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
2821 Scott Blvd, MS1785, Santa Clara, CA 95050, USA
deepak.bhatia@amat.com

Koray Doğan
Solvoyo Co.
77 Franklin Street Suite 504 Boston, MA 02110, USA
koray.dogan@solvoyo.com

Applied Materials, Inc. is the global leader in nanomanufacturing technology solutions, with a broad portfolio of innovative equipment, service and software products. The company supports its customers worldwide with an extensive service and parts network of over 100 locations. At the end of 2006, the company decided to evaluate and rationalize its network design in North America. A detailed optimization model (with 50,000 parts) was set up to develop a network and distribution strategy. To our knowledge, this is the first large-scale multi-echelon network design model that incorporates safety stock inventory costs with proper consideration of lead time and risk-pooling effects. Using the model, Applied Materials was provided with various improvement recommendations to reduce cost while maintaining or improving the service to customers. The recommendations included simplifying the distribution network through consolidating depot locations for certain customers and skipping an echelon for certain others, leading to a projected inventory reduction of $10 million. The company is currently acting on these recommendations and has already eliminated five depots. It is estimated that $5.24 million of inventory reductions in the first year of implementation can be attributed to these changes in the network and total savings during this period are $1.1 million.

Keywords: industries: semiconductor, service; facilities–equipment planning: location.

Introduction

This paper describes a network design project that was undertaken at Applied Materials for its service and parts network in North America. The objective of the project was to rationalize and evaluate the existing network in North America and provide senior management of the service and parts division a number of 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. Geographically, 20% of the orders come from North America, while 11% and 69% of the orders come from Europe and Asia, respectively.

The semiconductor equipment that the company manufactures are critical to the successful operations of its customers and it is thus important to have this equipment running at all times. In order to provide spare parts and service to customers in the event of equipment failures and for scheduled maintenance, the company has an extensive spare parts distribution network. The network consists 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 the 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. The company also manages consigned inventory in the stock rooms in facilities of its leading customers under agreements called Total Parts Management (see EDGE (2000)).

At the end of 2006, the company decided to evaluate and rationalize its existing service and parts network and develop a new distribution strategy in North America. The network in North America consists of over 50 locations serving several hundred customer ship-to destinations. The objective was to understand whether there were opportunities to reduce costs while maintaining or improving customer service. The costs under consideration were all costs relevant to the operation: inventory holding, transportation (within the network and outbound to customers), material handling, and warehouse costs. This was a comprehensive network design project that spanned the global inventory management and logistics functions at the company. We have to note here that even though the project scope was limited to North America, global demand and requirements needed to be considered because the CDC in North America also serves as a global procuring location for a significant portion 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/facility location decisions. However, the usual trade–off in these models is between transportation and facility costs; the cost of holding inventory due to demand uncertainty is not considered. 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 here is that the required inventory at any node in the network is a non–linear function of the demand during (effective) lead time, which itself is a decision variable in the problem. A node's lead time (first component of the demand during lead time) depends on which node upstream it is assigned to, and the demand it faces (second component of the demand during lead time) depends on which nodes downstream are assigned to it, both of which are endogenously determined in the model. Several approaches are suggested in the literature to deal with these non–linearities. For example, Daskin et al. (2002) use a Lagrangian based approach; Shen et al. (2003) employ column generation; Erlebacher and Meller (2002) 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, as they consider only a single product in the network. The approach in Candaş and Kutanoğlu (2007) may be used potentially for problems with a limited number of parts (the dataset in their computational study has 4 parts); but for a problem like ours, with tens of thousands of active parts, a new scalable approach was needed.

Understanding the difficulty of the problem and the unavailability of any commercial software in the market, Applied Materials decided to form a team that could undertake the development of a new solution. The core team (the authors of this paper) consisted of a faculty member at a research university, 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, is 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. It was decided that the functionality of this solution would be extended to incorporate the cost of holding inventory due to uncertain demand. While the major task in the project was to come up with a mathematical model and a software solution for the inventory–location problem, a significant amount of time and effort was spent on data collection and needed the participation of various Applied Materials employees throughout the project. The project was kicked off in January 2007 and successfully finished in July 2007, with a new scalable solution for the inventory–location problem developed and implemented for the service and parts network in North America. The results showed a potential of 10 million dollars in inventory savings through consolidating depot locations for certain customers and skipping an echelon for others without sacrificing customer service levels. The results were shared with Applied Materials executives in June 2007, and the company is currently moving forward with implementing the recommendations of the study. Based on these recommendations, five depots have been eliminated in the first year, leading to 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 different customer locations in North America. The network to serve these customers consists of a continental distribution center, Applied Materials–owned and –operated depots, and Applied Materials–managed stock rooms (consignments) in customer facilities. Applied Materials has 50,000 active parts composed of consumables and non–consumables, with a large variation in part costs. All parts are procured at the CDC and variation is also extremely high for supplier lead times (range from one to 270 days).

