Interfaces

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Applied Materials Uses Operations Research to Design Its Service and Parts Network

Alper Şen, Deepak Bhatia, Koray Doğan,

Interfaces
Vol. 40, No. 4, July–August 2010, pp. 253–266
ISSN 0092-2102 | EISSN 1526-551X | 10 | 4004 | 0253

informs®
DOI 10.1287/inte.1100.0493

Applied Materials Uses Operations Research to Design Its Service and Parts Network

Alper Şen
Department of Industrial Engineering, Bilkent University, Bilkent, Ankara 06800, Turkey, alpersen@bilkent.edu.tr

Deepak Bhatia
Service Parts Planning, Applied Materials, Santa Clara, California 95050, deepak_bhatia@amat.com

Koray Doğan
Solvoyo Company, Boston, Massachusetts 02110, koray.dogan@solvoyo.com

Applied Materials Inc. is the global leader in nanomanufacturing technology solutions. It has a broad portfolio of innovative equipment, service, and software products and supports its customers worldwide with an extensive service and parts network with more than 100 locations. At the end of 2006, Applied Materials decided to evaluate and rationalize the design of its North American network. It set up a detailed optimization model (including 50,000 parts) to develop a network and distribution strategy. To our knowledge, this is the first large-scale multiechelon network-design model that incorporates safety stock inventory costs while considering the effects of lead time and risk pooling. The company used the model's recommendations to reduce costs while maintaining or improving its service to customers. The recommendations included simplifying the distribution network by consolidating depot locations for specific customers and skipping an echelon for others, thus leading to a projected inventory reduction of $10 million. The company is currently implementing these recommendations and has already eliminated five depots. Applied Materials estimates that during the first year of implementation, inventory reductions of $5.24 million and total savings of $1.1 million can be attributed to these network changes.

Key words: industries; semiconductor; service; facilities-equipment planning; location.
History: This paper was refereed. Published online in *Articles in Advance* April 7, 2010.

This paper describes a network-design project that Applied Materials undertook for its service and parts network in North America. The project's objective was to rationalize and evaluate the existing network in North America and provide the senior management of its service and parts division with alternative designs to reduce supply chain costs such as holding inventory and transportation.

Applied Materials is the world's largest supplier of products and services to the global semiconductor industry. Its products include equipment, service, and software for the fabrication of semiconductor chips, flat panel displays, solar photovoltaic cells, flexible electronics, and energy-efficient glass. In 2007, Applied Materials recorded revenues of $9.73 billion. Its major customers are global semiconductor manufacturers. North America accounts for 20 percent of its orders, whereas 11 percent and 69 percent come from Europe and Asia, respectively.

The semiconductor equipment that the company manufactures is critical to its customers' operations and thus must run at all times. To provide spare parts and service to customers for scheduled maintenance or equipment failures, Applied Materials has an extensive spare-parts distribution network consisting of more than 100 locations around the globe. Three continental distribution centers (CDCs), one each in North America, Asia, and Europe, constitute the backbone of this network and are primarily responsible for procuring and distributing spare parts to depots and customer locations. Various depots in close proximity to customer sites provide faster service to these customers. The company also manages consigned inventory in stockrooms in facilities of its leading customers under agreements called Total Parts Management (Edge 2000).

At the end of 2006, Applied Materials decided to evaluate and rationalize its existing service and parts network and develop a new distribution strategy in North America, which consists of over 50 locations and serves several hundred customer ship-to destinations. Its objective was to determine if it could reduce costs while maintaining or improving customer service. The company considered all costs relevant to its operations: inventory holding, transportation (within the network and outbound to customers), material handling, and warehouse costs. This comprehensive network-design project spanned the company's global inventory-management and logistics functions. We must note here that although the project scope was limited to North America, global demand and requirements had to be considered because the CDC in North America also serves as a global procuring location for a significant number of active parts.

Since the seminal work of Geoffrion and Graves (1974), operations research models have been used extensively to help companies in their network-design and facility-location decisions. However, the usual trade-off in these models is between transportation and facility costs; they do not consider the costs of holding inventory because of demand uncertainty. Examples include Procter & Gamble (Camm et al. 1997), Digital Equipment Corporation (Arntzen et al. 1995), Volkswagen (Karabakal et al. 2000), and Hewlett-Packard (Laval et al. 2005). Despite its importance in service-parts logistics, the network-design problem with inventory considerations (the so-called inventory-location problem) has only recently been studied. The difficulty is that the required inventory at any node in the network is a nonlinear function of the demand during (effective) lead time, which itself is a decision variable in the problem. A node's lead time (i.e., the first component of the demand during lead time) depends on the upstream node to which it is assigned; the demand it faces (i.e., the second component of the demand during lead time) depends on which downstream nodes are assigned to it; both are endogenously determined in the model. The literature suggests several approaches to address these nonlinearities. For example, Daskin et al. (2002) use a Lagrangian-based approach, Shen et al. (2003) employ column generation, Erlebacher and Meller (2000) use a grid-based approach, and Candaş and Kutanoğlu (2007) approximate the fill-rate function using a step function. All these models except the last one are stylized models because they consider only a single product in the network. The approach in Candaş and Kutanoğlu (2007) may potentially be used for problems with a limited number of parts (e.g., the data set in their computational study has four parts); however, because our problem included tens of thousands of active parts, we needed a new, scalable approach.

Understanding the difficulty of the problem and the unavailability of any commercial software, Applied Materials decided to form a team that could develop a new solution. The core team (the authors of this paper) consisted of a faculty member at a research university, the founder and chief architect of a new supply chain-planning software company (Solvoyo), and an inventory manager and OR/MS practitioner at Applied Materials. Solvoyo's initial product, planLM (Solvoyo 2010), was designed as a platform (with a common data model) for various supply chain-planning activities and analysis. At the start of the project, planLM already had a solution called Supply Demand Optimizer, which provides decision support for strategic network-design decisions that involve transportation and facility costs. The team decided that it would extend the functionality of this solution to incorporate the cost of holding safety inventory. Although the major task in the project was to develop a mathematical model and a software solution for the inventory-location problem, the team spent a significant amount of time and effort on data collection; therefore, it required the participation of various Applied Materials employees throughout the project. It kicked off the project in January 2007 and successfully completed it in July 2007; during this period, it developed and implemented a new, scalable solution for the inventory-location problem for Applied Materials' service and parts network in North America. The results showed a potential of $10 million in inventory savings by consolidating depot locations for specific customers and skipping an echelon for others without sacrificing customer service levels. The company is currently implementing the project's recommendations. Based on these recommendations, it has eliminated five depots in the first year, achieving inventory reductions of $5.24 million and cost savings of $1.1 million.

