A technical framework for sense-and-respond business management

S. Kapoor
K. Bhattacharya
S. Buckley
P. Chowdhary
M. Ettl
K. Katircioglu
E. Mauch
L. Phillips

In this paper we present a technical framework that supports sense and respond (SaR), the approach that enables an enterprise to adapt to a rapidly changing business environment. To implement the SaR approach, an enterprise proactively monitors trends and uses effective decision-support tools to help it act in a timely manner. We describe two pilot projects in which we implemented SaR prototypes and applied them to solve business problems. In the first pilot project we helped the IBM Microelectronics Division deploy an automated inventory management system based on our inventory optimization model. In the second pilot project, we helped the IBM Personal Computing Division deploy a SaR system in support of demand/supply conditioning. One of the components of this SaR system is an order trend model that provides early warning of constraints and excesses in the supply chain and helps make demand/supply conditioning more effective. Early results from these projects are encouraging and show that significant gains in profitability are possible.

An on demand enterprise is able to adapt to a rapidly changing business environment. When faced with intense competition and changing customer preferences, this type of enterprise can preserve or extend its competitive advantage. In order to make timely, well-informed decisions its executives must have a window into the operational health of the business and suitable decision-support tools.

The sense-and-respond organization

Haeckel suggests that when customers' needs change faster than the company's ability to respond to them, the company must re-engineer itself to become a sense-and-respond (SaR) organization.[1] The same conclusion would be drawn if the company's product life cycle lagged behind that of its competitors. In Haeckel's view, the organizational hierarchy should be replaced by a dynamically configured network of modular capabilities. He defines *capability* as an organizational subsystem whose governance is based on context and coordination by people in roles accountable for outcomes rather than by command and control. The process of re-engineering the organization can be summarized as follows:

- *Set the context*—Identify your stakeholders (customers, investors, suppliers, employees) and define your goals with respect to each of them. This represents your company's reason for being. Define the five to ten principles your employees need to follow in order to ensure the continued existence of the organization in that form.

- *Design for change*—Define the business as a system of roles and responsibilities. Define the roles that will make the organization work and the results the person in a role will be accountable for. Define the interactions between the people according to their roles, and provide guidelines for people to negotiate and manage their commitments to one another. Assign the right person for each role.

- *Execute*—Perform the "adaptability loop":

    1. Sense: gather data from all sources to monitor your business environment—customers, business news, sensors, market research, and intelligence.

    2. Interpret: use a variety of technique to interpret the data, such as analytics and what-if scenarios.

    3. Decide: re-examine your company's goals and principles, as well as the defined roles and responsibilities, and decide if changes are necessary.

    4. Act: carry out the changes, if needed, or reaffirm the current goals, principles, roles, and responsibilities.

Model-driven architecture

In conformance with model-driven architecture**,[2,3] we view the model of a business as a hierarchy consisting of four modeling layers.

1. *Strategy layer*—The strategy model specifies what the business plans to achieve (this model is sometimes referred to as the "strategy execution model," to distinguish it from strategy formulation). It models business objectives in terms accessible to business people. For example, it might specify the objectives in terms of the well-known balanced scorecard[4] perspectives:

    a. Financial perspective: How should we look to our shareholders?
    b. Internal business perspective: What must we excel at?
    c. Innovation and learning perspective: How can we continue to improve and create value?
    d. Customer perspective: How do we want our customers to see us?

    This model is defined in terms normally used in business strategy discussions and is verified and refined through iterative interaction with strategy executives.

2. *Operations layer*—The operations model describes what the business is doing to achieve its strategic objectives and how it measures its progress toward them. It is typically developed by operations planners in collaboration with strategy executives, and it captures business operations and business commitments (these are tracked through a set of KPIs (key performance indicators) that are directly linked to the balanced scorecard associated with specific business operations and processes).

3. *Execution layer*—The execution model describes the workflows and information flows that the business uses to implement the operations model. It does not assume a particular implementation. It also models the monitoring and management aspects of the process, which are driven by KPIs, and tracks the business commitments and goals defined at the strategy and operations layers.

4. *Implementation layer*—The implementation model defines the actual information technology (IT) processes in a specific realization of the execution model. Tools are available today to construct portions of the implementation model directly from the execution model, much as a compiler translates a high-level language.

This four-layer modeling approach should be used for both the process component and the decision support component of a capability. A change at any layer requires verification with respect to higher layers as well as (semi-automated) propagation of the change to lower layers. Examining the functional relationships within an organization reveals complex relationships between processes. In order to evaluate any decision or action in an organization, it is necessary to identify all the important interactions and to determine the impact on the entire organization.[5]

Our approach

We have developed a technical framework and the supporting operating environment for implementing SaR systems. To test and refine our framework, we have implemented SaR prototypes and used them in pilot projects in two IBM businesses, the IBM Microelectronics Division (IMD) in Burlington, Vermont, and the Personal Computing Division (PCD) in Raleigh, North Carolina. To ensure that the implementation is driven by business requirements, we use the multilayer model-driven approach described earlier.

We also observe that reusable capabilities can be constructed using fundamental concepts from business components[6] and Web services.[7] According to business component expert Peter Herzum, "A business component represents a software implementation of an autonomous business concept or business process. It is composed of all software artifacts necessary to express, implement, and deploy a business component as an autonomous, reusable element of an information system." It is increasingly popular to build a business component as a Web service, exposing it through a service description interface such as WSDL.[8] The service interface describes the operations or services offered by the business component and the data needed to execute each operation. Web services can be publicized in a registry, which allows for flexible discovery of capabilities by other capabilities.

We define here a technical framework for SaR systems based on the foundation established by Web services and business components. Our work was performed in coordination with the recent IBM initiative in the area of Business Performance Management (BPM), a real-time, model-based discipline to optimize and adapt business operations and IT infrastructures based on dynamic performance targets.[9]

Run-time monitoring of a business process makes the KPIs visible to business managers. Violated commitments can be quickly detected. Using advanced analytics (sophisticated analysis of data), trends can be detected that may predict and prevent future violations of business commitments. Analytic trend detection allows a SaR system to support proactive business management, enabling critical changes in strategy and execution to be carefully planned well in advance, thereby preventing business exceptions well before they occur.

Analytics is an essential part of SaR systems because it detects business trends and helps diagnose difficult business problems. Analytic components are often tailored to handle specific industries or domains.

Another important aspect of our framework is optimization. Based on the trends and special situations identified through analysis, optimization components help plan the actions to be taken. Such actions fall into a number of categories, including:

- Notifying the appropriate business managers
- Changing the operational parameters or business rules
- Reallocating resources
- Invoking exception processes
- Initiating improvements of ineffective processes or strategies

The benefits of timely optimization are well known at IBM.[10] Similar to analytic components, optimization components are often tailored to handle specific industries or domains.

Customer requirements determined the aspects of our framework that were implemented in the two pilot projects. Whereas the full functionality of a SaR system would encompass four aspects—sense, interpret, decide, act—our implementations fully cover the first two (sense and interpret) and the third (decide) partially.

The rest of the paper is organized as follows. The next section covers our experiences within the IMD pilot project, whose goal was the deployment of a SaR prototype in support of an automated inventory management process. The section that follows covers the PCD pilot project, which involved the deployment of a SaR prototype for demand/supply conditioning (conditioning, for short). The last section contains a discussion and our conclusions.

IMD INVENTORY MANAGEMENT

This section describes a SaR system that was developed to support an automated inventory management process within IMD. The system monitors key supply chain events and KPIs (such as the inventory turn rate [the rate at which an inventory is sold], and on-time delivery) and helps proactively manage supply chain performance, thus achieving customer service requirements with the minimum inventory.

A main component of our system is an analytical model that retrieves its data from enterprise business applications and optimizes inventory in the IMD semiconductor supply chain. It complements existing planning applications and enables predictive monitoring of critical business events, which makes the development of contingency plans possible before such events impact the value chain. For example, presented with below-target expected deliveries, managers can take steps to reduce inventory costs and improve customer service.

Business environment

IMD business managers were looking for ways to improve operational performance and reduce the expenses associated with inaccurate original equipment manufacturer (OEM) forecasts, inefficient order flow, expedited shipments (caused by shortage of parts), and obsolete inventory. Like many other organizations, IMD was faced with the challenge of responding and adjusting to supply-chain events in a timely and intelligent fashion. It recognized the need for monitoring performance in order to identify early problem areas in the end-to-end supply chain. The management resolved to enhance supply chain visibility and develop a better understanding of the transactional data representing customer needs.