There are two types of customer orders. The first type results from equipment failures ("emergency orders") and is critical for the customer. These are usually satisfied from consignments (if such an agreement exists) or from depots. The CDC provides a second level of support for emergency orders that cannot be satisfied immediately from consignments or depots. The second type of demand occurs when customers request spare parts to be used in their scheduled maintenance activities ("regular orders"). The primary source to meet these demands is usually the CDC. However, for certain customers, local depots can also be used.

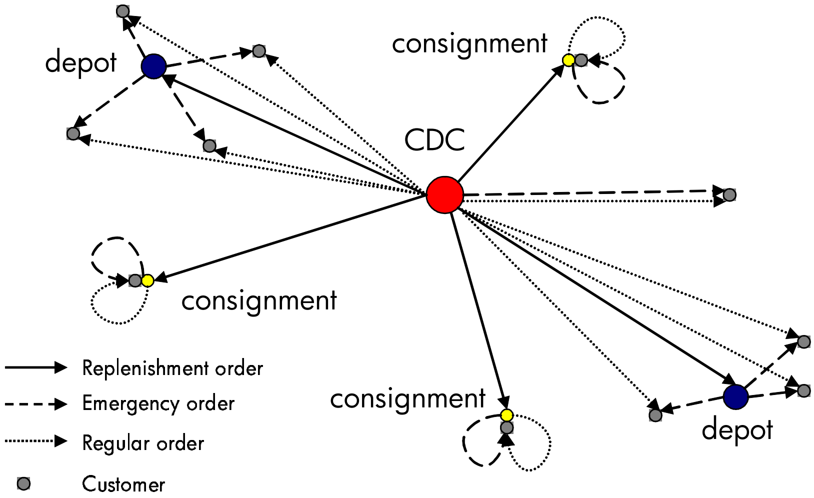


Figure 1 **Service Network in North America**

Both types of customer orders (emergency and regular) go through an order fulfillment engine, which searches for available inventory in different locations according to a search sequence specific to each customer. Emergency orders need to be satisfied immediately (their request date is the date of the order creation), whereas regular orders need to be satisfied at a future date. A depot may be facing emergency and regular demands simultaneously from a variety of customers. The orders that are placed by the consignments and depots to replenish their stocks are called "replenishment orders" and the CDC is used for this purpose. The CDC also faces emergency and regular orders from external customers. These are typically due to unsatisfied demand at the depots or consignments. But for a minority of customer locations that are closely located, the CDC may be the first source. The service network in North America is depicted in Figure 1.

Service level requirements vary for each customer site and type of order. Service levels at customer sites with consignment agreements are well structured and enforced through contracts. Required service from depots and the CDC are usually not mandated through contract, but through targets set by the executive team. Service levels are measured in terms of fill rate, i.e., percentage of orders satisfied from stock. Since 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. Inventory planning at the CDC is done by a global inventory management team. For depots and consignments, the global inventory management team suggests inventory levels and policies, but local teams are responsible for management of 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 a selection of carriers, including specialized courier services, parcel companies, air freight forwarders and other international logistics companies. Each carrier charges Applied Materials per transaction, based on the pick–up time, urgency, weight, and destination of the order.

The Mathematical Model and Assumptions

In order to develop a network and distribution strategy, the core team needed to formulate a mathematical model. The mathematical model would take various constraints into consideration in each scenario and come up with the optimal network configuration that would minimize the total transportation, inventory holding and material handling costs in North America. Note that the scope of the study was the service network in North America; particularly the number and locations of depots. As the consignments are co–located with customers, their number and locations were not decision variables in the study. Similarly, the location of the CDC was not part of the study and taken as fixed. Since the consignments are replenished only through the CDC and the total demand that the CDC has to serve does not change with a change in the network, the inventory at and the cost of transportation between the CDC and the consignments could be taken outside of the optimization model. Note, however, that some of the scenarios we studied require the service level (fill rate) at the CDC (β_C) to be increased, which lead to an increase in inventory at the CDC and a decrease in inventory at the consignments. These calculations are carried out externally and the costs are added to the costs that are obtained by running the mathematical program described below. For the inventory that is carried at the CDC, we had to incorporate demands from locations outside of North America in our model. For this purpose, 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 β_T across different consignments, while in reality each customer may be subject to customized consignment contracts with various terms, including different service level definitions and commitments.

Given the scope of the study stated above, the objective was to develop a model that would decide on the number and the locations of the depots to minimize inventory holding, transportation and material handling costs. In order to obtain a mathematical model that was scalable and whose data requirements were readily available, a number of assumptions had to be made. Since the model and its results would be used for strategic purposes (and not for daily operations), it was agreed that these assumptions would not hinder the validity of the recommendations. These assumptions for the optimization model are listed below.