The Service and Parts Network in North America

Applied Materials provides spare parts to several hundred customer locations in North America. The network to serve these customers consists of a CDC, depots that Applied Materials owns and operates, and stockrooms (consignments) that it manages in customer facilities. Its 50,000 active parts consist of consumables and nonconsumables with large cost variations. All parts are procured at the CDC, and the supplier lead times range from 1 to 270 days.

Customer orders are either emergency orders or regular orders. Emergency orders that result from equipment failures and are critical for customers are usually satisfied from consignments (if such an agreement exists) or from depots. The CDC also provides a second level of support for emergency orders if they cannot be satisfied immediately from consignments or depots. Regular orders occur when customers request spare parts for use in their scheduled maintenance activities. The CDC is usually the primary source for meeting these demands; however, local depots can also be used to meet the demands of some customers.

Both emergency and regular orders go through an order-fulfillment engine that searches for available inventory in different locations according to a search sequence that is specific to each customer. Emergency orders must be satisfied immediately (their request date is the date of the order creation), whereas regular orders can be satisfied at a future date. A depot may be required to handle both emergency and regular demands simultaneously and from a variety of customers. The orders that the consignments and depots place to replenish their stocks are called replenishment orders and are handled by the CDC. The CDC also handles emergency and regular orders from external customers. These are typically the result of unsatisfied demand at the depots or consignments. However, for a minority of customer locations that are in close proximity to the CDC, it could be the first source. Figure 1 depicts the service network in North America.

Service-level requirements vary for each customer site and type of order. Service levels at customer sites with consignment agreements are structured well and

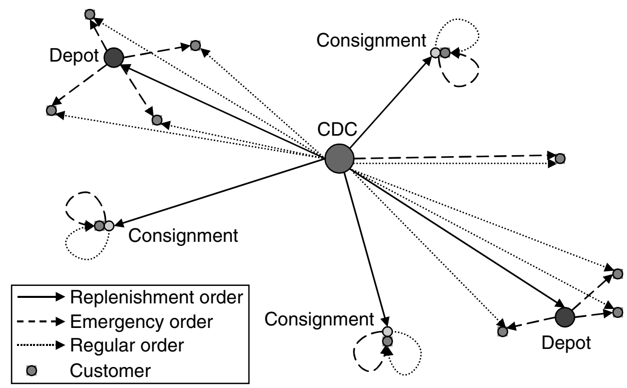


Figure 1: The service network in North America consists of a CDC, depots, and consignments.

enforced through contracts. Required service from depots and the CDC are usually mandated through targets that the executive team sets—not by contracts. Service levels are measured in terms of fill rate, i.e., percentage of orders satisfied from stock. Because each customer site is located within prespecified (time) windows from its assigned depot and the CDC, fill rates at these two levels can be translated into time-phased fill-rate measures for each customer. A global inventory-management team at the CDC does inventory planning. For depots and consignments, this team suggests inventory levels and policies; however, local teams are responsible for managing inventory at these locations.

Costs at the CDC and the depots include material handling fees for each inbound and outbound move and warehouse fees based on actual space usage. The transportation function in each location is carried out by multiple carriers, including specialized courier services, parcel companies, airfreight forwarders, and other international logistics companies. Each carrier charges Applied Materials per transaction, based on the pickup time, urgency, weight, and destination of the order.

The Mathematical Model and Assumptions

To develop a network and distribution strategy, the core team needed to formulate a mathematical model that would consider various constraints in each scenario and determine the optimal network configuration that would minimize the total transportation, inventory holding, and material-handling costs in North America. Note that the scope of our study was the service network in North America, particularly the number and locations of depots. Because the consignments are colocated with customers, their number and locations were not decision variables in the study. Similarly, the CDC location was not part of the study and was taken as fixed. Because the consignments are replenished only through the CDC and the total demand that the CDC must serve does not change when the network changes, we could consider the inventory at the CDC and the cost of transportation between the CDC and the consignments separately from the optimization model. However, note that some of the scenarios we studied require the service level (fill rate) at the CDC to be increased, thus leading to an increase in inventory at the CDC and a decrease in inventory at the consignments. We performed these calculations externally and added these costs to the costs obtained by running the mathematical program we describe in the appendix. For the inventory for which the CDC has responsibility, we had to incorporate demands from locations outside of North America in our model. Therefore, we created a dummy customer location and aggregated all non-North American demand for North America-sourced parts in this location. For the inventory carried at the consignment locations, we set an average service level across different consignments; in reality, each customer could have customized consignment contracts with various terms, including different service-level definitions and commitments.

Given the scope of the study, our objective was to develop a model that would determine the number and the locations of the depots to minimize inventory holding, transportation, and material-handling costs. To obtain a scalable mathematical model with readily available data requirements, we had to make several assumptions. Because the model and its results would be used for strategic purposes (and not for daily operations), we agreed that these assumptions, which we list below, would not hinder the validity of the recommendations.

Assumption 1. *All parts at a location have the same service level. The service-level commitments and targets that Applied Materials uses are typically not at the parts level. Service levels are defined and measured based on the total demand of all parts that a customer uses. In theory, Applied Materials may offer different service levels for different parts* (*e.g., a high service level for low-cost parts and a low service level for high-cost parts*), *thus averaging a customer's service-level commitment. However, we agreed that this level of detail would not alter our model's results and recommendations. We assume that each part at a location has the same service-level target. We selected the service levels for the baseline model based on historical performance and customer requirements* (β_T *for the consignments,* β_D *for the depots, and* β_C *for the CDC; the notations are described in the appendix*).

Assumption 2. *There is no differentiated service at the CDC. As we explained above, the CDC has multiple roles in the network*: (1) *to satisfy regular demand for nonconsignment customers*, (2) *to replenish consignments*, (3) *to replenish depots, and* (4) *to satisfy secondary emergency demand in the network. Although using rationing and offering differentiated service to different order types is possible in theory, this requires order-fulfillment software that is capable of reserving material for different types of demand*; *at the time of project implementation, Applied Materials did not use software with this functionality. Therefore, we assumed that all demand types receive the same service at the CDC.*

Assumption 3. *Weight distributions and rates will not change by reconfiguring the network. Most shipments, either for replenishment within the network or satisfaction of customer demand, involve multiple parts. Because the transportation cost for a shipment is usually a nonlinear function of the weight of the shipment, the average transportation costs per unit of weight depend on the distribution of shipment weights. In the* Data Collection *section, we provide a method to estimate these weight distributions and calculate the shipment costs per unit of weight for each origin and destination pair. We assume that a change in the network will not cause the weight distributions, and thus the transportation costs, to change. We also assume that a network change will not result in a change in carrier rates.*

Assumption 4. *Unsatisfied demand at a depot will be expedited only from the CDC. When a part has failed and the depot or the consignment to which the customer is assigned is out of stock for this part, order-fulfillment software searches the entire North American network* (*excluding some consignment locations for other customers*) *for* *the part. In the model, however, we assume that the part is expedited from the CDC only (perhaps after procuring it from the supplier), and transportation costs are incurred accordingly. This means that the model would not position inventory at a depot for unsatisfied orders at other locations and would not account for the delivery costs of customer orders expedited from elsewhere.*

Based on these assumptions, and after several meetings with the business users and the implementation team, we developed a mixed-integer linear programming model. In the appendix, we discuss the model, its features, and its size in terms of the number of variables and constraints. Note again that the model does not include the inventory at the CDC and the shipments between the consignments and the CDC. The related costs are calculated externally.

