Optimization Models for Restructuring BASF North America’s Distribution System

Slava Sery
ssery@mindspring.com
7500 Southwind Drive
Apartment 207
Chesterfield, Virginia 23832

Vince Presti
vjpresti@aol.com
BASF Corporation North America
3000 Continental Drive North
Mount Olive, New Jersey 07828

Donald E. Shobrys
dshobrys@home.com
186 Blackburn Road
Summit, New Jersey 07901

By 1995, annual distribution costs for BASF North America’s packaged goods were nearly $100 million. The firm explored trade-offs between customer service and operating costs in a redesign effort using linear-programming-based models. The project team adapted formulations to the extensive available data and used a series of formulations to cope with the scale of the project. A flexible modeling tool aided the team in implementing these formulations. The resulting revised distribution system reduced costs and improved customer service, and the modified distribution network took next-day deliveries from 77 percent to 90 percent. Although the team expected reduction in annual costs of 10 percent, subsequent customer service initiatives reduced the potential savings. In studies following the initial distribution changes the team estimated annual costs savings at six percent, but also identified a one-time nine-percent improvement in cash flow from inventory reductions. BASF also applied the models to operations in Scandinavia, Europe, and the Asia-Pacific area.

The BASF Group, with headquarters in Ludwigshafen, Germany, is one of the world’s leading chemical companies, with sales in 1997 of $32 billion and 100,000 employees worldwide. BASF offers a range of chemical and chemical-
related products in Europe, the NAFTA Region, South America, and Asia. BASF’s slogan, “We don’t make a lot of the products you buy. We make a lot of the products you buy better,” summarizes its diverse product mix, which includes chemicals, polymers, automotive coatings, colors, dyes, pharmaceuticals, nylon fibers, and agricultural products.

BASF Corporation is one of the strongest affiliates of the worldwide BASF Group, accounting for 21 percent of total sales, $6.9 billion in 1997. BASF Corporation and its affiliates, headquartered in Mount Olive, New Jersey, are the NAFTA region members of the BASF Group. BASF Corporation and its affiliates employ 15,000 people in the NAFTA region.

In the mid-1990s, the BASF Corporation examined its distribution of packaged goods in North America and found that it shipped 1.6 billion pounds of finished goods annually to customers from a network of 135 locations at an annual cost (transportation and warehousing) of nearly $100 million. Of the 135 shipping locations, 86 were distribution centers (DC) that received products from other locations and shipped to customers. Almost a billion pounds of products went directly to customers from 40 plants and 27 tollers, while 660 million pounds of products were shipped to customers through 86 DCs, 68 public and private warehouses, and 18 plants and tollers. Tollers are companies that provide additional production capacity on a contract basis. Hazardous, flammable, or temperature-controlled products required special storage at 56 percent of the facilities. Imported products made up one third of DC volume.

Acquisitions and divestitures during the 1980s left BASF with many operating divisions using independent distribution networks. Multiple facilities, many close to each other, served customers through outdated distribution networks that delivered only 77 percent of the shipped volume by the next day. BASF clearly needed to reconfigure its distribution facilities to serve customers more effectively. It initiated an effort to develop a strategic resource-planning tool for assessing the cost-versus-service consequences of distributing packaged goods in North America.

Project Scope

BASF gave the project team this problem:

Define the optimal number and location of warehouses and the corresponding material flows needed to meet anticipated customer demand and required delivery service times at the lowest overall cost.

Vince Presti of BASF corporate logistics systems led the project team. BASF used software and consulting services from Chesapeake Decision Sciences, Inc. Chesapeake staff included Slava Sery, the consultant who formulated the models, and Donald Shobrys, who as vice president of operations was responsible for consulting activities. BASF personnel from corporate warehousing and transportation and from the decision sciences support group contributed to data-collection and model-design efforts.

We performed this study under the direction of Nancy Carpenter, director of physical distribution, warehousing, and transportation. She reviewed model results and prepared presentations for Bernd Flickinger, vice president of logistics. Ms.
Carpenter and Mr. Flickinger worked with us in preparing model runs and arriving at the project’s final recommendations, which they presented to Dieter Stein, chairman of BASF North America.