Assumption 1 – All parts at a location have the same service level: The service level commitments and targets used by Applied Materials are usually not at the parts level. Service levels are defined and measured based on the total demand of all parts faced by a customer. 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) averaging the service level commitment to a customer. However, it was agreed that this level of detail would not alter the results and recommendations of the model. We assume that each part at a location will have the same service level target. The service levels for the baseline model were selected based on historical performance and customer requirements (β_T for the consignments, β_D for the depots and β_C for the CDC).
Assumption 2 – There is no differentiated service at the CDC: As explained before, the CDC has multiple roles in the network: *i-)* Satisfying regular demand for non–consignment customers *ii-)* Replenishing consignments *iii-)* Replenishing depots *iv-)* Satisfying secondary emergency demand in the network. While it is possible to employ rationing and offer differentiated service to different order types in theory, this requires an order fulfillment software that is capable of reserving material for different types of demand. The software at Applied Materials did not have this type of functionality at the time of project implementation; therefore it is 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 more than one part. Since the transportation cost for a shipment is usually non–linear in the weight of the shipment, the average transportation costs per unit of weight depend on the distribution of shipment weights. In Data 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 the weight distributions and thus the transportation costs per unit of weight for each origin and destination will not change as a result of a change in the network. We also assume that the rates provided by the carriers will not change with a change in network.
Assumption 4 – Unsatisfied demand at a depot will be expedited only from the CDC: When a customer is facing a down situation and the depot or the consignment that it is assigned to is out of stock for this part, the entire North American network (excluding some consignment locations for other customers) is searched for stock. 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, a mixed integer linear programming model was developed. In the Appendix, we discuss the model, its features and size in terms of the number of variables and constraints. Note again that the model described here does not include the inventory at and the shipments in between the consignments and the CDC. The costs related to these are calculated externally.

An important feature of the optimization model we use is the approach we take to calculate the inventory that is required 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 non–linear 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 assumed that the replenishment lead time is equal to the effective lead time for the depot: $\ell_j + (1 - \beta_C)L_p$). 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 of only safety stock. If the replenishment policy is (R, Q), then RI_{jp} includes cycle inventory as well as safety stock.

In order to deal with the non–linear RI_{jp} function in the objective function of the optimization model, we use the following approach as shown in Figure 2. First, we find the maximum variance that can be allocated to each depot for each part. In our application, this corresponds to the variance of the total emergency demand originating from non–consignment customers in North America, since one particular depot may 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 the problem with the linear cost function at iteration 1 gives us solution 1. At iteration 2, we find the safety stock requirement that corresponds to the allocated variance in solution 1 and approximate the function using a piecewise linear function (as shown in the second graph in Figure 2). Solving the problem with this piecewise linear function by incorporating binary variables gives us solution 2. At each iteration after this, 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 threshold, we continue and incorporate another piece to the approximation. If the difference is less than a threshold for all depot–part combinations, we stop and report the latest solution as the solution to the problem. 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 threshold value that is equal to 0.01% of the safety stock requirement in the previous iteration. With this threshold, the computation times were acceptable and the resulting inventory costs were very 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. In fact, we have attempted to use this approach initially to solve the problem. However, since all non–consignment, 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 non–linear required inventory curve. This led to introducing too many binary variables upfront and was prohibitive in successfully solving the problem with the commercial solvers we used.

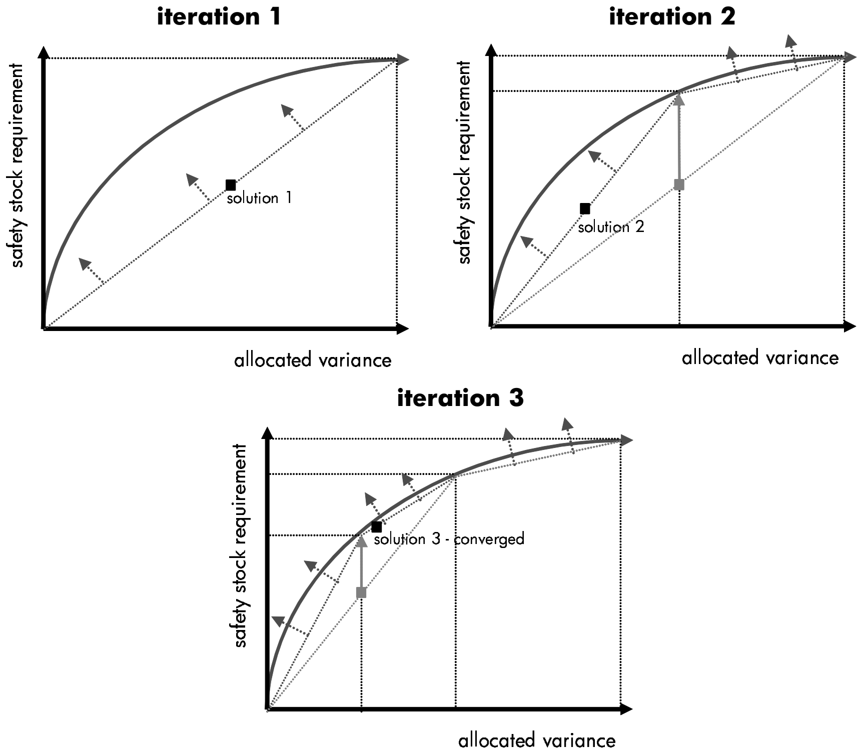
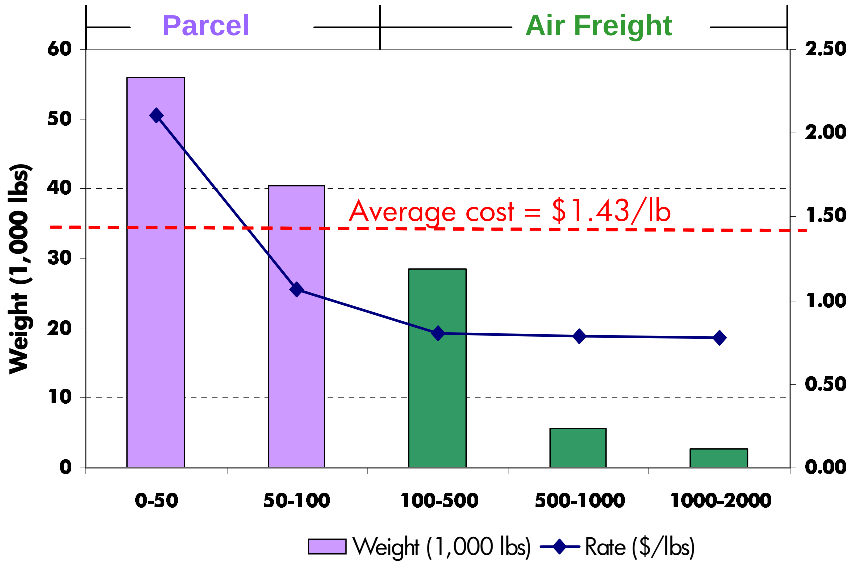


