

THE ONGOING CHALLENGE – AN ACCURATE ASSESSMENT OF SUPPLY LINKED TO DEMAND TO CREATE AN ENTERPRISE-WIDE END TO END DETAILED CENTRAL SUPPLY CHAIN PLAN

Ken Fordyce,
Alfred Degbotse
John Milne
Robert Orzell
Chi-Tai Wang
IBM
Essex Junction, VT USA 05452

ABSTRACT

Organizations can be viewed as an ongoing sequence of loosely coupled decisions where current and future assets are matched with current and future demand across the demand-supply network at different levels of granularity ranging from a placing a lot on a tool to an aggregate capacity plan across a five year horizon. Since the early 1990s detailed enterprise wide central planning has become a key member of this “decision suite.” Despite its importance, most organizations execute central planning with “limited levels of accuracy or intelligence.” Early in the evolution of “central planning engines” IBM determined that “extended accuracy” was an important component of supply chain efficiency and customer satisfaction and made a substantial investment to develop a central planning engine which could handle the scope (complexity) and scale (size) of large organizations. This paper covers the value from this investment and the technical details of combining heuristics and optimization.

1 INTRODUCTION

Most organizations can be viewed as an ongoing sequence of loosely coupled decisions where current and future assets are matched with current and future demand across the demand-supply network at different levels of granularity (Figure 1) ranging from a placing a lot on a tool to an aggregate capacity plan across a five year horizon.

Since the early 1990s detailed enterprise wide central planning (whose purpose is to create a detailed projected supply linked to detailed exit demand and synchronize the activities of the enterprise) has become a key member of this “decision suite.” Despite its importance, most organizations execute central planning with “limited levels of accuracy or intelligence” (smarts) which can and does create substantial inefficiencies or slack (Galbraith 1973) resulting in a firm operating below its efficiency frontier. Early in the evolution of “central planning engines” (the model and associated software which creates the plan) IBM determined that “extended

accuracy” was an important component of supply chain efficiency and customer satisfaction and made a substantial investment to develop a central planning engine which could handle the scope (complexity) and scale (size) of large organizations. That is a “smart planning engine” was the critical third leg (Figure 1) of supply chain efficiency (the other two legs being process and accurate data).

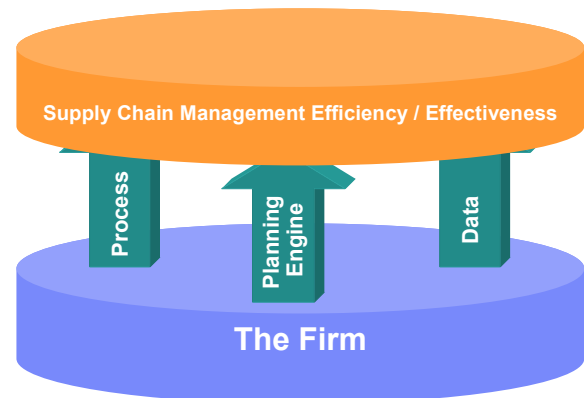


Figure 1: Three basic pillars supply chain planning

The complexities include, but are not limited to: carrying detailed lot and order information across a detailed multi-staged production path; alternative methods to produce the same part (either finished or intermediate) where the consumed part and resource required may or may not be the same; complex sorting where the exact identify of the part is not known until one or more tests occur; business policies such as fair share between customers and vendor allocation; lot sizing; required minimum starts (which mixes push and pull); attempting to meet customer request date without impacting the ability to meet customer commit date; maintaining lot level and exit demand level detail throughout the solution

Clearly, the complexity associated with managing demand-supply networks for semiconductor manufacturing makes it an ideal candidate for linear programming (LP) technology. However, the scale of the enterprise problem

is too large to be solved by a single LP model. Simple heuristics such as greedy algorithms can handle the scale. But they fail to handle even such simple complexities as binning and typically underutilize expensive assets. In order to build a central planning engine (CPE), the IBM team developed and implemented a series of significant advances, including an LP model for binning optimization that could be invoked by a heuristic, an advanced heuristic which provides highly effective decisions with regards to such items as demand priorities and binning (invoking the binning LP model), an advanced LP-based algorithm which handles complex trade-offs and nonlinearities such as lot sizing and preemptive demand priorities, an overall solution structure which dynamically blends both decision technologies and partitions the solution process in a manner that adds to the quality of the solution, and mechanisms to fully automate the process. The CPE is capable of solving enterprise-scale SCM problems with both function and speed. Its partitioning logic automatically classifies product structures and solves the most complex ones using large LPs, the moderately complex ones using small LPs, and the relatively simple ones using heuristics. The CPE achieves all this without any human intervention. The core of this work is described in the presentation.