An important feature of our optimization model is the approach we use to calculate the inventory necessary to maintain a service level at a depot, given the customers that are assigned to that depot. In our model, RI_{jp} is the function that corresponds to the required average inventory to maintain a service level (fill rate) of β_D at depot j for part p. Clearly, this is a nonlinear function of the mean and the variance of demand that is allocated to that particular depot. The function also depends on the transportation lead time l_j, the service level (fill rate) at the CDC (β_C), and the supplier lead time L_p of the part (we use the approximation in the METRIC model; Sherbrooke 1968) and assume that the replenishment lead time is equal to the effective lead time for the depot: $l_j + (1-\beta_C)L_p$. RI_{jp} is also a function of the replenishment policy (r_p) that is used for part p. Applied Materials uses two types of policies: $(S-1, S)$ and (R, Q). If the replenishment policy is $(S-1, S)$, then RI_{jp} consists only of safety stock. If the replenishment policy is (R, Q), then RI_{jp} includes cycle inventory as well as safety stock.

To address the nonlinear RI_{jp} function in the objective function of the optimization model, we use the approach that Figure 2 illustrates. First, we find the maximum variance that can be allocated to each depot for each part. This corresponds to the variance of the total emergency demand originating from nonconsignment customers in North America because one particular depot could potentially serve all such customers. Then, we find the safety stock requirement for this maximum variance and assume a linear cost function that passes from 0 and this value. Solving this problem with the linear cost function at iteration 1 gives us solution 1. At iteration 2, we find

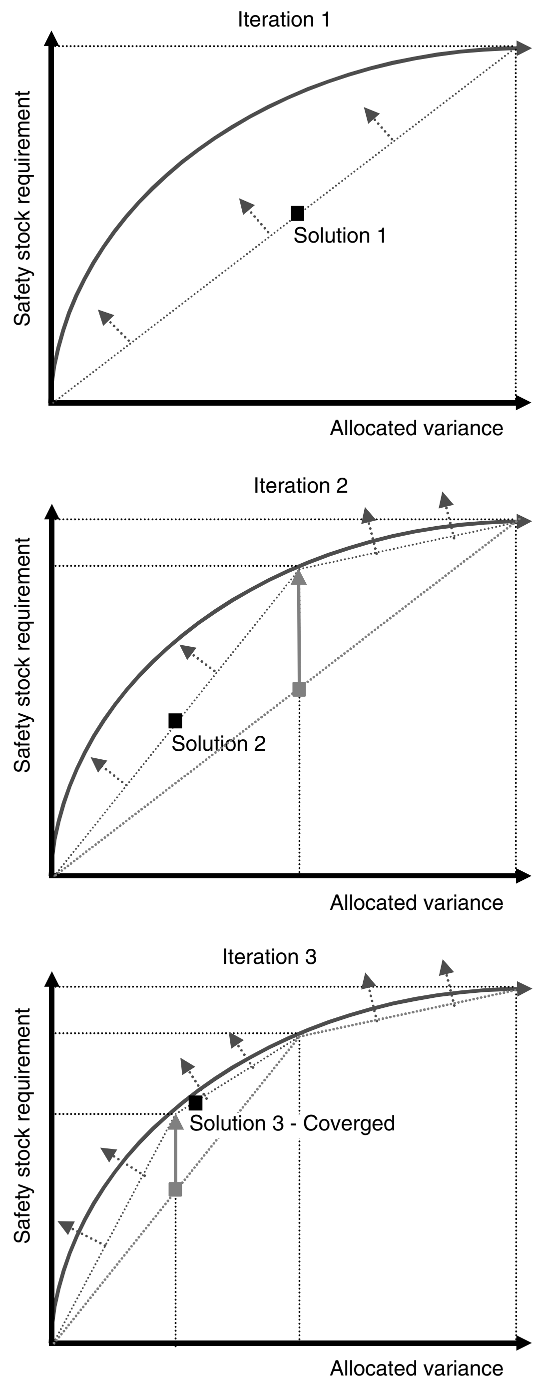


Figure 2: The nonlinear objective function is addressed using an iterative procedure.

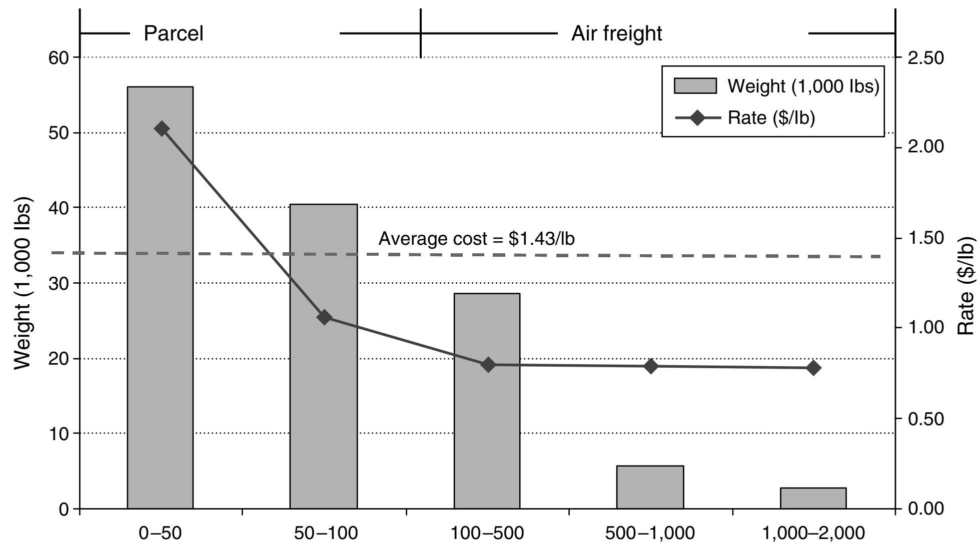
the safety stock requirement that corresponds to the allocated variance in solution 1 and approximate the function using a piecewise-linear function, as the second graph in Figure 2 shows. Solving the problem with this piecewise-linear function by incorporating binary variables gives us solution 2. At each succeeding iteration, we find the difference between the safety stock requirement using the piecewise-linear approximation and the true safety stock required for that allocated variance. If the difference is more than a predefined threshold, we continue and incorporate another piece to the approximation. If the difference is less than this threshold for all depot-part combinations, we stop and report our last solution as the final solution. In general, the threshold value can be determined based on the user's trade-off between the computation time and the quality of the approximation of the RI_{jp} function. In our implementation, we use a threshold value that is equal to 0.01 percent of the safety stock requirement in the previous iteration. With this threshold, the computation times were acceptable and the resulting inventory costs were accurate. An alternative approach to our iterative algorithm is to fit a piecewise-linear function for the required safety stock upfront, with a predetermined number of pieces. Initially, we tried this approach; however, because all nonconsignment, emergency demand in North America can be potentially assigned to one particular depot, the piecewise-linear functions required many pieces for a reasonable approximation of the required nonlinear inventory curve. This led to introducing too many binary variables upfront; thus, this was prohibitive if we were to successfully solve the problem using commercial solvers.