Within the IMD end-to-end supply chain, there were three main processes that were targeted for improved monitoring and visibility:

1. *Demand management*—Demand management focuses on changes to the demand forecast. The demand management process at IMD had no capabilities to monitor actual OEM demand and compare it with prior demand evaluations, the demand forecasts (obtained through advanced forecasting algorithms), or the forecasts input at earlier levels. The analysis performed was entirely manual.

2. *Supply management*—Supply management focuses on changes to work-in-process or production specification parameters.

3. *Inventory management*—Inventory management focuses on changes to business policies, such as inventory "days of supply" (which is also impacted by changes in the manufacturing process, such as shorter cycle times). Based on inventory reorder points, the inventory management process controls the manufacturing of wafers, devices, and modules.

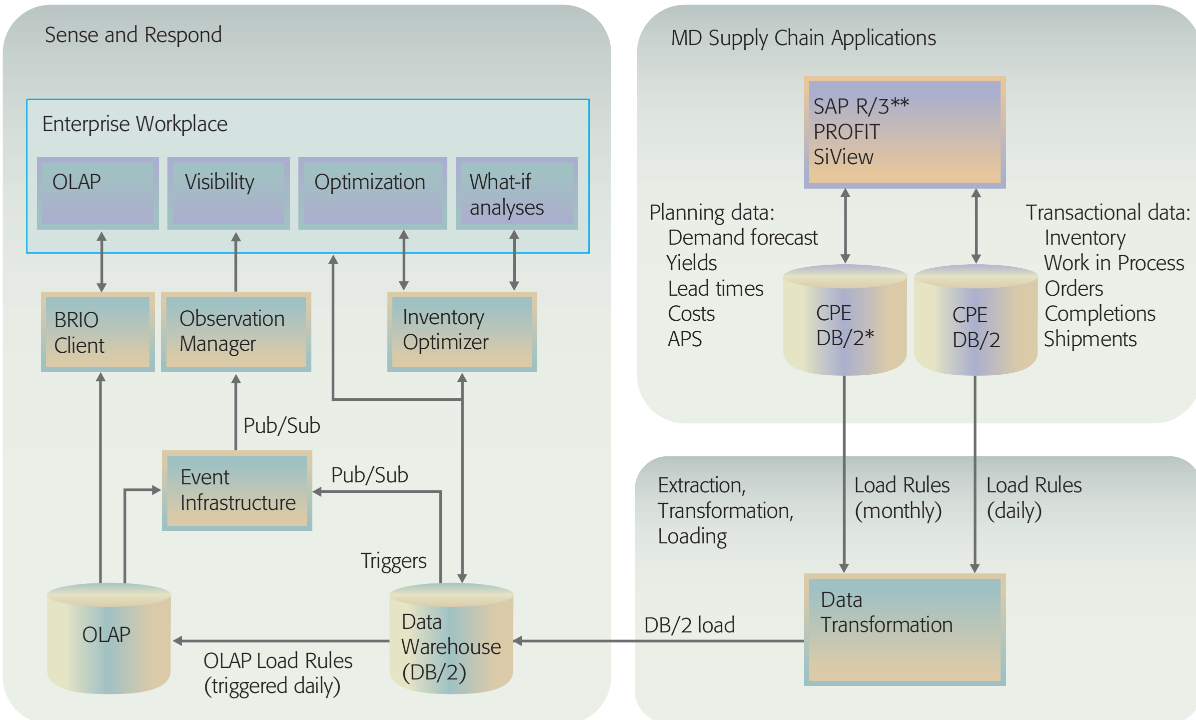
The sense-and-respond system

To improve these management processes, we developed an analytical supply-chain model that optimizes target inventory levels at different stages of manufacturing. The model helps identify potential shortages of finished goods and avoid delinquent customer deliveries. The model was able to improve IMD's inventory management process by diagnosing supply shortfalls, backlog accumulation, and inadequate inventory levels at the strategic stocking points.

We subsequently integrated the analytical supply chain model into a SaR system that identifies exceptions by monitoring demand forecast, inventory, and shipments and comparing these against predefined objectives. When the performance metrics deviate from acceptable limits, the system automatically generates alerts to inventory planners so that contingency plans can be developed before events impact the supply chain. ***Figure 1*** shows the architecture of the IMD SaR prototype.

The design and implementation of the SaR system required domain expertise in supply-chain management, data warehousing, online analytical processing (OLAP), and J2EE**-compliant (Java 2 Platform, Enterprise Edition-compliant) application servers. The system acquires and analyzes the data, detects exceptions, and invokes the appropriate actions. It has five main components:

1. A relational *data warehouse* that serves as the primary data repository of business event "trails"



APS: Advance planning and scheduling, BRIO: Corporate reporting tool (IBM internal), PROFIT, CPE, SiView: Custom applications (IBM internal), Pub/Sub: Publish/Subscribe
*Trademark or registered trademark of International Business Machines Corporation.
**Trademark or registered trademark of SAP AG.

Figure 1
Architecture of the IMD prototype

(history) from enterprise applications and advanced planning systems. It contains up-to-date profiles of business metrics for event engine processing and flexible performance reporting. The warehouse also contains operational manufacturing parameters such as bills of materials, lead times, process yields, demand forecasts, and supply commitments, which are used as inputs to the inventory optimization module.

2. A *data transformation* component that performs automated data extraction, transformation, and loading functions for gathering production planning and execution data from enterprise applications (such as the data marketed by SAP), legacy systems, and supply-chain-planning solutions (such as the data marketed by i2 Technologies, Inc.). Events and transactions that occur within the business, such as a release of a purchase order, receipt of a sales order, shipment of a customer order, or the completion of a manufacturing lot, demand high levels of data integrity and transaction processing that are fully supported by our infrastructure.

3. An *observation manager* event-driven component that determines whether events violate enterprise business commitments. Violations of business commitments are detected by monitoring standardized and configurable KPIs, and performance alerts are posted in real time.

4. An *inventory optimizer* module that provides business intelligence and analytics for improving the performance of the enterprise. It incorporates existing business processes and cost structures, and it recommends optimized inventory policies. The recommendations can be viewed by business process owners. They can see the expected impact of planning decisions, assess the profit risk and rewards of proposed actions, and evaluate alternative options.