This CPE is in regular use within IBM, Analog Devices, and other clients – generating millions in “savings”. Perhaps the best testimony came from Jerry Dundon (senior VP supply chain management, Analog Devices, Inc.) after the IBM CPE was successfully installed at Analog: “Much has been written about improving supply chain operations through reducing demand volatility and improving forecast accuracy. We have put equal focus on understanding and modeling the much-neglected supply side of the equation. Having a robust, credible model of supply capability allows us to be more flexible and responsive to inevitable changes to demand”.

2 OVERVIEW OF BEST-CAN-DO ENTERPRISE WIDE CENTRAL PLANNING ENGINES

The best-can-do (BCD) enterprise wide detailed central planning engine (CPE) is the control point for the flow of material or product within an organization, and focuses on how to best meet prioritized demand without violating temporal, asset (WIP and inventory), or capacity constraints. A CPE application minimizes prioritized demand tardiness and some aspects of cost, establishing a projected supply and synchronized targets for each element of the supply chain.

The core of the CPE process is matching assets with demand, which refers to aligning assets with demand in an intelligent manner to best meet demands. The alignment or match occurs across multiple facilities within the boundaries established by the manufacturing

specifications, process flows, and business policies. Assets include, but are not limited to, starts (manufacturing releases), work in progress (WIP), inventory, purchases, and capacity (manufacturing equipment and manpower). Demands include, but are not limited to, firm orders, forecasted orders, and inventory buffer. The matching must take into account manufacturing/production specifications and business guidelines. Manufacturing specifications and process flows include, but are not limited to, build options, bill of material (BOM), yields, cycle times, anticipated date on which a unit of WIP will complete a certain stage of manufacturing (called a receipt date), capacity consumed, substitutability of one part for another (substitution), the determination of the actual part type after testing (called binning or sorting), and shipping times. Business guidelines include, but are not limited to, frozen zones (no change can be made on supplies requested), demand priorities, priority tradeoffs, preferred suppliers, and inventory policy. Many of the manufacturing specification and business guideline values will often change during the planning horizon (time effective).

The creation of a CPE plan requires a solver (sometimes referred to as a model or an engine) with the following core features:

1. method(s) to represent the (potential) material flows in production, business policies, constraints, demand priorities, current locations of asset, etc., and relate all this information to exit demand;
2. capture asset quantities and manufacturing specifications (parameters);
3. search mechanism(s) to generate a balanced supply chain plan, relate the outcome to demand, and modify the plan to improve the match;
4. display and explain the results of the best-can-do match.

3 OVERVIEW OF THE IBM CPE

This section describes the following core aspects of the CPE solution mechanism, with a focus on its capability to dynamically mix and match linear (mathematical) programming (LP) and advanced heuristic resource allocation technology. These core elements are:

1. IMEX (implode-explode) heuristic and its ability to dynamically invoke the Binning LP model to handle simple binning with downgrade substitutions (but not complex situation such as alternative BOM, complex binning, and general substitution).
2. SCOPE (Supply Chain Optimization Planning Engine), an advanced LP based engine which handles the BSA, lot sizing, and preemptive priorities;
3. multidimensional partitioning strategy; and
4. Structures and algorithms to dynamically mix and match both resource allocation decision technologies (IMEX and SCOPE) to balance scope and scale.