Data Collection

In this section, we explain the methods we used to estimate four critical data elements.

Transportation costs from depots to customer locations. Estimating the cost per unit between depots and customer locations was not a trivial task. First, we created shipments from a depot to customer locations based on specific customer orders that could have different line items and thus different weights. Therefore, shipment costs, and thus cost per pound (or cost per unit), vary for each customer even when the delivery time is fixed. Second, based on a shipment's weight and delivery-time requirement, Applied Materials chooses a transportation service provider based on the provider's cost and availability at the time of the shipment. To address the first issue, we had to determine a weight distribution of shipments to customer locations. (Note that this weight distribution should be independent of the origin, i.e., depot location.) To develop the weight distribution, we used the complete shipment history for the 12-month period prior to the implementation. To estimate the shipment cost of an emergency order, we selected two representative carriers based on a carrier-routing guide at the service and parts division: a major parcel company for shipments weighing less than 100 pounds and an airfreight forwarder for shipments weighing more than 100 pounds. The rates the parcel company charges are based on zones (each origin and destination zip code pair falls in one of 15 zones), type of service (priority overnight, standard overnight, next day, second day, etc.), and weight (for each pound break). We determined the corresponding zone of each possible depot location and customer pair. Because the shipments from depots to customer locations are all emergency orders, we used priority overnight service. We estimated the actual weight distribution of shipments below 100 pounds using two intervals: 0–50 and 51–100. We then calculated the average shipment weights and used them to select the rates from rate tables. We used a similar approach for the airfreight forwarder, which uses a similar zone system and an all-units pricing scheme based on five weight breaks. Finally, we combined these to calculate the average transportation cost for each potential depot site, customer location, and part number.

Figure 3 shows a set of shipments from a potential depot site to a customer site (both the shipments and the rates are illustrative and do not represent actual customer data). The average weight for shipments in the 0–50 and 51–100 intervals are 10.67 and 71.27 pounds, respectively. Therefore, we select the rates corresponding to 11 and 72 pounds, respectively, from the parcel provider's rate tables. For shipments of more than 100 pounds, we use the weight breaks of the airfreight forwarder and consider the total weight



Range	0–50	50–100	100–500	500–1,000	1,000–2,000
Total weight	56,000	40,480	28,600	5,646	2,806
Average weight	10.67	71.27	185.71	564.68	1,403
Rate ($/lb)	2.11	1.06	0.80	0.79	0.78

Figure 3: Transportation costs are estimated using historical weight distributions and representative carriers.

of shipments that are in these weight ranges. This calculation results in an average freight rate of $1.43 per pound. We calculate the unit freight rate for each part using standard part weights (available in Applied Materials' Enterprise Resource Planning system) plus a packaging factor.

Transportation costs from the CDC to depots and customers. We followed a similar process to calculate the transportation costs from the CDC to depots and customers. One major assumption we made was that network changes will not cause changes in the weight distributions used in calculating these costs. A more accurate method would be to calculate the weight distributions of incoming replenishment orders to a depot for each possible scenario (i.e., for each set of customers that are assigned to that depot). However, calculating the weight distributions for all scenarios would be computationally prohibitive because we had 10–50 depot locations and 200–500 customer sites. We understood that relaxing the constant weight-distribution assumption would work in favor of the study's general recommendations, which suggest consolidating depot locations because consolidation will lead to heavier shipments—and therefore shipments that would be cheaper per unit of weight—from the CDC to depots. However, it was possible, in theory, that the constant weight-distribution assumption might have an impact on our recommendation of which depots to close. However, as we show in the *Model Validation, Recommendations, and Results* section, transportation costs are heavily dominated by inventory holding, and material-handling costs in our implementation. Therefore, we believe that these costs were the primary drivers of the decision regarding which depots to close; the assumption had little or no effect.

Demand. For demand history, we used the year just prior to the implementation to determine the demand distributions at the part, customer, and order-type levels. This required classifying parts as slow moving, medium moving or fast moving—the classification that Applied Materials uses in its current operations. Its classification is based on the number of customer requests within the previous two years, the time between customer requests, and the time since the last request. For slow-moving parts, we assumed that demands are distributed using a Poisson distribution with a mean annual rate equal to the total demand in the previous year. For medium- and fast-moving parts, we assumed that the annual demands are distributed normally—again with a mean equal to the total demand in the previous year. We calculated the standard deviation by using a constant coefficient of variation, which was the average value for these parts. Our assumptions were consistent with the approach that the inventory-management team took to make its operational inventory decisions.

Replenishment policy. We assumed that the replenishment policy is of type $(S-1, S)$ for slow-moving parts and type (R, Q) for medium- and fast-moving parts. To calculate the Q, we used a standard economic order quantity (EOQ) approach. Again, these assumptions were consistent with the approach that the inventory-management team took for its actual operational decisions. The replenishment-policy choice impacts the amount of inventory required for each part at each depot (RI_{jp} is defined in *The Mathematical Model and Assumptions* section) and the inventory required for each part at the CDC and the consignments (which is calculated outside of the model). That is, the required (average) inventory for a slow-moving part is calculated based on the $(S-1, S)$ policy and only includes safety stock; however, for a medium- or fast-moving part, it is calculated based on the (R, Q) policy and includes safety and cycle inventory. The transportation costs, however, are calculated based on the historical weight distributions of replenishment orders (which may consist of multiple parts) and Assumption 3 in *The Mathematical Model and Assumptions* section. For the EOQ calculations, we used the same order cost for all parts and locations (the inventory-management team used a similar approach for its operational decisions). Annual inventory holding cost for a part was calculated by multiplying the part's standard cost by the annual inventory holding-cost rate. Because our customers usually operate their equipment at full capacity, the spare-parts demand did not exhibit any seasonality; in addition, we saw no apparent trend or seasonality in the one-year demand history. Therefore, given a depot location and its customer assignment, the model generates only one set of parameters for each part's replenishment policy.