We examined two major areas of performance: customer service and cost. Customer service was measured in terms of same-day and next-day delivery of product shipments. The costs included fixed costs associated with the distribution centers, variable handling and storage costs at the distribution centers, freight costs for replenishing distribution centers, and freight costs for shipments to customers. Since the distribution system relies on both truckload (TL) and less-than-truckload (LTL) shipments, one mechanism for reducing costs was to identify network configurations that allow the use of truckload shipments instead of the more expensive less-than-truckload shipments.

Use of Optimization

Decision making in the real world always contains subjective elements since decisions occur in an uncertain world and often address conflicting objectives. We faced conflicting objectives of minimizing distribution costs and at the same time improving customer service. The decision maker must assess the trade-offs between these conflicting objectives in selecting a course of action. The challenge the project team faced was to build a quantitative mathematical model that was realistic enough to help the decision maker in this qualitative decision-making process.

The first decision we faced was whether to attack the problem with a single, all-inclusive model or to take a more approximate approach using a series of formulations. People differ on the role of optimization in supporting decision making. Some emphasize the importance of optimality, with success stories to support their views [Arntzen et al. 1995; Geoffrion and Powers 1995]. They are willing to deal with the size and complexity of the model needed to generate a single-pass, optimal solution. Although the optimizers and computing environments required to solve problems with millions of decision variables are available [McBride 1998], this approach has potential problems:

—The massive effort required to collect accurate and complete data,
—The need for complex and time-consuming solution methods,
—The difficulty of finding a human decision maker willing to trust the results generated by a complex tool, and
—The investment required to develop the systems and user skills needed to collect and validate data and to interpret solutions.

Others [Camm et al. 1997] have successfully used more approximate approaches that make strong but reasonable assumptions about data, and rely on one or more simplified models. The human decision maker makes the final decision based on postsolution analysis of the models’ response to a number of different scenarios. This approach leads to quick model development and often provides valid insights on the system being modeled, but it also has some potential shortcomings:

—Different people may interpret the results differently and thus the final decision may depend on a specific decision maker and his or her experience.
—The project team or the decision maker may overlook some scenarios that are important to understanding the system being modeled.

Actual practice often falls between these two extremes. Many optimization projects migrate from simple to more complex models over the course of the project to ease data validation and help the decision maker to understand and gain confidence in the solution methods being used.

We decided to take a simplified approach for the following reasons:
(1) The multiyear time horizon associated with strategic decisions increases uncertainty about the data being used and reduces the justified level of detail.
(2) Collecting and validating data was extremely time consuming.
(3) This tool would be used infrequently, making it hard to justify large investments in data integration and user training.
(4) The decision makers had little experience with optimization-based tools and would find a simplified approach more credible.

In performing the steps of data collection and analysis, model definition, model development, and presentation of results with acceptance by top management, we expended 2.5 workforce years of effort over a 14-month period.

Data Collection and Aggregation
Disparate business activities and systems complicated data collection. BASF’s distribution network served 17 operating divisions with a wide range of legacy systems and data definitions. It distributed a large number of products to many customer locations.

We collected historical data on product demand and costs from many sources, made them consistent, and codified them to provide distinct groupings for model computations and reports. We then aggregated 25,000 SKUs (stock-keeping units) into 382 business product categories, and 15,000 customer demand points into 287 postal-zip-code geographical regions. The design of the aggregation process considered the need to report results to business units, determine transport-mode costs for truckload and less-than-truckload quantities, develop storage costs across different categories of materials, and set costs for handling activities.

Personnel from the corporate transportation department helped develop meaningful transportation costs. They analyzed active carrier contractor rates and prepared the cost structures used in the models. Transportation costs were in dollars per mile for truckload (TL) shipments, and dollars per hundredweight for less-than-truckload (LTL) shipments. Truckload distances between aggregated postal-code centers came from available mileage tables, and the rates used were those for known carriers. We incorporated the ability to change the mix between direct and DC shipment volumes by freezing the direct linkage between plants and customers.