3.1 IMEX Implode/Explode Best Can Do Heuristic

IMEX (Figure 2) executes three major steps. Step 1 is an MRP (material requirements planning) type of explosion of demand across the entire BOM supply chain where hints are established and the binning points are optimized with the “Binning LP” (BINLP). Hints refer to required manufacturing starts, tentative assignment of WIP or inventory to support a specific start, and demand priority associated with each level of exploded demand (Dangat et al. 1999). The (optional) second step allows the files generated during the first step to be modified to influence the final creation of anticipated supply. For example, the wafer start file could get modified. The third step executes a BOM implosion to create a final best-can-do solution that meets planning requirements such as demand class, fair share, etc.

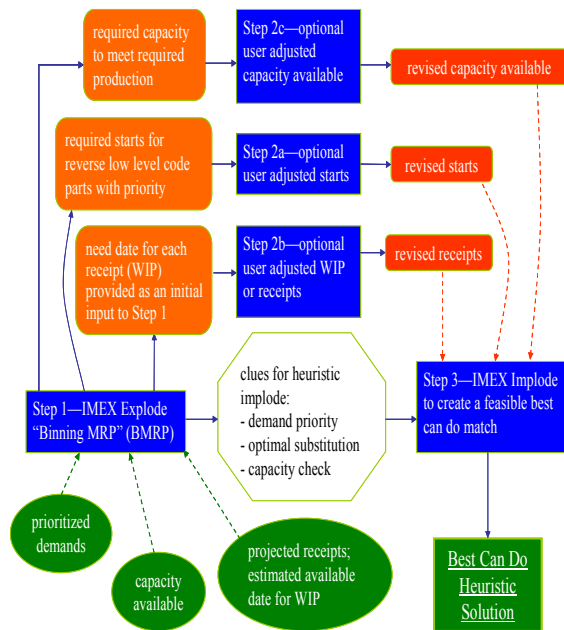


Figure 2: IMEX smart heuristic

3.2 Supply Chain Optimization Planning Engine

SCOPE is a linear programming (LP) based supply chain solution developed and deployed by IBM in the 1990s that has been continuously enhanced since. This solution is primarily used to handle the complex portions (complex binning, general substitution, and alternative BOM) of the demand supply network for which the IMEX and other heuristic based solutions are in general unable to identify optimal solutions. The SCOPE solves a supply chain LP model with a cost minimization formulation. So a minimum cost solution is obtained when the model is solved to optimality. SCOPE considers such costs as backorder, processing, inventory holding, material

substitution, part shipment (ship to customers or ship to other manufacturing locations within the enterprise), and a few others. In addition to costs, the SCOPE requires the usual input such as customer demand, scheduled vendor shipment, yield, capacity, cycle time, and so on. The input data structure is very much the same as that for the IMEX. Also as in the IMEX, most input data can be provided as time effective that the value would change over time to reflect the real situation.

The decision variables in the SCOPE LP model are chosen to correspond to actual supply chain activities: customer shipment, backorder, substitution, manufacturing start, interplant shipment, inventory level, and sourcing. In any feasible solution, these variables must hold values to satisfy the following five types of constraints:

- Material balance equations maintain a flow balance between the creation and consumption (arrival and departure) of any part number at any stocking point, in any period, and at any manufacturing location. These equations handle all the complexities associated with binning, general substitution, and alternative BOM structures.
- Backorder conservation constraints keep track of unsatisfied demand throughout the planning horizon. That is, in any period they capture the total unsatisfied quantity contributed by all demands with the same part number, customer location, and demand class so it can be backordered and met in a future period.
- Capacity control equations safeguard capacity utilization, ensuring no resource is over-utilized by starts requiring the same resource.
- Sourcing constraints enable the SCOPE users to control the deviation from sourcing targets, each of which is specified by two numbers,
- Non-negativity constraints require all decision variables to be greater than or equal to zero.

3.3 “Divide & Conquer” – Decomposing the Problem

Despite the constant improvements in hardware performance, determining an optimal central plan for a large enterprise with just the SCOPE (or LP-based solutions in general) along is not even remotely close to being realistic. If the IMEX heuristic is used by itself, the quality of the decisions would most likely be less than desirable for complex product flows. Besides performance and ability to handle complex product flows, there are business reasons to partition and sequence the solution flow. For example, executing an explosion step identifies a prioritized set of required starts / releases needed to satisfy all demand in time. This is essentially a prioritized “wish list.” This information would not be available with a single, monolithic solution process, since the only starts ever calculated are feasible with respect to capacity and component availability. The information of prioritized starts is particularly helpful when a planner is attempting to

determine what actions to take to meet certain demands that are currently behind schedule or to establish minimum starts. Therefore, developing a fully automated, dynamically partitioning-and-sequencing algorithm which makes the best use of both heuristics and LP technologies was the logical course of action. For details see Fordyce et al (2008).