In addition to these four items, the core team worked extensively with personnel from the logistics and marketing teams to gather data such as material-handling and facility setup costs, inventory holding cost rates, and customer-service requirements. All data templates were populated with the necessary data, and the final data were approved by the entire team.

Software Solution

We implemented our mathematical model and the solution using the Solvoyo planLM Supply Chain Planning and Analysis platform. Solvoyo planLM provides the graphical user interface, application server infrastructure, and data interface to a supply chain data model that includes all the data requirements for the mathematical model given in the appendix. Therefore, no changes to the data model were necessary. To speed up our development, we implemented the mathematical model as an extension to the planLM Supply Demand Optimizer application, which is designed for strategic network-design problems as well as tactical master planning. Although planLM provides interfaces to most third-party solver engines, we used ILOG CPLEX 11.0 on a dual-core 64-bit Windows XP platform. The run times for solving the mixed-integer linear program (see the appendix) varied between two and eight hours, depending on the specific scenario, and used a maximum of 23 GB of memory. We found planLM to be particularly useful in enabling us to quickly create and compare the scenarios and run sensitivity analyses. Its user interface and other reporting features were also useful in speeding up our analyses. Figure 4 illustrates a planLM screen that shows comparisons of different scenarios.

Model Validation, Recommendations, and Results

After gathering the data and using planLM to run the model, we began the validation task by establishing a baseline scenario corresponding to the current depot locations and customer-depot assignments. We used the most current service-level performance numbers on depots (β_D) and the CDC (β_C) and a constant β_T service-level target for the consignments,

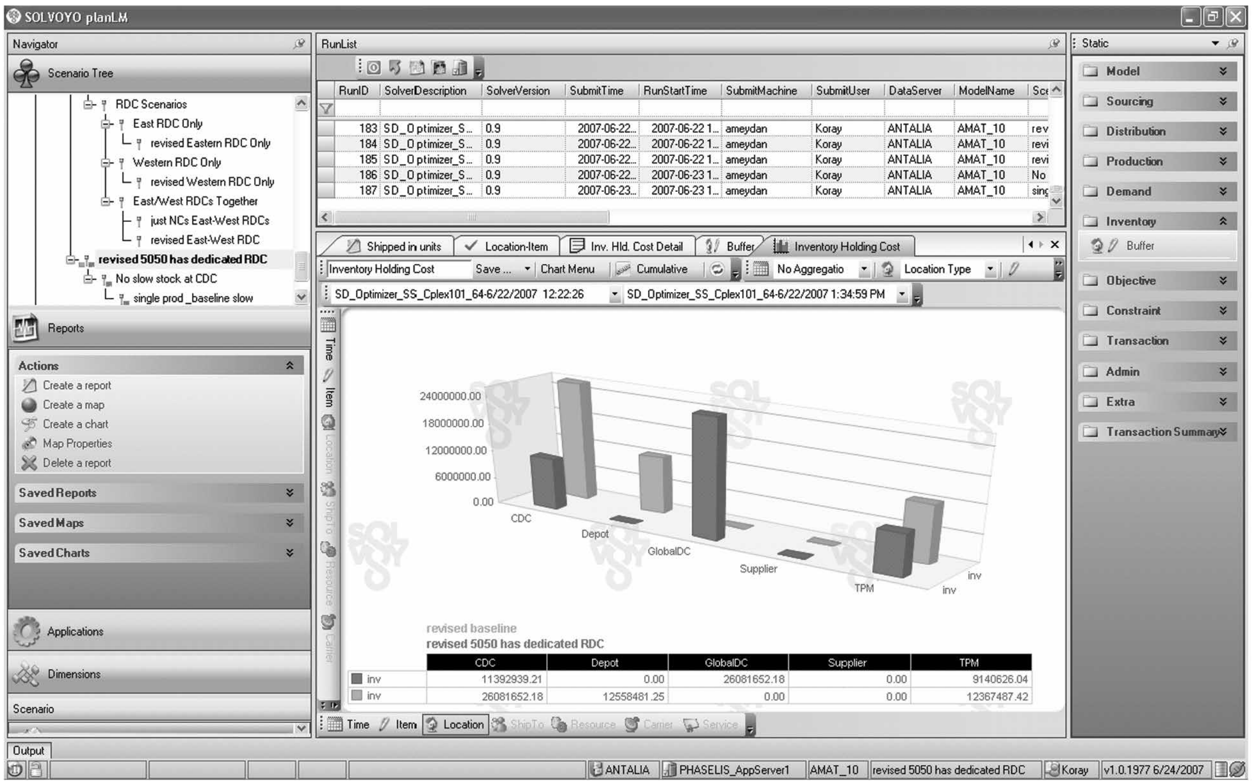


Figure 4: The user interface of planLM speeds up the analyses.
Source. **Screen shot of SOLVOYO planLM Planner Interface. Used with permission.**

and we then validated the results of the baseline scenario against the actual inventory, transportation, and material-handling costs. The model predicted the total costs to be 9.8 percent more than the actual total costs and generated values for the transportation and material-handling costs that were close to actual costs. The inventory carried at the CDC and the consignments also closely matched the actual numbers (both were within 2 percent). The model suggested keeping 29.2 percent more inventory in depots, primarily because it forces a service level of (β_D) for each part; however, the field objective was to maintain this service level at a location level (see Assumption 1 in *The Mathematical Model and Assumptions* section). The differences were acceptable because the model would be used only for strategic planning purposes.