5. An *enterprise workplace* personalized dashboard-type portal that provides a view of the overall

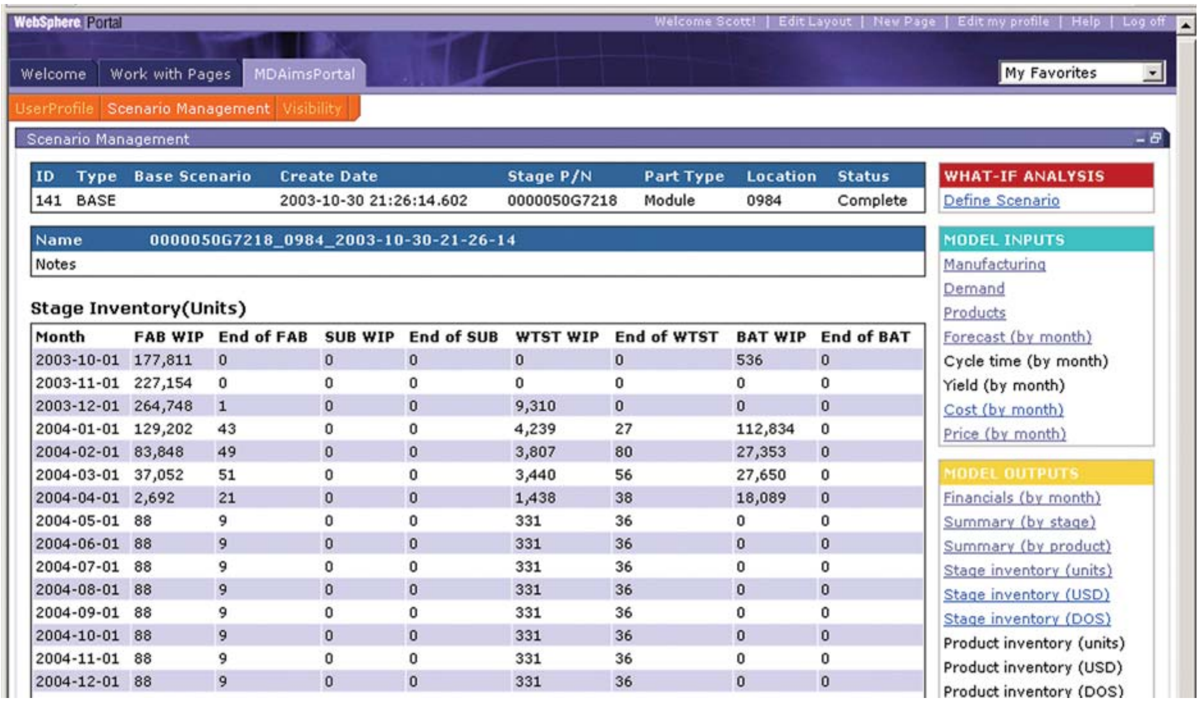


Figure 2
IMD enterprise dashboard

health of the business. (A dashboard is an interactive user interface resembling an automobile's instrument panel.) Through a Web browser-configurable dashboard, portlets deliver real-time visibility into key business process metrics (a portlet is an application that creates and displays a small window on a portal page). In addition, the dashboard provides "drill-down" capability for analyzing the cause of the error and provides suggested responses to resolve the situation. It supports role-based portal views in a multi-user environment, including portal views for inventory analysts, product line managers, supply-chain executives, and financial executives. Furthermore, the dashboard allows users to carry out "what-if" analyses and assess the impact of business decisions before they are implemented.

The screenshot in ***Figure 2*** illustrates a dashboard that displays KPIs and analysis capabilities provided by the system.

Inventory optimization model

The supply chain optimization model applies to typical semiconductor manufacturing operations. It has a three-echelon (three-level) structure with additional assembly functions implemented in the downstream node. The stages in manufacturing are FAB (wafer fabrication), WFT (wafer test), SUB (substrate) and BAT (bond-assembly and test). ***Figure 3*** shows a graphical illustration of the model and the solution approach.

Manufacturing starts at FAB for a single wafer design. The tested modules (finished products) after the BAT phase exhibit varying processing speeds and are assigned different part numbers according to their speeds (fanning arrows in Figure 3 illustrate this step). In case of shortage, high-speed products can be downgraded to satisfy demand for low-speed products. Demands for multigrade products are viewed as nonstationary processes with independent random variables and normal distributions. Prices and manufacturing costs can change over time as well. Inventory is replenished according to base-stock policies in all stages. The objective is to minimize inventory subject to a service requirement measured as on-time delivery to customers within an allowed lead-time window.

The model uses demand forecasts, manufacturing cycle times, yields, costs, lot sizes, inventory

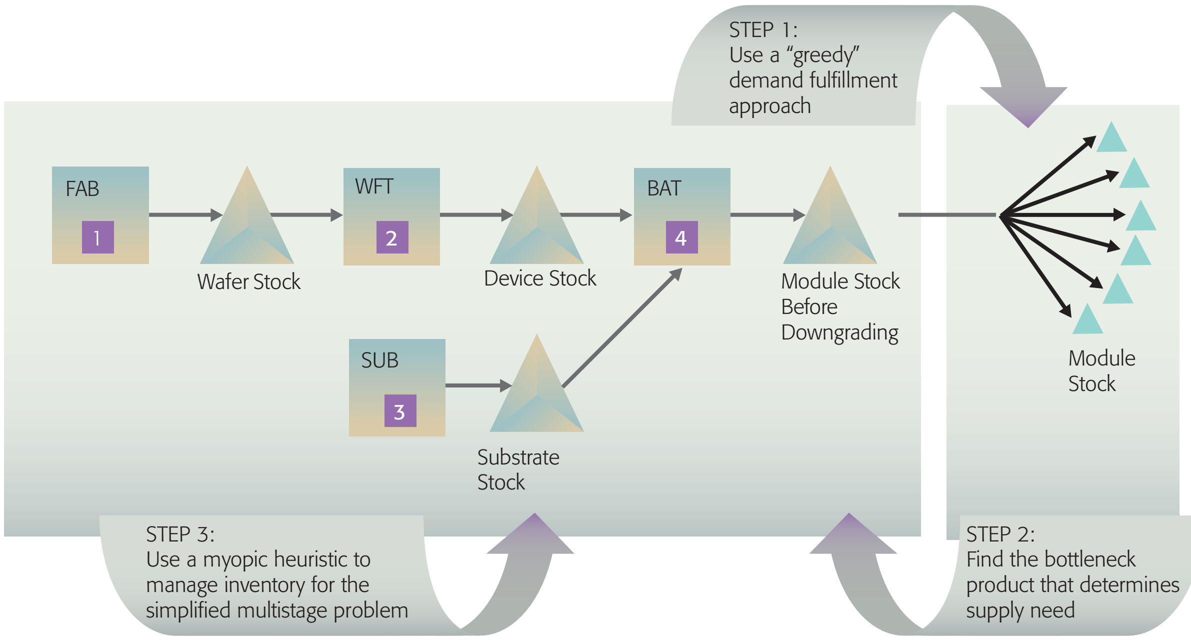


Figure 3
Inventory optimization model

policies, contractual buffers, customer service targets, product prices, and the rates of change in prices and costs. Based on all these input parameters, it calculates and reports operational and financial performance for business managers and inventory planners. The performance reports comprise numerous financial and operational performance metrics. The hierarchical diagram in ***Figure 4*** illustrates the way various variables are calculated. The solid lines indicate how a "parent" (higher level) KPI is determined based on the values of the "children" (lower level) KPIs. Some lower-level KPIs have more impact than others on the value of the higher-level KPI, and this is indicated by "decision weights." Values for these parameters are projected weeks or months into the future. The analytical model also determines optimal days-of-supply policies at strategic stocking locations in wafer fabrication and module assembly and testing.

Because the model has to handle numerous parameters, the calculation of optimal inventory levels is extremely difficult.[11] A careful approach to simplifying the problem without significant loss of accuracy had to be developed. The simplification was achieved through the following steps:

1. Simplify the demand/supply matching problem at BAT:
   a. Use a greedy supply allocation starting from the highest speed product.
   b. Always downgrade when needed.
2. Divide the problem into two separate problems:
   a. Calculate supply needed to meet mean demand at BAT.
   b. Use supply needed as the demand of a single product.
   c. Simplify to a multi-echelon, single product problem.
3. Use myopic solutions (base stocks that minimize inventory cost during lead time).

The calculation of the optimal inventory policies follows a heuristic that consists of the preceding three steps. First, the "greedy" allocation of supply tries to satisfy all the demand by downgrading inventory as much as possible. This step is illustrated by the arrow labeled STEP 1 at the top in Figure 3. Then, based on this allocation, it determines the product grade that causes the supply bottleneck, illustrated by the arrow labeled STEP 2 at the bottom of the figure. This effectively reduces the problem to a single-variable problem, which is easier to solve. In the last step, the optimal initial

