
- time-to-market, 289

- topography of a supply chain, 70
- transit stock, 236, 385
- Transport Planning, *see* Distribution and Transport Planning
- transportation lot-sizing stock, 58–59
- transshipment point, 229
- typology
 - distribution type, 68
 - functional attribute, 66
 - integration and coordination, 70
 - procurement type, 66
 - production type, 66
 - sales type, 68
- structural attribute, 69
- upstream, 294
- value driven APS implementation, 298
- vehicle loading, 241
- vehicle scheduling, 234, 241
- VMI (vendor managed inventory), 75, 89, 265, 351
- Winters' method, 146, 154, 463
 - seasonal coefficient, 146, 154
- WIP (work-in-process), 61, 289

About Contributors

Frank Altrichter is a manager at j & m Management Consulting, Mannheim. From October 1997–August 2000, he was a consultant at Numetrix GmbH (quired 1999 by J.D.Edwards). He worked as a technical implementor, solution architect, and project manager in a large, international APS implementation projects, mainly in the process industry. He studied mechanical engineering at Darmstadt University of Technology and holds a certification for production and inventory management (CPIM) from APICS. He can be contacted at [⟨frank.altrichter@jnm.de⟩](mailto:frank.altrichter@jnm.de).

Tanguy Caillet is a senior consultant at j & m Management Consulting, Mannheim. From January 2000-October 2001, he was an in-house consultant in Warehouse Management at CrownCork & Seal Inc. a leader in consumer packaging goods. Since he joined j & m, he works as a technical implementor, solution architect and project manager in large SAP APO projects with a focus on SNP. He holds a Masters Degree of Industrial Engineering from the French Engineering School, ICAM (Institut Catholique des Arts et Metiers), in Toulouse. He can be contacted at [⟨tanguy.caillet@jnm.de⟩](mailto:tanguy.caillet@jnm.de).

Prof. Dr. Bernhard Fleischmann holds a chair for Production and Logistics at the University of Augsburg. 1978–1991 he was a professor of Operations Research at the University of Hamburg. 1971–1978 he worked in the Operations Planning department of Unilever Germany. His research interests include the development and application of systems for production planning, transportation and distribution planning and inventory management. He can be contacted at [⟨bernhard.fleischmann@wiso.uni-augsburg.de⟩](mailto:bernhard.fleischmann@wiso.uni-augsburg.de).

Prof. Marc Goetschalckx is Associate Professor in the School of Industrial and Systems Engineering of the Georgia Institute of Technology. His research interests are in the areas of analysis and design of material flow networks, ranging from the design of global supply chains to the dispatching of vehicles to pick orders in a warehouse. He has written numerous articles, is a frequent speaker at international meetings and short courses, consults, and has developed decision support software in these areas. He can be contacted at [⟨marc.goetschalckx@isye.gatech.edu⟩](mailto:marc.goetschalckx@isye.gatech.edu).