Figure 2 Approximating the Non–linear Objective Function for the Model

Data Collection

We now explain the source and the estimation methods for four critical data elements:
Transportation costs from depots to customer locations: Estimating cost (per unit) between depots and customer locations was non–trivial. First, shipments from a depot to customer locations were created based on specific customer orders, which could consist of different line items and thus different weights. Therefore, shipment costs and thus cost per pound (or cost per unit) vary from one customer order to another even when the delivery time is fixed. Second, based on the weight and the delivery time requirement of a shipment, Applied Materials chooses from several transportation service providers. The decision is made at the time of the shipment, based on the cost and the availability of the transportation service providers. In order to solve the first problem, we had to come up with a weight distribution of shipments to customer locations. Note that the weight distribution of a customer site should be independent of the origin (depot location). To get the weight distribution, we used the complete shipment history for the 12–month period just prior to the implementation. We used two representative carriers to estimate the shipment cost of an emergency order, based

on a routing guide for the choice of the carrier at the service and parts division: a major parcel company for shipments weighing less than 100 lbs and an airfreight forwarder for shipments weighing more than 100 lbs. The rates for the parcel company 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). First, we found the corresponding zone of each possible depot location and customer pair. Since the shipments from depots to customer locations are all emergency orders, we used Priority Overnight service. The actual weight distribution of shipments below 100 lbs was estimated with two intervals: 0–50 and 51–100. The average shipment weights were calculated and used to pick the rates from the rate tables. A similar approach is used for the airfreight forwarder, which has a similar zone system and an all–units pricing scheme based on five weight breaks. Combining these, the average transportation cost for each potential depot site, customer location and part number was calculated.



Range	0-50	50-100	100-500	500-1000	1000-2000
Total Weight	56000	40480	28600	5646	2806
Average Weight	10.67	71.27	185.71	564.68	1403
Rate ($/lbs)	2.11	1.06	0.80	0.79	0.78

Figure 3 **Estimating Freight Rates**

A fictitious example is provided in Figure 3 for shipments from a potential depot site to an undisclosed customer site, with fictitious rates. The average weight for shipments in the 0–50 and 51–100 intervals are 10.67 lbs and 71.27 lbs, respectively. Therefore, the rates corresponding to 11 lbs and 72 lbs are picked up from the rate tables of the parcel provider. For shipments of more than 100 lbs, weight breaks of the airfreight forwarder are used, and the total weight of shipments that are sent in these weight ranges are considered. The calculation

results in an average freight rate of $1.43 per pound. The unit freight rate for each part is calculated using standard part weights (available in the ERP system) plus a packaging factor.
Transportation costs from the CDC to depots and customers: A similar process was followed to calculate the transportation costs from the CDC to depots and customers. One major assumption here was that the weight distributions used in calculating the costs from the CDC to depots will not change as a result of changes in the network. A more accurate method would be to calculate the weight distribution 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 the calculation of the weight distributions for all scenarios would be computationally prohibitive as we had 10–50 depot locations and 200–500 customer sites. It was understood that relaxing the constant weight distribution assumption will work in favor of the general recommendations of the study, which suggest consolidating depot locations, since consolidation will lead to heavier (and therefore cheaper per unit of weight) shipments from the CDC to depots. However, it was possible, in theory, that the constant weight distribution assumption might have an impact on the details of the recommendation (i.e., which particular depots to close). However, as we show in the next section, transportation costs are heavily dominated by inventory holding and material handling costs in our implementation. Therefore, we believe that inventory holding and material handling costs were the primary drivers of the decision with regards to which depots to keep and which other depots to close, and the assumption had little or no effect.
Demand: We used one year demand history just prior to the implementation to determine the demand distributions at the part, customer and order type level. This required a classification of parts as slow–moving, medium–moving and fast–moving. We used the classification that Applied Materials is using in their current operations. Applied Materials classification is based on the number of customer requests within the last two years, the time between customer requests and the time since last request. For slow–moving items, it was assumed that demands are distributed with Poisson distribution with a mean annual rate equal to the total demand in the last year. For medium– and fast–moving items, it was assumed that the annual demands are distributed Normally again with a mean equal to the total demand in the last year. The standard deviation is found by using a constant coefficient of variation. This coefficient of variation was the average value for these parts and these assumptions were consistent with the approach taken by the inventory management team used in their operational inventory decisions.
Replenishment Policy: For slow–moving items, it was assumed that the replenishment policy is of type $(S-1, S)$. For medium– and fast–moving items, it was assumed that the replenishment policy is of type (R, Q). For the calculation of the Q, a standard EOQ approach was used. Again, these assumptions were consistent with the approach taken by the inventory management team for their actual operational decisions. The replenishment policy choice impacts the amount of inventory required for each part at each depot (RI_{jp} defined in the mathematical model section) and the inventory required for each part at the CDC and the