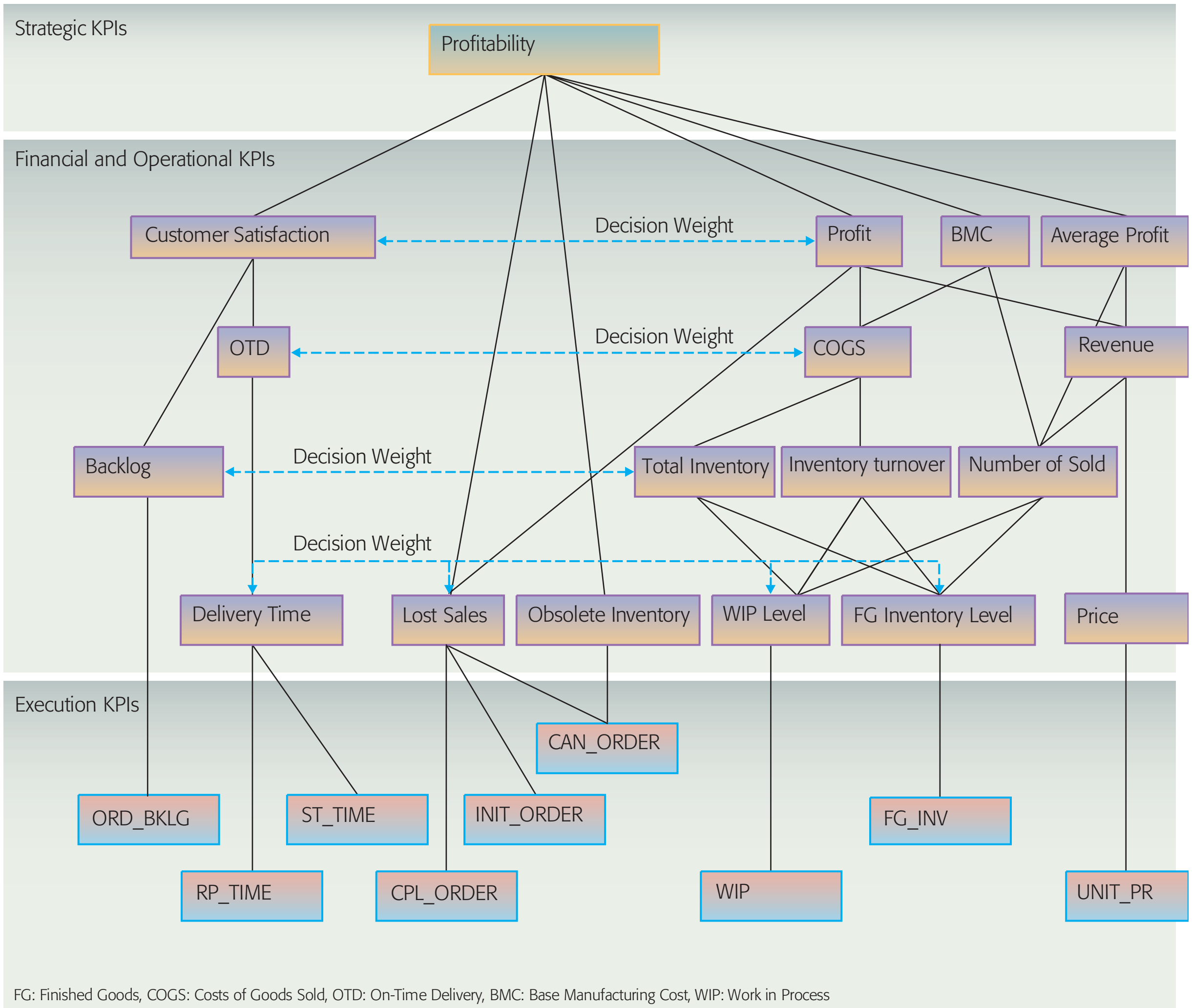


Figure 4
Calculation of financial and operational parameters by the inventory optimization model

quantities for each stage needed to supply sufficient stock for the bottleneck grade are determined. This method was tested in detail, and it provides a near-optimal solution.[12]

The model had to consider additional variables, such as manufacturing lot sizes, contractual buffer obligations for finished-product inventories, random lead times, and manufacturing yields. All these increase the complexity of the model and make the calculations harder. For further details on how we handled these challenges, see Reference 13.

We tested the performance of our optimization engine using real data for a subset of products. The sample covered a wide range of product groups. We collected all the necessary data needed to run our engine for each of these products. The data included the most current weekly forecasts, forecast errors, current yields, and manufacturing cycle times and their variability for each key stage in the manufacturing process (i.e. fabrication, wafer test, substrate, and bond and assembly stages). We ran our inventory optimization engine based on this data for all stages in order to calculate the optimal inventory policies that minimize overall system inventory levels and achieve target customer on-time-delivery objectives. The results showed that optimized inventory policies would give lower overall inventory and deliver more consistent service levels as compared to the policies that were in use at the time of the tests. ***Figure 5*** shows the inventory reduction

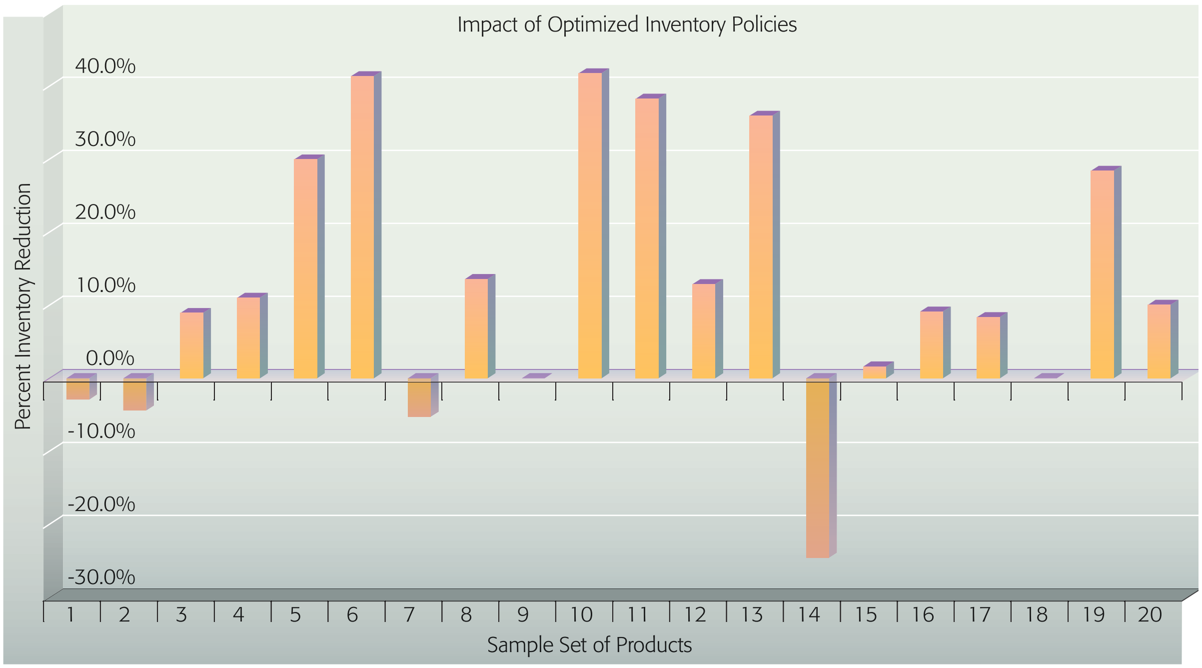


Figure 5
Inventory reduction projected for a selected set of products

projected by using optimal inventory policies for a sample set of products (for confidentiality reasons, we omitted the product names and dollar value of inventory reduction).

Optimal policies took into account product-specific characteristics when policies were determined. Current policies did not reflect the impact of such characteristics. For instance, some products had much higher demand forecast than others, and some had much higher lead-time variability than others. Because of these characteristics, the recommended safety stocks were in fact higher than the safety stocks being held for some products—as can be seen in Figure 5. This brought much more data-driven intelligence to the process of determining safety stocks, and as a result, improved the consistency of on-time-delivery performance. Of course, the SaR system, by using such an engine, can update the entire safety stock levels almost instantly and much more precisely as a response to changing demand and manufacturing characteristics—something no manual process can compete with. In summary, our tests showed that the benefits of data-driven optimization of inventory policies were twofold: more consistent and accurate on-time-delivery performance and an overall reduction in inventory levels.

PCD DEMAND/SUPPLY CONDITIONING

Business performance can often be improved by the intelligent application of changes to supply-chain variables. Such actions, known as *conditioning* actions, may touch on three aspects of supply-chain operations: demand, supply, and product offerings. In this section we first describe the PCD business environment and the ongoing efforts to improve supply-chain operations, and then we describe our experience in deploying the SaR prototype in support of the existing conditioning activities. We also discuss a business analytics contribution made by our prototype involving an order trend model.

Business environment

The PCD conditioning project began as an end-to-end analysis of IBM's ability to deliver PCD products to customers. One of the results of this analysis was an initiative to minimize the impact of constrained supply on product deliveries. The PCD conditioning project team was chartered in 2002 to achieve this goal.