Tables of LTL costs at various weight breaks between origin and destination zip codes came from the Yellow Freight rating service. BASF transportation department
analysts adjusted interstate and intrastate rates to meet contracts in place. We used regression analysis to extend rate tables into regions not covered by current contracts. We calculated the effective rates between origin and destination points using historical demand to estimate the expected mix of LTL and TL deliveries. This greatly reduced the number of decision variables in the models.

Delivery times were keyed to distance. We used mileage tables available from the transportation department to constrain the volume delivered to a region using specified service levels, such as one day for 400 miles, two days for 700 miles, and so forth. BASF decision-sciences support staff worked with the model designers in developing dialogues that let the users selectively use mileage tables to freeze or unfreeze links between origins and destination. This let the users focus the model’s optimization capabilities on specific areas of interest by freezing the agreeable portions of the network. This selective-freeze capability reduced the number of decision variables in a given optimization run and sped up calculations, enabling the users to quickly explore and analyze a wide range of alternatives.

Analysts in the BASF warehousing group developed unit storage charges in dollars per square foot for owned and public facilities from actual data based on product type and footprint, pallet size, stacking, inventory velocity, and inventory turns. BASF staff compared these calculated values and quotes from third-party vendors to benchmarked industry standards for similar facilities in regions throughout North America. The warehousing group used a similar process to determine and validate unit handling charges (in dollars per pound) for owned and public facilities.

**Modeling Assumptions**

During the data-collection effort, the design team assessed the distribution networks of each business and reached the following conclusions.

The delivery of packaged goods through the BASF distribution network consists of the following steps:

—Produced products go to the plant’s warehouse;
—Products move from the plant warehouses directly to customers, or as replenishments to distribution centers; and
—Products move from distribution centers to demand locations.

The models assume that all the production from a plant moves through the plant warehouse, and production costs include the cost of this movement. The models treat any plant warehouses distributing products produced at other plants as distribution centers. The ports of entry for import products have the same attributes as plants in the models, and the ports of shipment for export products have the same attributes as distribution centers.

For the reasons stated above, a single echelon model is a good representation of the BASF distribution network. Any distribution center serves as a delimiter and connector of two steps in product delivery:

—Replenishment from plants to distribution centers via truckload shipments; and
—Delivery to demand regions from distribution centers, which is based on the expected mix of truckload and less-than-
truckload modes and is represented by effective rates.

We assumed that the proportion of truckload and less-than-truckload shipments going to a demand region is an attribute of each demand region, since it is largely determined by the size distribution of customers in each demand region.

We used an annual, single-time-period formulation because we examined major flow patterns across the distribution network and did not consider operational issues at any single location.

**Mathematical Formulations**

The model developers and decision makers used this three-step development process to clearly understand how each component of the final formulation represented operations in the actual distribution network. We give details of the formulations in the Appendix.

Step 1 was a typical single-echelon distribution model that minimized the sum of the following costs:

—Production costs at plants,
—Transportation costs from plants directly to markets,
—Transportation costs from plants to DCs,
—Transportation costs from DCs to markets,
—Handling and storage costs at the DCs, and
—Penalty costs for any demand shortages.

The model considered the following constraints:

—All production goes to DCs or to customers,
—Demand is fully satisfied with a penalty on any shortages,
—All products received at DCs go to customers within a year,
—The total shipment volume to a DC defines its throughput,
—Throughput for each DC is within its capacity,
—Throughput at each DC for products with special storage requirements is within the DC's capacity for that type of storage (for example, flammable), and
—The sum of special storage capacities used at a DC is within its total storage capacity.

This initial formulation introduced Ms. Carpenter and Mr. Flickinger to model behavior and validated model data, such as transportation rates, facility costs, and distances. Running the model with flows constrained to historical patterns showed the most valuable patterns of transportation flow, the busiest and least active DCs, attainable customer-service levels, and candidate DC locations that were good prospects.