In Figure 5, we illustrate different cost components as a percentage of total cost for the baseline scenario. The transportation costs (2.56 percent) are heavily dominated by the material-handling costs (48.46 percent) and inventory holding costs (48.98 percent), which include costs of inventory at all locations—CDC, depots, and consignments. This shows that the current network is designed to serve customers from locations in close proximity, leading to minimal transportation costs. This is accomplished by establishing an echelon of many depots to serve nonconsignment customers; these depots hold a substantial amount of decentralized inventory and incur heavy material-handling costs. The dominance of inventory holding costs over transportation costs, as shown in the figure, is typical of service-parts logistics; the majority of items are slow moving, inventory turnover rates are small, and obsolescence and scrap rates are high. According to a benchmark study of nine companies in the computer industry, transportation costs constitute only 8.4 percent of their relevant service-function costs (Cohen et al. 1997). As Figure 5 shows, the transportation costs are even less significant because parts are less bulky (the average weight is 6.22 pounds), and we

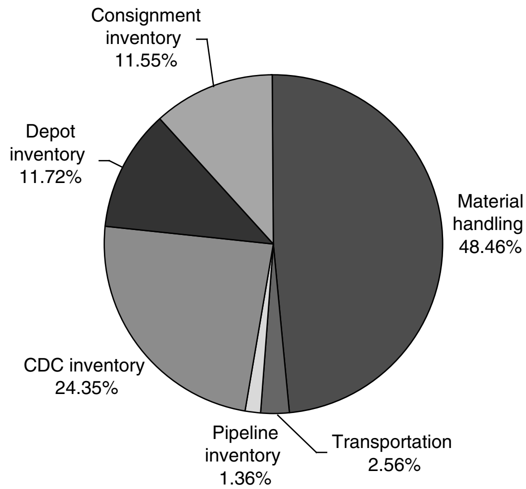


Figure 5: The transportation costs are heavily dominated by material-handling and inventory holding costs in the baseline scenario.

exclude the inbound freight costs (from the supplier to the CDC) in our model.

Some costs shown in Figure 5 are unavoidable. For example, the CDC in North America replenishes all parts for all locations and customers in North America and most parts for the rest of the world. Therefore, unless the service levels are changed, the inventory at the CDC cannot be changed by reconfiguring the network, and material-handling costs at the CDC cannot be reduced. Similarly, service levels and consignments are contractual obligations for Applied Materials. Therefore, unless the service level for the CDC is changed, reducing inventory costs for consignments by changing the service network is impossible. Therefore, we identified reducing the inventory carrying costs and material-handling costs by reconfiguring the network at the depot level to be a major opportunity for reducing costs. We investigated if and how we could accomplish this and also consider transportation costs by using the five scenarios described as follows.

Scenario 1: *Optimized baseline.* In this scenario, we optimize the depot customer assignments. We allow the solver to pick the CDC as a source for satisfying the emergency demand of a customer, and we keep the CDC service level at β_C. Thus, this scenario can potentially reduce the service levels to specific customers if they are assigned to the CDC.

Scenario 2: *Increased service level at the CDC.* We increase the service level at the CDC to β_D, enabling an unchanged service level for customers assigned to the CDC for emergency demand. However, because the CDC provides nondifferentiated service for different types of demand, as we explained above, this also leads to better service for other regions and regular orders. In addition, service levels for replenishments to consignments also increase, possibly leading to reductions in consigned inventory.

Scenario	Transportation	Inventory holding	Material handling	Total costs	Service level
0	2.56	48.98	48.46	100.00	99.0
1	2.87	37.13	47.32	87.33	99.0
2	2.87	44.55	47.32	94.74	99.0
3	2.87	44.06	47.32	94.26	99.0
4	2.87	44.76	47.32	94.96	99.0
5	4.11	47.38	53.68	105.17	99.0

Table 1: We developed a baseline scenario and five new scenarios to investigate the impact of a network reconfiguration.

Scenario 3: *Separate stock for emergency demand at the CDC.* We created a new stock location at the CDC to reserve separate stock for emergency demand, providing a β_D service level for those customers who are assigned to the CDC. The rest of the network operates as it did previously.

Scenario 4: *Regional distribution center for North America.* We introduced a new distribution center for North America to satisfy only the North American demand. This regional distribution center (RDC), which is a virtual location at which the CDC is located, will provide β_D to all types of demand in North America.

Scenario 5: *Two RDCs for North America.* We established two RDCs in North America, one on the east coast and one on the west coast. These RDCs will provide β_D and will be replenished by the CDC.

In each of these scenarios, the customer-depot assignments are decided by the mathematical model without any restrictions. A customer may also be assigned to the CDC or an RDC rather than a depot. Table 1 shows the results of the solution for these scenarios. All costs are represented as a percentage of the total costs in the baseline scenario (Scenario 0).

Our results show that all scenarios, except Scenario 5, result in savings in total costs. These savings are largely in inventory holding, and partly in material-handling, costs. Inventory holding costs at the depots are reduced because of risk pooling via consolidation of depots and (or) assigning customers to distribution centers (CDC or RDC). When a customer is assigned to a CDC or virtual RDC, material-handling costs for that customer are also eliminated at the depots. Moreover, the increases in transportation costs are comparatively small.

Based on these results, we initially selected Scenario 2 as our recommended solution. Although Scenario 1 gives a lower total cost, this alternative will potentially reduce customer service levels for customers who are assigned to the CDC. Scenario 3 also leads to slightly lower total costs; however, we decided that the benefits of increasing the service level at the CDC for all order types globally outweighs this difference in cost. Before presenting Scenario 2 to Applied Materials management as our recommendation, we ran one final iteration because of the increase in service time that a specific set of customers would see under the new network design. In their existing system, these customers received support for their emergency orders within four hours by a special ground service that the specific depots to which they were assigned provided. In the new network, the customer parts might be replenished from the CDC using a "priority overnight" service provided by a parcel company, hence increasing the time to fulfill the customer's order. Applied Materials decided to maintain the status quo for these customers, at least during a transition period. We set up a new scenario and forced the solution such that it would use the existing depot–customer assignments for these customers. Scenario 6 in Table 2 shows the results (again, all costs are represented as a percentage of total costs in Scenario 0). Overall, Scenario 6 was expected to result in a 2.88 percent savings in total costs. It recommends that the company close six depots and serve their customers from the CDC—but at an increased service level. This would result in eliminating the inventory at these depots, reducing the inventory at the remaining sites (at consignments and open depots because of the increased service level at the CDC), and increasing the inventory at the CDC. Total inventory would be reduced by 4.55 percent. In addition, material-handling costs would be reduced by 1.9 percent because of shipping parts directly from the CDC. However, the transportation costs were expected to increase by 10.5 percent because customer orders would be sent over longer distances.

Scenario	Transportation	Inventory holding	Material handling	Total costs	Service level
6	2.83	46.75	47.54	97.12	99.0

Table 2: The recommended scenario (Scenario 6) reduces costs by 2.88%.

Finally, we analyzed the sensitivity of the results with respect to two important parameters in the analysis. The first was the inventory holding-cost rate. The core team was reminded by Applied Materials that there could be important changes in the financial opportunity cost, which was a big portion of the inventory holding-cost rate. The second was the material-handling cost per unit of shipment. Changes are also likely for this parameter.