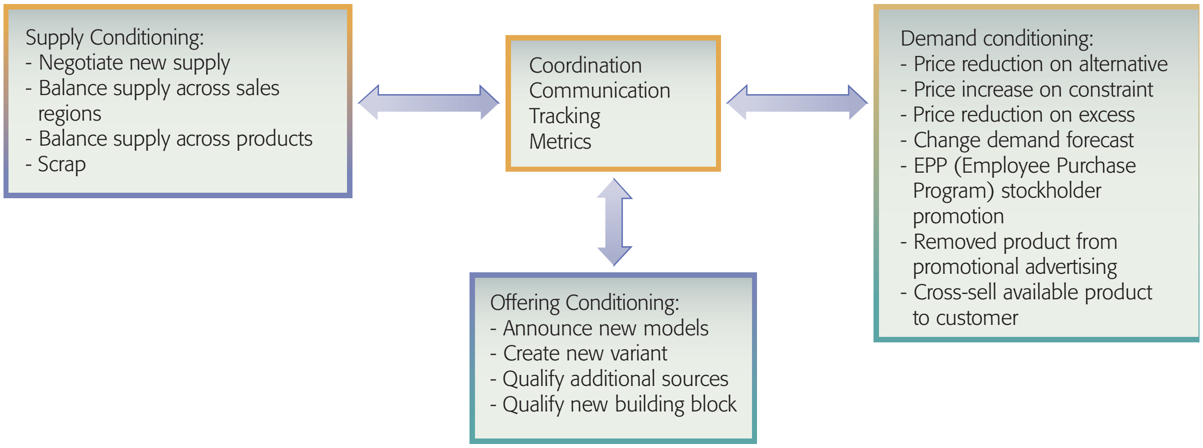


Figure 6
PCD conditioning process

The success of the conditioning process depends on all the aspects involved. The supply should be somewhat flexible in order to react to customer demand, which is not totally predictable. The demand, as reflected in the sales plan, should be flexible in the sense that it can be affected in case of supply imbalances. The product offerings must also have flexibility so that reasonable alternatives can be offered to customers in case of supply imbalances.

The PCD conditioning system illustrated in ***Figure 6*** represents a management system for affecting the three principal dimensions of PCD conditioning. As shown in the figure, there is one control and coordination component and three functional components, for each of the principal dimensions involved.

- *Supply conditioning*—a collection of supply-based actions to resolve supply imbalances, such as negotiating for additional supply or rebalancing supply between “geographies” (sales regions).
- *Offering conditioning*—a collection of offering-based actions such as creating new product models that either use surplus components or substitutions for demand-constrained components.
- *Demand conditioning*—a collection of actions that affect demand in a desirable way. Examples of demand-conditioning actions are a sale promotion for a surplus product or reducing the price on an alternative product in order to transfer demand from a constrained product.

When a supply imbalance is to be addressed, the degree to which each of the three principles is involved in the solution is carefully considered and optimized. Because the process involves people in different organizations at dispersed locations, communication among teams is important so that all participants are in agreement and have a clear understanding of what needs to be done. When the solution arrived at is to be executed, monitoring is needed to ensure that the solution is executed properly. Finally, KPIs involving customer orders must be tracked to ensure that the solution is effective.

The PCD conditioning process is owned and executed by the Integrated Supply Chain (ISC) Worldwide Fulfillment Operations team in Research Triangle Park, North Carolina. The execution of the process revolves around a weekly “core team” meeting. The core team has a team leader and includes representatives from the PCD Brand, PCD Operations, ISC Procurement, and PCD Finance organizations. This team identifies supply imbalances, creates a conditioning plan in partnership with the three “geographies” (regional sales) organizations, and manages the execution of the plan. ***Figure 7*** shows the execution of the conditioning process before the deployment of the SaR prototype.

The conditioning process was put in place in August 2003. Although the process has been successful, it can be improved in many areas. One such improvement involves the timely and proactive

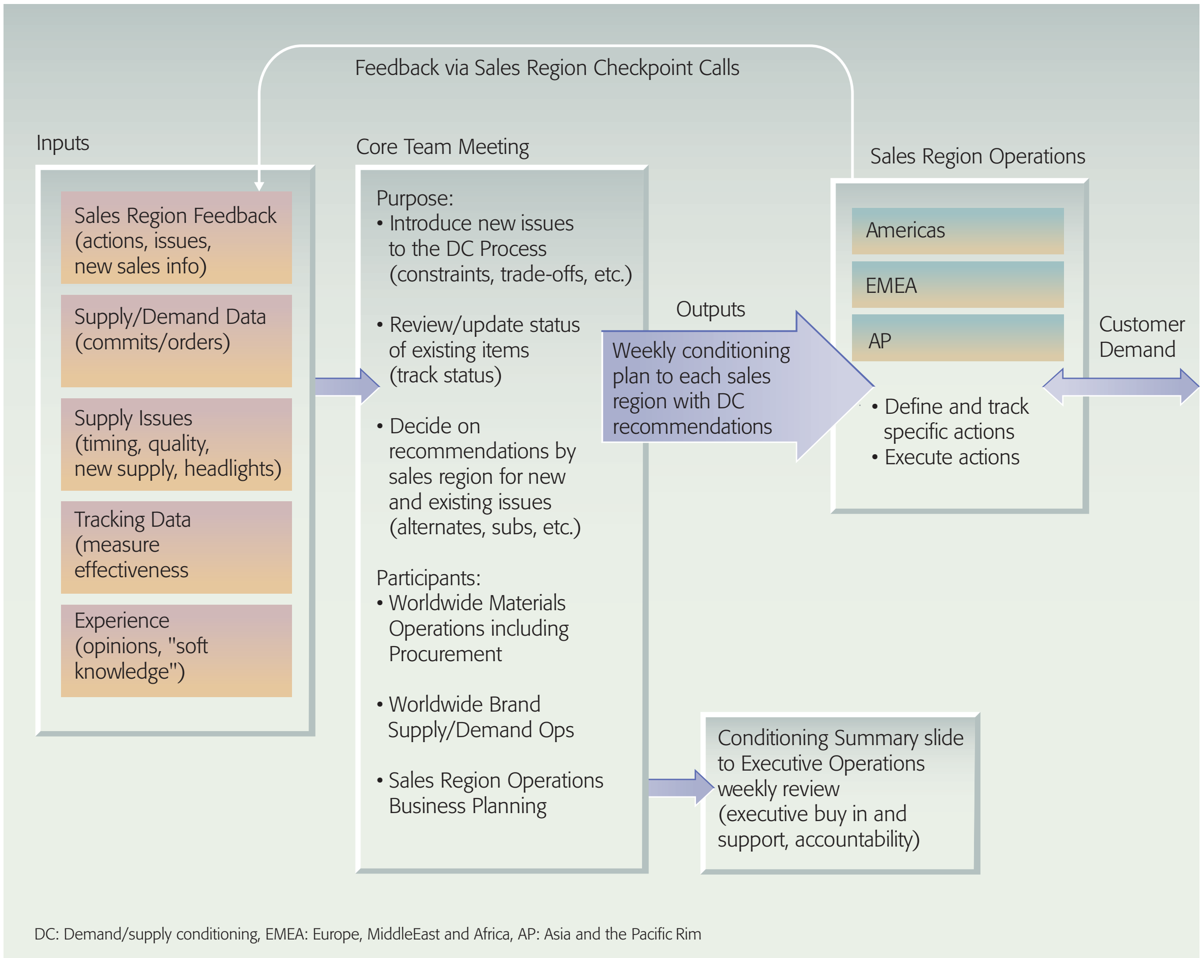


Figure 7
Execution of the PCD conditioning process

identification of supply imbalances, as proactive identification makes it possible to develop more effective plans. The immediate target was identified: analyzing trends and patterns in customer orders, with the goal of predicting demand forecast inaccuracies four weeks in advance. This order-trend analysis is a critical element of the SaR pilot for the PCD conditioning process and is discussed next.

The sense-and-respond system

The main advantage of a SaR system for demand conditioning is to offer a proactive means of control for the supply chain by triggering corrective actions before a problem accelerates into a disaster. *Figure 8* shows the monitoring and control model for the PCD conditioning process as implemented in the pilot.

The SaR system uses available data, such as forecasts, customer orders, and supply commitments, and aims to provide an early warning system for conditioning. An important innovation in this pilot is a new algorithm that identifies potential problems by using historical information and future indicators to forecast trends for customer orders. The order-trend analysis is used to compare trends to the demand forecast as a lead indicator of future supply imbalances.