4 NEW SCIENCE EMERGES AND EXTENDS THE BORDERS OF BOUNDED RATIONALITY

Today IBM and some of its clients use a combination of linear programming and some pretty clever "heuristics" to establish a daily plan for their enterprise. These firms cannot think of "life without these decision models." This same level of success of decision models can be seen in a number of areas. We believe we are witnessing a growing awareness of the importance of decision models and already competing on analytics (Davenport 2006).

That said, there is still a long way to go. People are comfortable with their guesses and decision scientists often fail to deliver real value. Politely, life is much cleaner if your modeling work remains an academic exercise. What decision scientists offer is the potential to "Extend the Borders of Bounded Rationality."

Herbert Simon (Nobel Prize Winner in Economics) observed, "as humans, we have 'bounded rationality' and break complex systems into small manageable pieces." The challenge for organizations is to integrate information and decision technology to push boundaries out and improve performance.

The challenge for the Operations Research community is to understand the task is not to optimize a solution point in the process, but the process and a sequence of decisions.

REFERENCE

- Dangat, G., A. Gokhale, S. Li, R. Milne, R. Orzell, R. Reid, X. Tang, and C. Yen. 1999. Best can do matching of assets with demand in microelectronics manufacturing. U.S. Patent 5,971,585 (10/26/1999).
- Fordyce, K., A. Degobtse, R. Milne, R. Orzell, and C. Wang 2008. Ongoing challenge. Forthcoming as chapter in *Handbook of Production Planning*.
- Davenport, T. 2006. Competing on analytics. *Harvard Business Review* 1–9.
- Galbraith, J. 1973. *Designing complex organizations*. Addison-Wesley.
- Lyon, P., R. Milne, R. Orzell, and R. Rice 2001. Matching assets with demand in supply-chain management at IBM Microelectronics. *Interfaces* 31(1):108–124.

AUTHOR BIOGRAPHIES

KENNETH FORDYCE joined IBM in 1977 and is involved in the integrated application of decision and information technology to improve organizational performance especially for planning, scheduling, and dispatch. Along with his co-authors he has received awards from IBM, INFORMS, POMS, and AAAI; been a key contributor to IBM's advanced supply chain planning solution used inside and outside of IBM; and published numerous papers on planning, scheduling, and dispatch. He is an adjunct at Columbia University; serves as a research team member at Einstein College of Medicine; and has a PhD from Union.

ALFRED DEGOBTSE is senior engineer at IBM with a wide range of responsibilities in Operations Research and Statistics much of them in planning and scheduling. He serves on the invention review board, holds a number of patents, and is an Edelman Judge for Informs. He has a PhD from Virginia Tech.

R. JOHN MILNE develops commercially available supply chain optimization software for IBM and other high technology companies. He has over 30 patents granted and pending at the US Patent and Trademark Office and is the VP of Practice for Informs. He holds a PhD in Decision Sciences and Engineering Systems from Rensselaer Polytechnic Institute.

ROBERT ORZELL is a Senior Engineer with over 30 years of experience at IBM. He has held a variety of positions, from technical positions developing software to support production planning and inventory management, to a management position in Production Control charged with rolling out advanced software. His current assignment is as a technical leader in the design and development of leading edge supply chain management and decision support software for IBM where he has been a key contributor to both heuristics and LP applications.

CHI-TAI WANG has been working at IBM since 1999 developing and enhancing advanced supply chain solutions and supporting their production use at IBM and IBM's clients. He is also an editor for Journal of the Chinese Institute of Industrial Engineers. He received PhD and Master of Science Degrees in Industrial and Operations Engineering from the University of Michigan at Ann Arbor.