Our extensive sensitivity analysis, for which we used planLM, showed that although changes in these two parameters impact the magnitude of the savings, the ordering of our alternatives (i.e., in terms of total cost) would not change. For all scenarios, Table 3 shows the sensitivity of the results to changes in the inventory holding cost rate and the material-handling fee per shipment, where i and f represent the values that we used in the study for these parameters. Our analysis showed that the savings would still be significant using Scenario 6 when the parameters change (percentage savings of Scenario 6 over Scenario 0 ranges from 1.80 to 4.06 percent). Sensitivity analysis also showed that the total inventory (and where it is carried) in Scenario 6 does not change unless the inventory carrying-cost rate is significantly lower

Inventory carrying-cost rate	Handling fee	Total cost of scenario 0	1	2	3	4	5	6
i	f	100.00	87.33	94.74	94.26	94.96	105.17	97.12
$0.8i$	f	90.20	79.89	85.83	85.44	86.00	95.69	87.77
$0.6i$	f	80.41	72.47	76.92	76.63	77.05	86.22	78.42
$0.4i$	f	70.61	65.04	68.01	67.81	68.09	76.74	69.07
$0.2i$	f	60.82	57.62	59.10	59.00	59.14	67.27	59.72
$0.04i$	f	52.98	51.75	52.06	51.95	51.98	59.69	52.34
$0.4i$	0	22.15	17.41	20.40	20.49	20.77	23.06	21.25
$0.2i$	0	12.36	9.94	11.44	11.68	11.82	13.59	11.86

Table 3: Ordering of scenarios is insensitive to changes in holding-cost rate and material-handling cost.

than the rate that we used in this project and (or) the material-handling fee is completely eliminated.

We presented our project results, including the recommendation to change the network, to the senior management team of the Applied Materials service and parts division on June 26, 2007. The management team approved our approach and provided positive feedback about our recommendation. The project was officially declared complete in July 2007.

Applied Materials started implementing our recommendations in early 2008. In 2008, it eliminated three depots in North America; in 2009, it closed two additional depots. The core team carried out a thorough value analysis in July 2009 to measure realized savings because of restructuring the network in North America. The analysis involved comparing the costs and inventory before the changes were initiated with the costs and inventory of the current quarter. Note that the changes in business conditions such as decline in demand might have also led to reductions in inventory holding, transportation, and material-handling costs. To fairly assess the savings resulting from the network restructuring, the impact of these changes are eliminated from the analysis. The results show that of the total inventory reduction since closing the depots, $5.24 million can be attributed to network redesign recommendations from this project. The analysis estimates that the savings in inventory carrying, material-handling, and transportation costs so far are $1.1 million. The current level of restructuring will also lead to additional savings of $1.38 million annually in the future. These savings estimates are conservative because the impact of higher service levels at CDCs on inventory levels at consignments and regions outside of North America are not considered. The value analysis also showed that there has been no impact on service levels because of these changes in the distribution network. A consolidated network also brings other benefits to Applied Materials; these include, but are not limited to, a reduction in network complexity potentially resulting in increased visibility within the network and an agile supply chain.

Conclusion

This paper describes a strategic network-design project that was carried out for the service and parts division of Applied Materials. We developed a novel approach that simultaneously considers inventory and logistics costs. This approach extends the functionality of an existing software solution to solve a large-scale problem and evaluate alternatives. To our knowledge, this is the first multiechelon network-design solution of this scale that incorporates safety stock inventory costs and also considers lead-time and risk-pooling effects. The senior management team at Applied Materials approved our approach and our study's recommendations, which included simplifying the company's network by consolidating depot locations for some customers and skipping an echelon for others. We projected that these measures would lead to significant inventory reductions. Applied Materials is currently implementing the recommendations and has already eliminated five depots from its parts and service network in North America. Moreover, it estimates that $5.24 in inventory reductions during the first year of implementation can be attributed to these network changes.

Many companies face network-design problems for which safety stock inventory holding costs constitute a large portion of the total costs. Item-level large-scale models may be necessary to solve these problems and evaluate different network strategies. We believe that such companies can benefit from the approach and the solution we have described in this paper.

Appendix. Mathematical Model

The notations we used are listed below.

K: set of customers, indexed by k.

J: set of potential depot sites, indexed by j.

P: set of parts, indexed by p.

d_{kp}: mean demand of customer k for part p.

r_p: replenishment policy for part p.

σ^2_{kp}: variance of demand for customer k for part p.

h_p: inventory holding cost for part p per unit per year.

c_{jp}: unit transportation cost from the CDC to depot j for part p.

c_{jkp}: unit transportation cost from depot j to customer k for part p.

c_{0kp}: unit transportation cost from the CDC to customer k for part p.

m_j: material-handling cost per unit at depot j.

m_0: material-handling cost at the CDC.
l_{jk}: delivery time from depot j to customer k.
l_j: delivery time from the CDC to depot j.
L_p: supplier lead time to the CDC for part p.
f_j: fixed operating costs of depot j.
C_j: flow capacity of depot j.
β_C: service level for the CDC.
β_D: service level for depots.
RI_{jp}: required inventory for depot j and part p as a function of allocated demand, allocated variance, service level, effective lead time, and replenishment policy.

The following variables are the decision variables of the problem.

v_{jp}: allocated variance at depot j for part p.
x_{jkp}: amount of flow from depot j to customer k for part p.
x_{0kp}: amount of flow from the CDC to customer k for part p.
z_{jp}: amount of flow from the CDC to depot j for part p.

$$q_{jk} = \begin{cases} 1 & \text{if customer } k \text{ is assigned to depot } j, \\ 0 & \text{otherwise.} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if depot } j \text{ is used,} \\ 0 & \text{otherwise.} \end{cases}$$

The mathematical program can be written as

$$\begin{aligned} \min \sum_{p\in P}\sum_{j\in J}\sum_{k\in K}(c_{jkp}+m_j)x_{jkp} + \sum_{p\in P}\sum_{k\in K} c_{0kp}x_{0kp} \\ + \sum_{p\in P}\sum_{j\in J}(c_{jp}+m_0)z_{jp} + \sum_{p\in P}\sum_{j\in J\cup\{0\}}\sum_{k\in K} h_p l_{jk} x_{jkp} \\ + \sum_{p\in P}\sum_{j\in J} h_p l_j z_{jp} + \sum_{p\in P}\sum_{j\in J} h_p RI_{jp} \\ \cdot\left(\sum_{k\in K} d_{kp}q_{jk}, v_{jp}, \beta_D, l_j + (1-\beta_C)L_p, r_p\right) \\ + \sum_{j\in J} f_j y_j \end{aligned} \tag{1}$$

$$z_{jp} - \sum_{k\in K} x_{jkp} = 0 \quad \text{for all } j\in J,\ p\in P, \tag{2}$$