In step 2, we searched for the best DC candidates. We introduced additional constraints that forced the model to select between storage locations in adjoining regions. The economies of scale associated with DC handling and storage costs drove this consolidation. Transportation costs dominate these economies of scale in the overall solution, so we did not explicitly model the cost reductions from consolidating adjacent DCs.

We ran different scenarios with the model to define a final list of potential DC locations. The resulting list included 40 of the existing 86 DCs and additional new locations.

In step 3, we addressed three related questions:

—How do transportation costs and
customer-service levels vary with the number of DCs in the final solution?
—If the distribution network is to have a specified number of DCs, which locations should be selected and who should they serve?
—What products and what product volumes should be stocked at each distribution center?

We added a restriction on the required number of DCs to the model, along with the fixed cost associated with using (opening) a new distribution center. In step 1 and step 2, we omitted these fixed warehouse costs because we used the transportation-flow patterns required to serve customer demand as the primary criteria for selecting candidate DC locations.

**Implementation**

We used MIMI (Manager for Interactive Modeling Interfaces) to implement these models. MIMI is a decision-support tool kit from Chesapeake Decision Sciences, Inc. [MIMI User’s Guide 1994]. It provides a data-manipulation language, configurable graphical user interfaces, optimization, and an expert system. These capabilities allow the creation of data-handling routines, screens, dialogs, model formulations, solution procedures, and reports. MIMI can efficiently model all of the features of a distribution network and its attendant replenishment rules and contractual obligations.

By integrating LP generation and optimization tools with an expert system, we were able to accurately model the unique aspects of the BASF distribution-network problem. We could also create and analyze a wide range of different scenarios easily. The ability to quickly configure GUIs (graphical user interfaces) within MIMI let us easily view distribution flows at an aggregate or detailed level by type of movement, drill down to underlying data or a tabular view of results, and evaluate service levels. The ability to build alternative scenarios quickly by reconfiguring the network or changing data, such as transportation modes or rates, was a key requirement.

**Solution Analysis**

The original network configuration incurred total annual distribution costs of $98.2 million. Of that, $46.1 million came from movements through distribution centers, including both transportation and facility costs. This network configuration delivered 22 percent of the shipped volume on the same day and only 77 percent of the shipped volume by the next day.

We evaluated many alternate configurations. The results clearly indicated an opportunity to both improve delivery times and lower costs (Figure 1). Results of this sort are common when evaluating trade-offs between conflicting objectives. While the solution that provides the optimal value for one objective frequently shows poor performance for the conflicting objectives, usually a number of solutions exist that provide near-optimal values for both objectives [Cohon 1978].

Reconfiguring the DC portion of the network from 86 DCs to 12 facilities simultaneously reduces costs and improves customer service. The reduction in facility and transport costs is $10 million while the volume delivered within one day increases from 77 percent to 90 percent and the level of same-day deliveries improves slightly. Adding another 10 DCs (for a total of 22)
Figure 1: The alternative configurations outperformed the original (86 DCs) in annual operating costs and levels of customer service. The upper line shows the percentage of next-day deliveries, and the lower line shows the percentage of same-day deliveries.

improves same-day deliveries from 22 percent to 30 percent and next-day service to 95 percent but at an increase in total cost over the 12-DC configuration.

Two factors drive BASF's desire to improve customer service. The minimum inventory levels the plant is comfortable with decrease with reductions in transport lead times. This reduces inventory costs for the customer, or for BASF if it manages the inventory at the customer location. Also, faster response to customer orders differentiates a company from its competitors, improving its ability to retain existing customers and attract new ones. Faster response increases revenues.

We could both reduce cost and improve customer service because of the way the existing distribution network evolved. The 86 DCs were not part of a single integrated design. Products moved through 17 different distribution networks (one for each division), which had each grown incrementally.

In addition to supporting the redesign of the distribution network, this model also supports negotiations with contract warehouses by providing a basis for analyzing cost quotes. The model considers the relative costs of the different locations in determining potential use of a candidate warehouse.