The system receives daily customer orders (that is, order loads) and shipments, weekly supply commitments, and demand forecasts. It correlates and analyzes the information, detects early “headlights” into supply constraints and excesses, alerts the appropriate role players, and recommends correc-

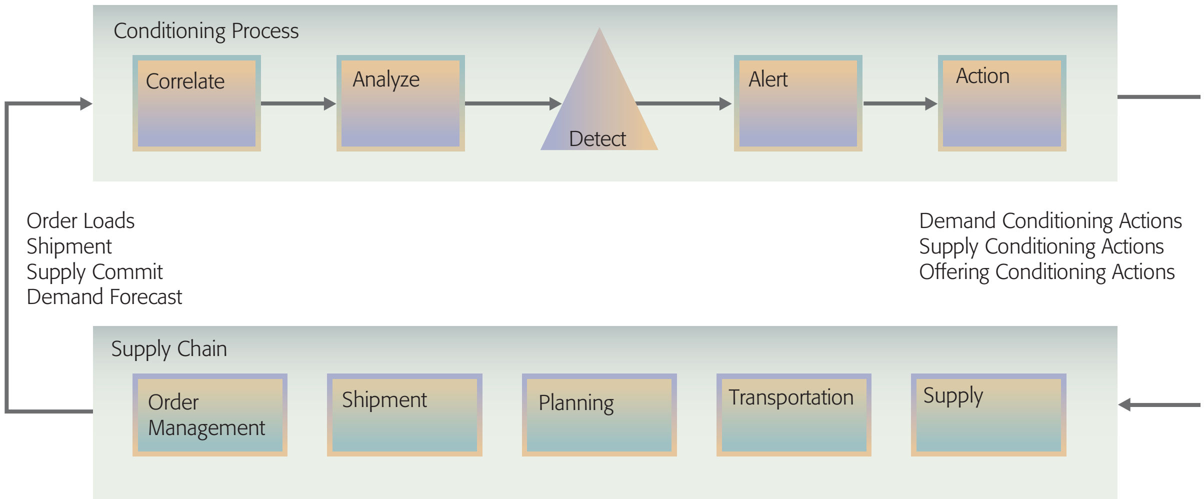


Figure 8
Monitoring and control for the PCD conditioning process

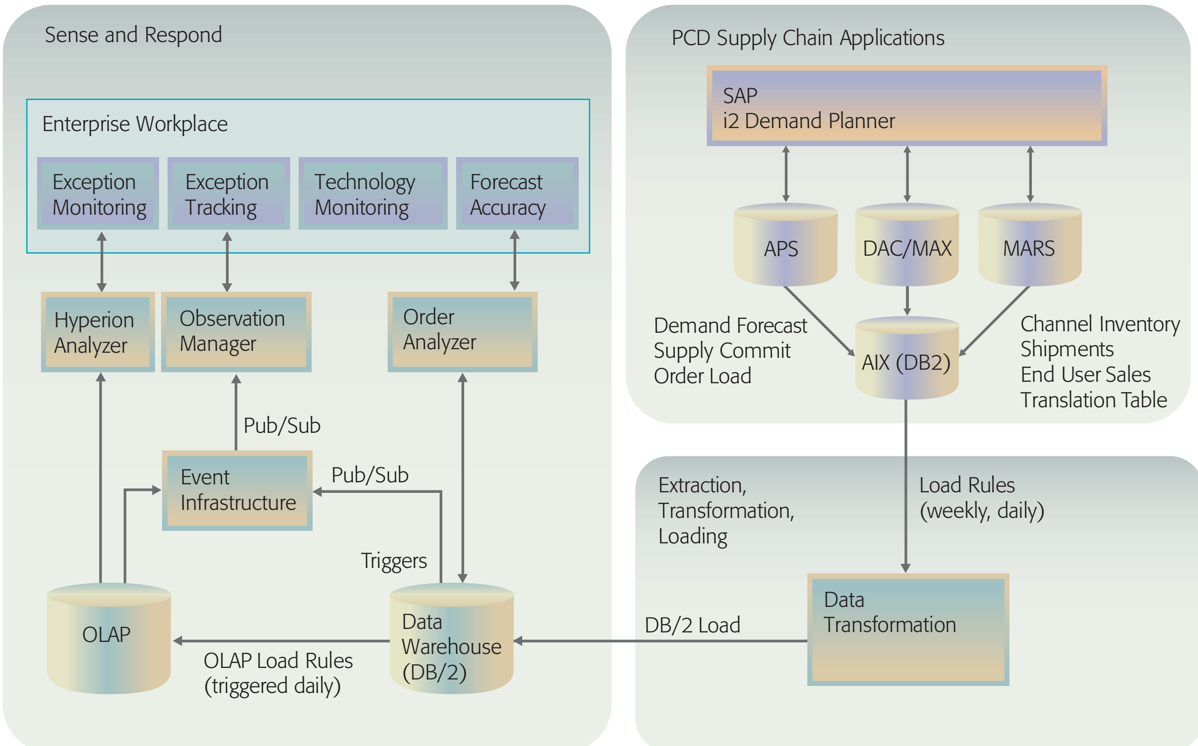
tive actions. It generates order trends based on historical and future demand-related indicators to create volume projections over a time horizon. For assessing the accuracy of the order trend and comparing it to the demand forecast, the system analyzes the demand forecasts and order trends predicted over the trailing 13 weeks and compares it with the actual results. It uses industry standard metrics like MAPE (Mean Absolute Percentage Error) and MFE (Mean Forecast Error) to report the accuracies of the forecast and order trend.[14]

Tracking capabilities are provided that record a snapshot of the data at the time the alert was generated and compare it to the data over a predetermined time horizon. This provides benefits in two respects:

1. It can track the effect of actions invoked in response to a business exception.
2. It builds a rich history of actions in the data warehouse, on which intelligent mining operations can be performed later in order to learn and recommend future actions.

The architecture as described in ***Figure 9*** consists of the following components:

- *Extraction, transformation, loading*—Demand and supply information is made available by transactional systems and processed for the SaR system by applying extraction, transformation and loading rules. In this process, which uses a bill-of-materials document, the transactional data available at the system-level data are transformed into records corresponding to product parts.
- *Order-trend analyzer*—This component generates order trends and provides early warning about supply constraints and excesses. It combines standard forecasting techniques with future indicators, such as order coverage and channel inventory, in order to identify repetitive historical patterns of orders in a four-to-eight-week time horizon, factoring in seasonality and product life cycle.
- *Observation manager*—This component detects business exceptions by comparing various indicators of supply and demand over a rolling 13-week time horizon. Business rules are defined for the cumulative differences between supply and demand, which result in detecting supply shortages or overages over a rolling forward-looking 13-week time horizon.
- *Data warehouse and OLAP*—A relational data warehouse that captures the order loads, shipments, and planning data (at both the system and the product-part level) is augmented by an OLAP component in order to provide root-cause-analysis capabilities.



APS: Advanced planning and scheduling, DAC/MAX: Reporting tool (Direct Alliance Corporation), MARS: Custom application (IBM internal), Pub/Sub: Publish/Subscribe

Figure 9
Architecture of the PCD prototype

- *Enterprise workplace*—A dashboard-type portal that provides the conditioning team an end-to-end view of the imbalances between supply and demand. It allows for customization and administration by different role players. It also recommends actions based on alerts generated and provides capabilities to track these actions, for enhanced effectiveness. The screenshot in ***Figure 10*** illustrates the tracking capabilities provided by the enterprise dashboard.

Order-trend model

Unlike long-range forecasting, the goal of order-trend analysis is to identify repetitive historical patterns of customer orders and obtain accurate short-term projections through emphasis on the data available in order execution systems, such as actual demand and customer order inflow. Coupled with improved data integration and the Web-based management dashboard, order-trend analysis enables a current view of key supply and demand metrics for each component of a PCD product.