$$-v_{jp} + \sum_{k\in K} \sigma^2_{kp} q_{jk} = 0 \quad \text{for all } j\in J,\ p\in P, \tag{3}$$

$$\sum_{j\in J} x_{jkp} + x_{0kp} - d_{kp} = 0 \quad \text{for all } k\in K,\ p\in P, \tag{4}$$

$$x_{jkp} - \beta_D d_{kp} q_{jk} \le 0 \quad \text{for all } j\in J,\ k\in K,\ p\in P, \tag{5}$$

$$q_{jk} - y_j \le 0 \quad \text{for all } j\in J,\ k\in K, \tag{6}$$

$$\sum_{p\in P}\sum_{k\in K} x_{jkp} - C_j \le 0 \quad \text{for all } j\in J, \tag{7}$$

$$x_{jkp} \ge 0 \quad \text{for all } j\in J\cup\{0\},\ k\in K,\ p\in P, \tag{8}$$

$$v_{jp} \ge 0 \quad \text{for all } j\in J,\ k\in K,\ p\in P, \tag{9}$$

$$q_{jk} \in \{0,1\} \quad \text{for all } j\in J,\ k\in K, \tag{10}$$

$$y_j \in \{0,1\} \quad \text{for all } j\in J. \tag{11}$$

The objective in Equation (1) minimizes the sum of transportation costs, material-handling, inventory, and facility costs. The first term represents the cost of shipments from depots to customers and material-handling fees at depots. The second term represents the cost of shipments from the CDC to customers. The third term represents the cost of shipments from the CDC to depots and the material-handling fees at the CDC. The fourth and fifth terms represent the holding costs for the outbound and inbound pipeline inventory for the depots. The sixth term corresponds to the holding costs for the average required inventory that must be kept at the depots to satisfy the service-level requirements. The last term in the objective function represents the facility costs. The constraints in Equation (2) are for conservation of flow in depots. The constraints in Equation (3) ensure that the variance of demand from one customer is fully allocated to the depot to which it is assigned. The constraints in Equation (4) ensure that the flow from the CDC to customer locations captures the demand that is not satisfied at depot locations and the demand of customers who are not assigned to any depot (except the CDC). The constraints in Equation (5) ensure that the flow from a depot to a customer is equal to the fill-rate portion of that customer's demand. The constraints in Equation (6) ensure that a customer is assigned only to an open depot. The constraints in Equation (7) ensure that the capacity of a depot is not exceeded. The constraints in Equations (8)–(11) are the usual nonnegativity and integrality constraints.

We can only provide approximate numbers about the number of locations and customer sites. The number of potential sites ($|J|$) was in the range of 10–50 and the number of customer locations ($|K|$) was in the range of 200–500. There were about 50,000 parts

($|P| = 50,000$). We had about 190,000 part–customer location pairs with positive demand leading to a model with about three million continuous variables, three thousand binary variables, and about three million constraints. This represents the initial size of the model at iteration 1. As we described above, we add binary variables and constraints at each iteration as we increase the number of pieces that we use to approximate the nonlinear functions RI_{jp}.

Acknowledgments
D. Bhatia's current affiliation is the Decision Technology Group, Applied Materials, Santa Clara, California 95050.

References

Arntzen, B. C., G. G. Brown, T. P. Harrison, L. L. Trafton. 1995. Global supply chain management at Digital Equipment Corporation. *Interfaces* **25**(1) 69–93.

Camm, J. D., T. E. Chorman, F. A. Dill, J. R. Evans, D. J. Sweeney, G. W. Wegryn. 1997. Blending OR/MS, judgment, and GIS: Restructuring P&G's supply chain. *Interfaces* **27**(1) 128–142.

Candaş, M. F., E. Kutanoğlu. 2007. Benefits of considering inventory in service parts logistics network design problems with time-based service constraints. *IIE Trans.* **39**(2) 159–176.

Cohen, M. A., Y.-S. Zheng, V. Agrawal. 1997. Service parts logistics: A benchmark analysis. *IIE Trans.* **29**(8) 627–639.

Daskin, M. S., C. R. Coullard, Z.-J. M. Shen. 2002. An inventory-location model: Formulation, solution algorithm and computational results. *Ann. Oper. Res.* **110**(1–4) 83–106.

Edge. 2000. Applied Materials' Total Support Package helps accelerate production ramp-up at LSI Logic. *EDGE: Work-Group Comput. Rep.* (March 6), http://findarticles.com/p/articles/mi_m0WUB/is_2000_March_6/ai_60034084/.

Erlebacher, S. J., R. D. Meller. 2000. The interaction of location and inventory in designing distribution systems. *IIE Trans.* **32**(2) 155–166.

Geoffrion, A. M., G. W. Graves. 1974. Multicommodity distribution system design by Benders decomposition. *Management Sci.* **20**(5) 822–844.

Karabakal, N., A. Günal, W. Ritchie. 2000. Supply chain analysis at Volkswagen of America. *Interfaces* **30**(4) 46–55.

Laval, C., M. Feyhl, S. Kakouros. 2005. Hewlett-Packard combined OR and expert knowledge to design its supply chains. *Interfaces* **35**(3) 238–247.

Shen, Z.-J. M., C. Coullard, M. S. Daskin. 2003. A joint location-inventory model. *Transportation Sci.* **37**(1) 40–55.

Sherbrooke, C. C. 1968. METRIC: A multi-echelon technique for recoverable item control. *Oper. Res.* **16**(1) 122–141.

Solvoyo. 2010. planLM—Supply Chain Planning and Analysis PlatforM. Retrieved February 5, 2010, http://www8.solvoyo.com/webfront/planlm.html.

Cassio Conceicao, Vice President, Service Products Group, Applied Global Services, P. O. Box 58039, Santa Clara, California 95052-8039, writes: "This letter is to confirm the benefits obtained from a project led by Dr. Alper Şen, Dr. Koray Doğan and Deepak Bhatia at Applied Materials to rationalize our service parts network in North America in 2007.

"By evaluating and rationalizing our North American service parts network, we substantiated the belief that we could provide improved service at lower costs through redesigning our network. This result, obtained by a team that included both academics and business practitioners, demonstrated clearly the potential savings, and succeeded in shifting our 'inventory-centric' perspective to one of 'total-cost optimization.'

"We have since consolidated our depots in North America. We believe that this network rationalization effort has led to an inventory reduction of $5 Million and an estimated cost reduction of $1.1 Million so far. We have also incorporated network design and optimization into a short list of key strategies as we expand our existing service business and add new service businesses to our portfolio. In designing the service support network for our solar-panel equipment business, for example, we were able to leverage our new total-cost modeling techniques. By keeping total costs down as we provide superior service support, we contribute to the success of our customers in a very tangible manner."