Uncertainty is a natural component of any decision-making process, and the level of uncertainty often increases with the number of sources providing data for an analysis. Ideally one addresses uncertainty with probabilities and distribution functions in a fuzzy approach to model development.

The data required to explicitly characterize uncertainty were not readily available. In addition, time was limited and decision makers often need experience with solution approaches that explicitly incorporate uncertainty to develop confidence in the results. We dealt with uncertainty by using sensitivity analysis that relaxed constraints and modified costs. At each step of the development process, the decision maker defined scenarios based on his or her knowledge of the business and evaluated scenario solutions. Friendly user interfaces were critical to the what-if analyses of a variety of different scenarios.

Conclusions

This type of analysis requires both model-solution and model-management capabilities. Solutions come from optimizers that solve strictly linear, special-ordered-set, and MIP formulations. Model management facilitates aggregation strategies, allows rapid generation and analysis of scenarios, and provides navigation guides for the end user through model data and results. Configurable GUIs are
very effective aids for scenario creation and model navigation.

This analysis was the first step in the redesign process. The model remains available to help the business examine any future events that affect distribution, including mergers, acquisitions, new products, or new markets.

BASF learned a number of lessons as it moved to capture the potential benefits identified by the study. It made initial changes to the North America distribution network by opening a new distribution center in the northeastern United States, eliminating several warehouses, and consolidating storage and handing activities. Network modifications resulted in the expected 15-percent improvement in the volume of goods delivered next day.

Several factors complicated BASF’s realizing the accompanying predicted cost reductions. Experience with consolidating different businesses into common DCs showed that implementation and facility-operating costs were higher than anticipated. Customer-service initiatives, such as vendor-managed inventory, just-in-time delivery, and customized labeling, further constrained the movement of goods through the redesigned networks. Recent benchmark studies that considered inventory levels revisited the cost consequences of the redesign effort. Initial results show a six-percent reduction in transport, facility, and inventory carrying costs and a one-time nine-percent improvement in cash flow from a reduction in the working capital tied up in inventory. These studies show that inventory reductions account for much of the total benefits and that additional customer-service requirements reduce potential transport and facility-cost savings.

Another indication of the value of this redesign effort is from “echoes” in Europe [Söder 1997], Scandinavia, and the Far East. BASF adapted the tools we developed for use in those regions, with resulting changes in their distribution networks.

**APPENDIX**

We describe the formulation of the model over the course of the project. Let us first define notation.

**Indices**
- PRD = set of product groups,
- PLA = set of plants,
- DC = set of distribution centers or warehouses,
- MKT = demand markets,
- LOC = all locations, the superset of PLA, DC, MKT, and
- TYP = special types of storage requirements (for example, flammable).

**Data**
- LINK ⊆ [LOC × LOC] all sensible transportation links between plants, warehouses, and demand markets.
- LPD ⊆ LINK will represent links between PLA and DC,
- LDM ⊆ LINK—links between DC and MKT, and
- LPM ⊆ LINK—links between PLA and MKT.

- PDEFF = effective transportation rate per unit shipped between a plant (PLA) and a distribution center (DC),
- PMEFF = effective transportation rate per unit shipped between a distribution center (DC) and market (MKT),
- DMEFF = effective transportation rate per unit shipped between a distribution center (DC) and market (MKT),
- DEMAND = demand for each product in each market,
- HAND = per-unit handling cost for each product and distribution center,
MAXPROD = maximum production capacity at plants,
N = required number of distribution centers,
STOR = per-unit storage costs for each product and distribution center,
TRNCON = per-unit storage requirements for each product,
PRODCOST = production costs for each product,
 PENALTY = penalty costs for unsatisfied demand,
DCCAP = total storage capacity at each DC,
PRDTYP = per-unit special storage requirements for each product and special storage type, and
DCTYP = storage capacity for each special type of storage at each DC.