consignments (which is calculated outside of the model). That is, if the part is a slow–moving item, the required (average) inventory is calculated based on the $(S-1, S)$ policy and only includes safety stock. On the other hand, if the part is a medium– or fast–moving item, the required (average) inventory 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) as explained above and Assumption 3 in mathematical model section. For the EOQ calculations, we used the same order cost for all parts and locations (a similar approach is used by the inventory management team for their operational decisions). Annual inventory holding cost for a part was calculated by multiplying the part standard cost with the annual inventory holding cost rate. Since the equipments at customer sites are usually operated at full capacity at all times, the spare parts demand did not exhibit any seasonality. There was no apparent trend in one year demand history either. Therefore, given a depot location and its customer assignment, the model generates only one set of parameters for the replenishment policy for each part.

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

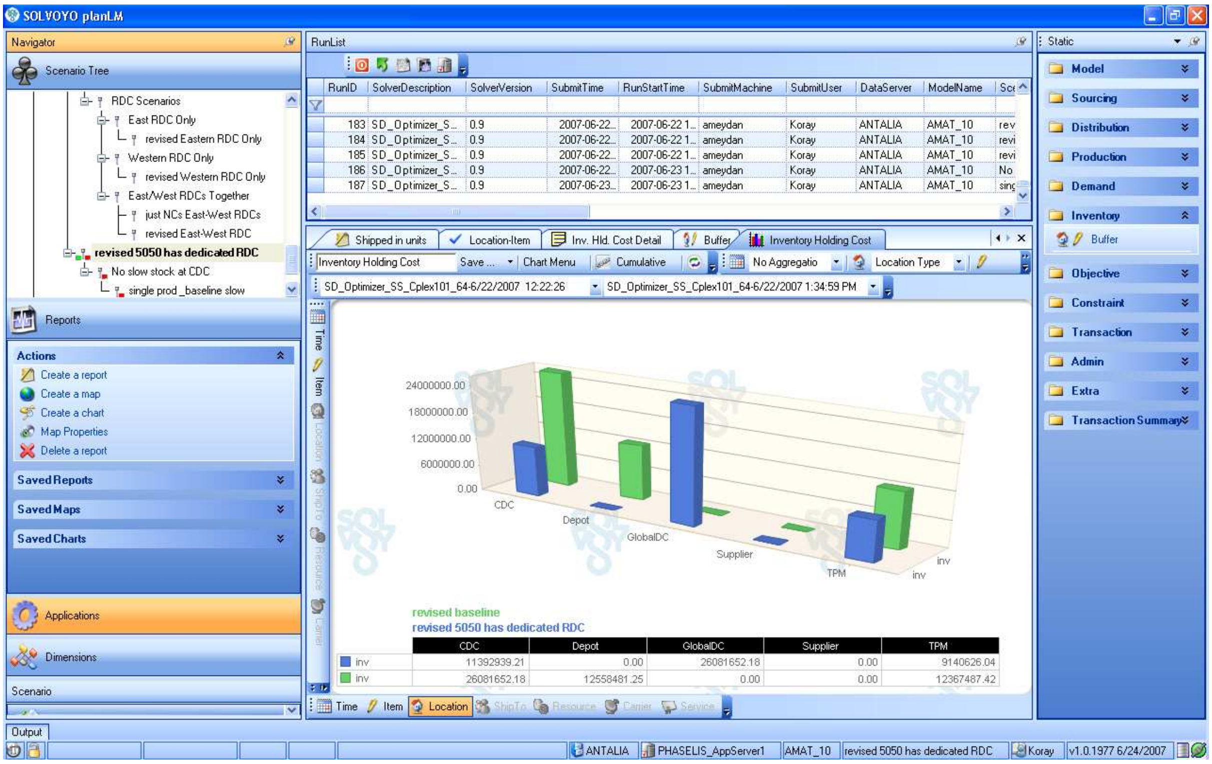
Software Solution

We implemented the mathematical model and the solution procedure in Solvoyo planLM Supply Chain Planning and Analysis Platform. Solvoyo planLM provided the graphical user interface, application server infrastructure, and data interface to an extensive supply chain data model that already included all the data requirements for the mathematical model given in Appendix. Therefore no changes were necessary for the data model. To speed up the development efforts, the mathematical model was implemented as an extension to planLM Supply Demand Optimizer application, which is designed for strategic network design problems as well as tactical master planning. While planLM provides interfaces to most third–party solver engines, the development was conducted using ILOG CPLEX 11.0 on dual core 64bit Windows–XP platform. The run times for solving the mixed integer linear program given in the Appendix varied between two to eight hours, depending on the specific scenario, and utilized about 23GB memory at peak. planLM's scenario management capabilities were particularly useful in quickly creating and comparing the scenarios and running sensitivity analysis. The user interface and other reporting features of planLM were also useful in speeding up the analysis. A screen shot of the user interface that shows comparisons of different scenarios is given in Figure 4.

Model Validation, Recommendations and Results