Order-trend analysis is based on a model that uses historical and future-demand-related indicators to project short-term order trends. The model estimates the effects of seasonality, order skew within a quarter, product life cycle, and repetitive order trends from historical data. It also provides point estimates, percentiles, and confidence intervals for risk management. Order-trend analysis combines traditional statistical forecasting techniques with demand-related indicators visible in the current time period. It improves on baseline forecasts. This is because the order trends are operational forecasts that provide a more accurate picture of demand for the next four-to-eight weeks, which is the most critical time for deployment. The three indicators that are integrated into the analysis are:

- *Total order load:* the current amount of unfilled customer orders with a customer-requested shipment date some time in the future.
- *Order coverage:* the current amount of supply-committed customer orders with a confirmed future shipment date.

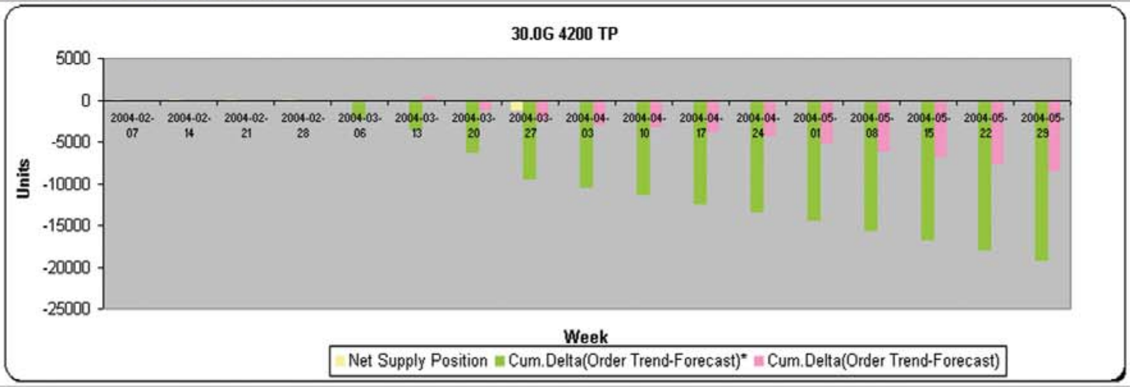
Exception Tracking

Quoted At : 2004-03-05

Summary Detail

Date	Alert Description	Brand	Commodity	Technology
2004-02-28	Supply Shortage(Order Trend vs. Demand Forecast)	ThinkPad	Hard Drives	30.0G 4200 TP

Technology Tracking Summary



Technology Tracking Detail

Week	CRSD Volume (units)	Net Supply Position (units)	Order Trend* (units)	Demand Forecast* (units)	Cum Delta* (units)	Order Trend (units)	Demand Forecast (units)	Cum Delta (units)
2004-02-07	931	-	-	-		-	-	
2004-02-14	899	-	-	-		-	-	
2004-02-21	1,054	-	-	361		-	-	
2004-02-28	1,289	(1,163)	-	412	0	-	3,033	
2004-03-06	681	-	2,271	255	(2,016)	-	1,300	0
2004-03-13	-	-	1,643	216	(3,443)	808	1,153	345
2004-03-27	-	-	3,300	101	(9,420)	2,133	765	(2,455)
2004-04-03	-	-	1,155	168	(10,407)	917	477	(2,895)
2004-04-10	-	-	1,041	80	(11,368)	810	477	(3,228)
2004-04-17	-	-	1,090	80	(12,378)	837	328	(3,737)
2004-04-24	-	-	996	80	(13,294)	767	326	(4,178)
2004-05-01	-	-	1,222	80	(14,436)	1,022	-	(5,200)
2004-05-08	-	-	1,189	-	(15,625)	914	100	(6,014)
2004-05-15	-	-	1,056	-	(16,681)	817	100	(6,731)
2004-05-22	-	-	1,297	-	(17,978)	1,011	100	(7,642)
2004-05-29	-	-	1,122	-	(19,100)	914	100	(8,456)

* Data at the time when the Alert was detected

Complete Tracking

Disable Conditioning Tracking

New Conditioning Action

Select Conditioning Category:

<Select Category>

Enter a new response:

Select Conditioning Action:

Add Conditioning Action

Conditioning History

Date	Action Category	Action Subcategory	User	Details

Figure 10
PCD enterprise dashboard

Table 1 Forecast accuracy results for a subset of PCD commodity groups

	Double Exponential Smoothing		Winter's Method		Order Trend Analysis	
	(monthly)	(quarterly)	(monthly)	(quarterly)	(monthly)	(quarterly)
Hard Drives	29.4%	17.5%	27.0%	17.8%	23.3%	15.0%
Optical Drives	34.8%	23.2%	32.1%	23.0%	29.2%	22.2%
Planars	46.5%	44.9%	35.3%	35.5%	34.5%	29.3%

Note: Entries in the table are in MAPE (mean absolute percentage error)

- *Channel inventory:* the current amount of inventory stocked at a business partner's warehouses to fill future customer demand.

The order trend analysis projects order volumes for IBM PC components over a future time horizon of N weeks—weeks $c+1, c+2, \ldots, c+N$, where c denotes the current week. The base algorithm proceeds in three steps as outlined next:

Step 1. Analyze historical order coverage.
   a. For all historical quarters $q = 1, 2, \ldots, Q-1$, compute the order coverage ratio in week c of quarter q.
   b. Compute the sample mean and sample standard deviation of the order coverage in week c in the current quarter Q.

Step 2. Compute cumulative order trends.
   a. Compute the first and second moment of the random variables $G_c^{(i)}$ which denote the cumulative order volume for weeks $c+1$ through $c+i$ for $i = 1, 2, \ldots, N$.
   b. Adjust the cumulative order volumes $G_c^{(i)}$ by taking into account the actual sales transactions, order load, and channel inventory available in the supply-chain execution systems.

Step 3. Compute weekly order trends. Compute the first and second moment of the random variables $F_c^{(i)}$, which denote the marginal order volume for week $c+i$ from the cumulative order volumes $G_c^{(i)}$ determined in Step 2.

The analytical model is executed daily for producing new order trends. The order trends are compared to the official demand forecast as a lead indicator of future supply imbalances. Part of the weekly conditioning process is to select technologies where this indicator shows a potential issue, and review the forecast with the U.S. "geography" (sales region) planning team.

To investigate the usefulness of the alerts and to determine whether the PCD conditioning team should use order-trend analysis rather than traditional forecasting methods, we conducted a numerical study. The order-trend model is compared with two standard time-series-based forecasting methods: exponential smoothing with a linear trend (double exponential smoothing) and exponential smoothing with a linear trend and seasonality (Winter's method).[14] The forecast error was evaluated by calculating a MAPE of the forecasted volumes generated by each method against actual sales volumes. All point forecasts were generated eight weeks ahead of the period being forecasted to account for frozen zones imposed by component suppliers.

Table 1 summarizes the results obtained for a set of 15 PC components in three commodity groups including hard drives, optical drives, and planars. The data show that the order-trend model produces a lower MAPE than the two traditional methods for both monthly and quarterly volume projections. We further observe that the order-trend model tends to be much more responsive in capturing sales trends as well as intra-quarter seasonality. ***Figure 11*** illustrates the weekly point forecasts that resulted from an order-trend analysis of a selected optical drive.