**Decision Variables**

\[ 0 \leq \sum_{j \in PLA} p_{ij} = \text{MAXPROD}_{ij} = \text{amount of product } i \text{ produced at plant } j, \]

\[ s_{ij} \geq 0 = \text{volume of product } i \text{ shipped from plant } j \text{ to demand market } l, \]

\[ t_k \geq 0 = \text{throughput of product } i \text{ at distribution center } k, \]

\[ 0 \leq c_k \leq 1 = \text{fraction of capacity utilized at distribution center } k, \]

\[ 0 \leq y_{km} \leq 1 = \text{fraction of special type of storage capacity } m \text{ utilized at distribution center } k \text{ and} \]

\[ u_i \geq 0 = \text{(slack) unsatisfied demand of product } i \text{ in demand market } l. \]

In step 1, we used this typical formulation for one echelon distribution model.

\[
\begin{align*}
\min z &= \sum_{i \in PRD} \sum_{j \in PLA} \text{PRODCOST}_{ij} \cdot p_{ij} + \\
& \quad \sum_{(j,k) \in LIP} \text{PDEFF}_{jk} \cdot \sum_{i \in PRD} s_{ijk} + \\
& \quad \sum_{(j,k) \in LPM} \text{PMEFF}_{jk} \cdot \sum_{i \in PRD} s_{ijl} + \\
& \quad \sum_{(k,l) \in LDM} \text{DMEFF}_{kl} \cdot \sum_{i \in PRD} s_{ijkl} + \\
& \quad \sum_{i \in PRD} \sum_{k \in DC} (\text{HAND}_{ik} + \text{STOR}_{ik}) \cdot t_k + \\
& \quad \sum_{i \in PRD} \sum_{l \in MKT} (\text{PENALTY}_{il} \cdot u_{il}),
\end{align*}
\]

Subject to all production is replenished and shipped to customers:

\[
p_{ij} - \sum_{k \in DC} s_{ijk} - \sum_{l \in MKT} s_{ijl} = 0, \quad i \in PRD, j \in PLA,
\]
demand is fully satisfied with penalty on shortages (\( u \) penalty vector):

\[
\sum_{j \in PLA} s_{ij} + \sum_{k \in DC} s_{ijkl} + u_i = \text{DEMAND}_{jl}, \quad i \in PRD, l \in MKT,
\]

all products received at DCs are shipped to customers within a year:

\[
\sum_{j \in PLA} s_{ijk} - \sum_{l \in MKT} s_{ijkl} = 0, \quad i \in PRD, k \in DC,
\]

the total shipment volume to a DC defines its throughput:

\[
\sum_{l \in MKT} s_{ijkl} - t_k = 0, \quad i \in PRD, k \in DC.
\]

DC capacity should meet or exceed the storage area requirements for throughput:

\[
\sum_{i \in PRD} (\text{TRNCON}_{ik} \cdot t_k) - \text{DCCAP}_k \cdot c_k \leq 0, \quad k \in DC,
\]

DC capacity should meet or exceed any storage requirement for special types of products:

\[
\sum_{i \in PRD} (\text{PRDTYP}_{ik} \cdot t_k) - \text{DCTYP}_{km} \cdot y_{km} \leq 0, \quad k \in DC, m \in TYP,
\]

balance of a DC special storage capacity and full capacity:

\[
\sum_{m \in TYP} \text{DCTYP}_{km} \cdot y_{km} - \text{DCCAP}_k \cdot c_k \leq 0, \quad k \in DC.
\]

In step 2, we placed special-ordered-set restrictions on the warehouse capacity vector \( c \) to select between storage locations that were in adjoining regions. We did this because of the economies of scale associated with warehouse handling and storage.
costs. These economies of scale are dominated by transportation costs in the overall solution, so we did not model them explicitly.

In step 3, we added the following restriction on the required number of DCs:

$$\sum_{k \in DC} c_k = N,$$

where $c_k$ were defined as binary variables and $N$ is the required number of warehouses. In addition, we added the following fixed warehouse costs to the objective function:

$$\sum_{k \in DC} FXCOST_k \cdot c_k.$$

References


MIMI Users Guide 1994, Chesapeake Decision Sciences, Inc. (now the Chesapeake Supply Chain Division, Aspen Technology), New Providence, New Jersey.