Once the data was gathered and the model was implemented in planLM, our next task was validation. For this purpose, we established a baseline scenario corresponding to the current depot locations and customer–depot assignments. We used the latest actual service level performance on depots (β_D) and the CDC (β_C) and used a constant β_T service level target for the consignments. The results of the baseline scenario were validated against the actual inventory, transportation and material handling costs. The model predicted the total costs to be 9.8 % more than the actual total costs. The model generated close to actual values for the transportation and material handling costs. The inventory carried at the CDC and the consignments also matched the actual numbers closely (both within 2 %). The model suggested to keep 29.2 % more inventory in depots, primarily due to the fact that the model forces a service level of (β_D) for each part, while the field objective was to maintain this service level at a location level (See Assumption 1 in model section). The differences were considered acceptable as the model would be used for strategic purposes only.

In Figure 5, we report different cost components as a percentage of total cost for the baseline scenario. The transportation costs (2.56%) are heavily dominated by the material handling costs (48.46%) and inventory holding costs (48.98%) at different echelons in the network, showing that the current network is designed to serve customers from locations in close proximity, incurring minimal transportation costs. This is accomplished by establishing an echelon of many depots to serve non–consignment customers, holding a substantial



RunID	SolverDescription	SolverVersion	SubmitTime	RunStartTime	SubmitMachine	SubmitUser	DataServer	ModelName	Sce
183	SD_Optimizer_S...	0.9	2007-06-22...	2007-06-22 1...	ameydan	Koray	ANTALIA	AMAT_10	rev
184	SD_Optimizer_S...	0.9	2007-06-22...	2007-06-22 1...	ameydan	Koray	ANTALIA	AMAT_10	revi
185	SD_Optimizer_S...	0.9	2007-06-22...	2007-06-22 1...	ameydan	Koray	ANTALIA	AMAT_10	revi
186	SD_Optimizer_S...	0.9	2007-06-22...	2007-06-23 1...	ameydan	Koray	ANTALIA	AMAT_10	No
187	SD_Optimizer_S...	0.9	2007-06-23...	2007-06-23 1...	ameydan	Koray	ANTALIA	AMAT_10	sing

	CDC	Depot	GlobalDC	Supplier	TPM
🟦 inv	11392939.21	0.00	26081652.18	0.00	9140626.04
🟩 inv	26081652.18	12558481.25	0.00	0.00	12367487.42

Figure 4 **User Interface: Charts**

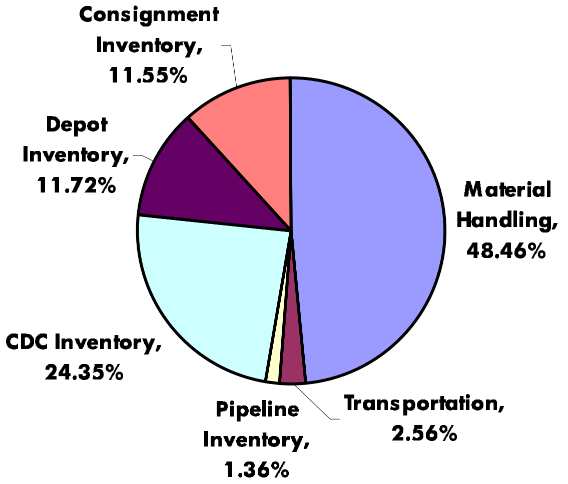


Figure 5 Baseline Scenario

amount of decentralized inventory and incurring heavy material handling costs. Dominance of transportation costs by inventory holding costs as demonstrated here is typical of service parts logistics, as majority of items are slow moving, inventory turnover rates are small and obsolescence and scrap rates are high. According to a benchmark study of 9 companies in the computer industry, transportation costs constitute only 8.4% of all relevant costs in service functions of these companies (Cohen et al. 1997). The reported transportation costs are even less significant in Figure 5, as parts are less bulkier (average weight is 6.22 lbs) and we exclude the inbound freight costs (from supplier to the CDC) in our model.

A portion of the costs in Figure 5 are unavoidable. For example, the CDC in North America is replenishing all locations and customers in North America for all parts and in the world for most parts. Therefore, the inventory at the CDC cannot be changed by reconfiguring the network unless the service levels are changed. Material handling costs at the CDC also cannot be reduced for the same reason. Similarly, service levels and consignments are contractual obligations for Applied Materials. Therefore any reduction in inventory costs for consignments is not possible by a change in the service network unless the service level for the CDC is changed. With these in mind, reducing the inventory carrying costs and material handling costs by reconfiguring the network at the depot layer was identified to be a major opportunity for cost reduction. Whether and how this could be accomplished also considering transportation costs was investigated using four different scenarios. We describe these scenarios below.

Scenario 1 – Optimized Baseline: In this scenario, we optimize the depot customer assignments. In doing this, the solver can pick the CDC as a source for satisfying the emergency demand of a customer. We keep the CDC service level at β_C. Thus, this scenario can potentially reduce the service level to certain 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 if they were assigned to the CDC for emergency demand. However, since the CDC provides non–differentiated service to different types of demand, as explained before, this also leads to better service to other regions and regular orders. In addition, service level for replenishments to consignments is also increased, which may lead to reductions in consigned inventory.

Scenario 3 – Separate Stock for Emergency Demand at the CDC: A new stock location at the CDC is created to reserve separate stock for emergency demand, providing β_D service level for those customers that are assigned to the CDC. The rest of the network operates as before.

Scenario 4 – Regional Distribution Center for North America: A new distribution center is introduced for North America to satisfy the North American demand only. This regional distribution center (RDC) is a virtual location where the CDC is located, and will provide β_D to all types of demand in North America.

Scenario 5 – Two Regional Distribution Centers for North America: Two RDCs are established in North America: One on the east coast and the other 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 instead of a depot. The results of the solution for these scenarios are given in Table 1. All costs are scaled by the total costs in the baseline scenario (Scenario 0).

Table 1 **Comparison of Scenarios**

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
6	2.83	46.75	47.54	97.12	99.0

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

Based on these results, Scenario 2 was initially picked as a recommended solution. While Scenario 1 obtains a lower total cost, the customer service levels will be diminished by this alternative since the customers that are assigned directly to the CDC will receive lower service. Scenario 3 also leads to slightly smaller total costs, but it was decided that the benefits of increasing the service level at the CDC for all order types globally outweighs this difference in cost. One final iteration was done on Scenario 2 before presenting it as a recommendation. This was due to the service time that would be provided to a certain set of customers under the new network design. Under the former system, these customers received support for their emergency orders within four hours by a special ground service that was provided by specific depots to which they were assigned. With the new network, the customers may be replenished from the CDC which would mean a "Priority Overnight" type of service provided by a parcel company, hence increasing the time to fulfill their request. It was decided that the status quo will be maintained for these customers, at least for a transition period. A new scenario was set up and the solution was forced to use the existing depot–customer assignments for these customers. The results are shown in Scenario 6 in Table 1. Overall, Scenario 6 was expected to lead to a 2.88% savings in total costs. Scenario 6 recommends closure of 6 depots and serving their customers from the CDC at an increased service level. This was to lead to the elimination of inventory at these depots, a reduction of inventory at the remaining sites (at consignments and open depots due to the increased service level at CDC) and an increase in inventory at the CDC. Total inventory was to be reduced by 4.55%. Other major savings will be on material handling which will go down by 1.9% due to shipping parts direct from the CDC. The transportation costs, on the other hand, was expected to increase by 10.5% as fast and small sized customer shipments would now be sent over longer distances.

One final investigation was the sensitivity of the results with respect to two important parameters in the analysis. One was the inventory holding cost rate. The core team was reminded 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.

Table 2 **Sensitivity of Results**

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

Extensive sensitivity analysis that was carried out in planLM showed that while changes in these two parameters impact the magnitude of the savings, ordering of the alternatives would not change. Table 2 shows the sensitivity of the results for all scenarios to changes in the inventory holding cost rate and the material handling fee per shipment where i and f represent the values of these parameters that were used in the study. The analysis showed that the savings would be still significant with Scenario 6 when the parameters change (percentage savings over Scenario 0 are in the range 1.80%–4.06%). Sensitivity analysis also showed that the total inventory (and where they are carried) in Scenario 6 does not change unless the inventory carrying cost rate is significantly lower than what is used in the project and/or the material handling fee is completely eliminated.

The project results, including the recommendation to change the network, were presented to the senior management of the service and parts division on June 26, 2007. The approach taken by the core team was approved in the meeting. The feedback that was received with regards to the recommendation was also positive. The project was officially declared complete in July 2007.

Applied Materials started implementing the recommendations of this project in early 2008. In 2008, three depots were eliminated in North America. In 2009, two other depots were closed. A thorough analysis was carried out in July 2009 to measure realized savings due to the restructuring of the network in North America. The analysis involved a comparison of 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. In order to fairly assess the savings due to network restructuring, the impact of these changes are eliminated in our analysis. The results shows that out of the total inventory reduction since the depots were closed, \$5.24 million can be attributed to network redesign recommendations coming from this project. It is estimated 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 further savings of \$1.38 million annually going forward. These savings estimates are conservative as the impact of higher service level at CDCs on inventory levels at consignments and regions outside of North America was not considered. Analysis conducted also showed that there has been no impact to the realized service level due to these changes in the distribution network. A consolidated network also brings other benefits to Applied Materials including but not limited to a reduction in network complexity potentially leading to 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. A novel approach that simultaneously considers the inventory and logistics costs was developed. This approach was implemented to extend the functionality of an existing software solution to solve Applied Materials' large–scale problem and evaluate different alternatives. To our knowledge, this is the first multi–echelon network design solution of this scale that incorporates the safety stock inventory costs with proper consideration of the lead time and risk–pooling effects. The approach taken in the project and the recommendation of the study were approved by the senior management team at Applied Materials. The recommendations included simplifying the network through consolidating depot locations for some customers and skipping an echelon for others. These measures were projected to lead to significant inventory reductions for Applied Materials. The company is currently implementing the recommendations and has so far eliminated five depots from its parts and service network in North America. It is estimated that $5.24 million portion of inventory reductions over the first year of implementation can be attributed to these changes in the network.

Many companies face network design problems for which safety stock inventory holding costs constitute a big 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 here.

Appendix. Mathematical Model

First, we introduce the notation 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,
$\sigma_{kp}^2:$ 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,
$\ell_{jk}:$ Delivery time from depot j to customer k,
$\ell_j:$ 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,
$\beta_C:$ Service level for the CDC,
$\beta_D:$ 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:

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

$$\begin{aligned} z_{jp} - \sum_{k \in K} x_{jkp} &= 0, \text{ for all } j \in J, p \in P && (2) \\ -v_{jp} + \sum_{k \in K} \sigma_{kp}^2 q_{jk} &= 0, \text{ for all } j \in J, p \in P && (3) \\ \sum_{j \in J} x_{jkp} + x_{0kp} - d_{kp} &= 0, \text{ for all } k \in K, p \in P && (4) \\ x_{jkp} - \beta_D d_{kp} q_{jk} &\le 0, \text{ for all } j \in J, k \in K, p \in P && (5) \\ q_{jk} - y_j &\le 0, \text{ for all } j \in J, k \in K && (6) \\ \sum_{p \in P} \sum_{k \in K} x_{jkp} - C_j &\le 0, \text{ for all } j \in J && (7) \\ x_{jkp} &\ge 0, \text{ for all } j \in J \cup \{0\}, k \in K, p \in P && (8) \\ v_{jp} &\ge 0, \text{ for all } j \in J, k \in K, p \in P && (9) \end{aligned}$$

$$q_{jk} \in \{0,1\}, \text{ for all } j \in J, k \in K \tag{10}$$
$$y_j \in \{0,1\}, \text{ for all } j \in J \tag{11}$$

The objective in (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 needs to 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 (2) are for conservation of flow in depots. The constraints in (3) ensure that the variance of demand from one customer is fully allocated to the depot to which it is assigned. Constraints in (4) ensure that the flow from the CDC to customer locations captures the demand that is not satisfied at depot locations and customers that are not assigned to any depot (but the CDC). Constraints in (5) ensure that the flow from a depot to a customer is equal to the fill rate portion of the demand of that customer that the depot serves. Constraints in (6) ensure that a customer is assigned to a depot only if it is open. Constraints in (7) ensure that the capacity of a depot is not exceeded. The constraints (8–11) are the usual non–negativity and integrality constraints.

We can only provide rough figures regarding the number of locations and customer sites. The number of potential sites ($|J|$) was in the range 10–50 and the number of customer locations ($|K|$) was in the range 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 3 million continuous variables, 3,000 binary variables and about 3 million constraints. This is the initial size of the model at iteration 1. As described before, we add binary variables and constraints at each iteration as we increase the number of pieces that we use to approximate the non–linear functions RI_{jp}.

References

Arntzen, B. C., G. G. Brown, T. P. Harrison, L. L. Trafton. 1995. Global supply chain management at Digital Equipment Corporation. *Interfaces* **25** 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** 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 Transactions* **39** 159–176.

Cohen, M. A., Y. Zheng, V. Agrawal. 1997. Service parts logistics: a benchmark analysis. *IIE Transactions* **29** 627–639.

Daskin, M. S., C. R. Coullard, Z.-J. Shen. 2002. An inventory-location model: Formulation, solution algorithm and computational results. *Annals of Operations Research* **110** 83–106.

EDGE. 2000. Applied Materials' Total Support Package helps accelerate production ramp-up at LSI Logic. *EDGE: Work-Group Computing Report* **March 6**.

Erlebacher, S. J., R. D. Meller. 2002. The interaction of location and inventory in designing distribution systems. *IIE Transactions* **32** 155–166.

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

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

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

Shen, Z. J., D. Coullard, M. S. Daskin. 2003. A joint location-inventory model. *Transportation Science* **37** 40–55.

Sherbrooke, C. C. 1968. METRIC: A multiechelon technique for recoverable item control. *Operations Research* **16** 122–141.

Cassio Conceicao, Vice President, Service Products Group, Applied Global Services, Applied Materials Inc., 2821 Scott Blvd., Santa Clara, CA 95050, writes: “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 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. ”