A number of functional enhancements are currently under way. First, we are enhancing the order-trend model to provide improved volume predictions for new product introductions and end-of-life situations based on technology transitions maps. Second, we are developing capabilities to record the actions that the conditioning team is taking to resolve supply imbalances in order to form a knowledge base for data mining that will assist decision making for future supply imbalances. And third, we are building advanced analytics that go beyond demand planning and optimize inventory buffers at

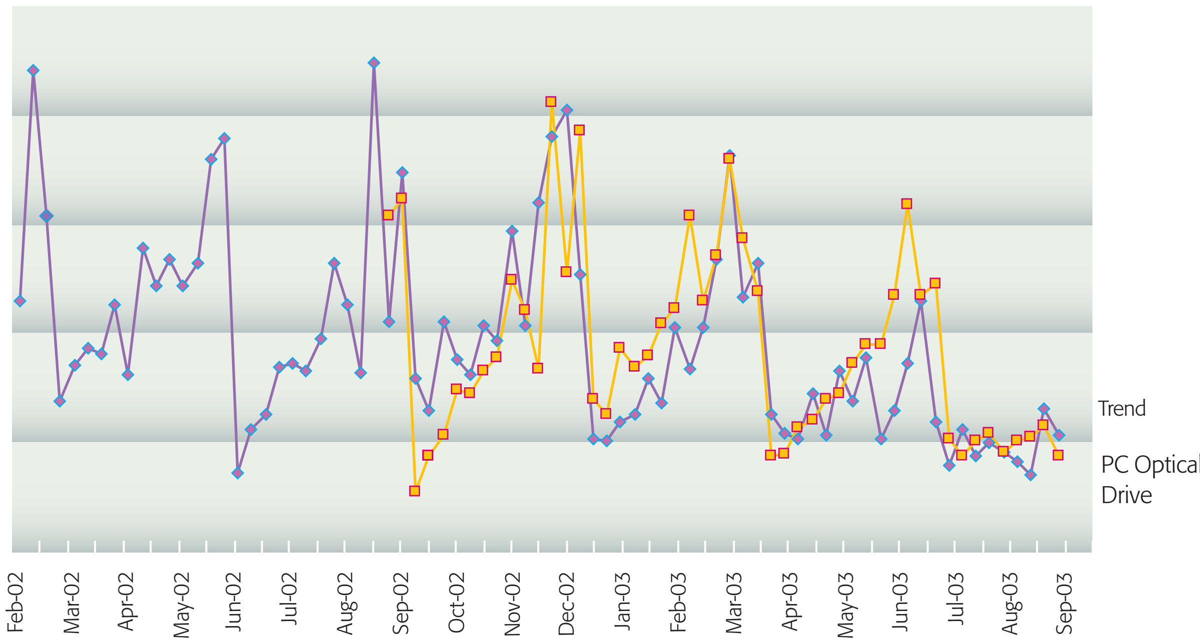


Figure 11
Comparison of order trend projections and actual sales of a PC Optical Drive

upstream component suppliers. The analytics will facilitate the monitoring of fulfillment activities and provide metrics and alerts that focus attention on serviceability issues.

Discussion and conclusions

This paper describes a technical framework and the supporting operating environment for implementing SaR systems, including two SaR prototypes that were applied to two IBM businesses, IMD and PCD.

IMD conducted a pilot implementation of optimal inventory policies recommended by our optimization engine from the fourth quarter of 2001 through the third quarter of 2002. Initial and updated policy recommendations were made during this time on 400 parts. For those parts on which the recommendation was implemented, an inventory reduction of 4 percent was documented. On-time delivery performance achieved during the pilot was substantially above the target levels.

PCD implemented three strategic initiatives in 2003 to improve profitability

1. Shortening the interval between the time the customer signs a contract and the time the customer can place an order;
2. Reducing the time from order entry to order delivery;
3. Improving planning and forecasting through conditioning across the enterprise.

Although it is too early to assess the financial impact, the conditioning initiative is estimated to provide a 10–15 percent increase in profitability through improved serviceability and customer retention and reduced inventory writedowns in a market environment where assets depreciate .5 percent per week.

The PCD conditioning process benefits the customer through improved delivery times, and it benefits IBM through higher inventory turns, which lead to higher profits. The benefit is measured at a high level by the number of orders that are unfulfilled due to unavailable supply. Over the past year the conditioning process has improved this KPI by 50 percent. By improving our ability to handle demand-supply imbalances through order-trend analysis and the resulting alerts from the SaR system, we expect to improve this KPI by another 40 percent.

The pilot projects described in this paper aim at improving the responsiveness of both the planning and the execution processes in the supply chain. The prototypes were built using components developed by the IBM Research Division.

Our experience in these pilot projects shows that analytics can be embedded into pluggable components that can be customized then for specific applications. In the IMD pilot we built an inventory optimizer that helped build contingency plans and make adjustments to inventory levels. In the PCD pilot we built an order analyzer that analyzes trends and issues proactive alerts.

We are currently developing new SaR pilots in other domains, including one in the customer relationship management (CRM) domain and one in the insurance industry. In the CRM project we are combining the SaR architecture with data-mining techniques in order to make the organization more proactive in managing all phases of the customer life cycle. The solution optimizes the placement of customers into appropriate segments and improves customer segment sales for enhanced revenue, profit, and customer satisfaction. Specifically, we developed a new predictive modeling technique that discovers optimum customer moves through customer segments based on known customer spending patterns and analytical correlations. The technique allows the proactive detection and alerting of dormant customers (as well as profiling customers who have left) so that an organization can act to retain customers at risk of leaving, thus reducing churn and attrition. In the insurance project we also use the four-layer modeling framework and transform the operational specification of the business into a platform-independent execution model and then into a platform-specific implementation model. Business performance models are incorporated into each level to validate the promise of the model-driven architecture.

The ultimate value of the SaR approach can be summed up as *business responsiveness*—the ability to quickly and effectively adapt to impending threats and opportunities. Although few businesses today can rightfully claim to be highly responsive, this concept is known and found under many names, such as on demand enterprise, adaptive enterprise, business agility, real-time enterprise, zero-latency enterprise, and enterprise of the future. We should point out that the ability to respond in real time to events does not ensure success. A single-minded focus on quick decisions ignores the importance of those decisions being intelligent and strategic. Sometimes "near real-time" or "right-time" responses are needed so that all key decision parameters can be assessed and effective synchronization established across multiple time zones and interlocked workflows that incorporate manual as well as automated processes.

Maintaining alignment between business design and IT solutions through a model-driven architecture has the potential to greatly reduce the "time to value" of business transformations. This alignment is a significant step towards closing the infamous "business-IT gap."[2] This linkage also has the potential to provide real-time visibility of business operations. This visibility, together with business-level optimizations and what-if analyses, enables the continued improvement of business operations.

As with any major business transformation, becoming a SaR enterprise requires a strategy and a roadmap. For most companies, a SaR roadmap involves a progression through four levels of maturity[15]:

1. *Automated*—The enterprise strives to automate business processes and improve the use of resources within individual functions and divisions.

2. *Visible*—Enterprise-wide business processes are established, helping individual organizations execute their roles within complex management processes. Information is integrated, and standardized measurements are clearly defined, enabling effective performance management throughout the extended enterprise.

3. *Controlled*—The enterprise identifies strategic customers, suppliers, and business partners. Service level agreements and scorecards are used, and corrective actions are taken when operational performance falls below acceptable limits. Event monitoring and alerts based upon rules are integrated into the overall workflow for decision support.

4. *Adaptive*—The enterprise works strategically with customers, suppliers, and business partners to determine jointly agreed-upon performance targets. Value chain activities are monitored

against the performance targets and thresholds, and deviations are automatically detected. Collaborations among the enterprises are defined and supported by advanced analytics and automated decisions to optimize cross-functional performance objectives.

An industry research study conducted by the Supply-Chain Council[16] recognized that supply-chain organizations must progress in sequence through the maturity levels by building on practices that they have established at each level. The study conducted benchmarks of supply-chain practices in the high-technology industry which confirmed that attempts to advance without a base of firmly established practices at preceding levels are rarely successful.

The four-level maturity model imposes a serious challenge because in the first few phases, the return on investment (ROI) is difficult to predict and measure. The expense of transforming the business to the automated and visible stages may at times outweigh short-term savings. The customer may have to reach the controlled stage in order to see a clear financial benefit.

A good way to avoid this problem is to introduce analytics and optimization into the process very early—perhaps in a manual, nonthreatening way. In many instances we have delivered spreadsheet-based analytics and optimization to customers in a pilot environment without drastically changing the existing process. Normally, after some period of time we can measure savings or revenue growth. We then make the case for deploying the new analytics and optimization in a dashboard with associated process transformation and visibility. This is much easier to sell because a baseline for the value of the technology has already been established.

ACKNOWLEDGMENTS

Lianjun An, Harriet Cao, Henry Chang, Opher Etzion, Dagan Gilat, Paul Huang, Jun-Jang Jeng, John Kearney, Grace Lin, Vladimir Shcherbina, Josef Schiefer, Roger Tsai, and Haifeng Xi made important contributions to the technical architecture. Anil Nigam, Nathan Caswell, Robert Guttman, Santhosh Kumaran, Frederick Wu, and Prabir Nandi greatly contributed to the four-layer modeling approach. Blair Binney, Alvin Gillis, Anthony Humphrey, John Konopka, Paul Leslie, Doug Melton, Steve Molella, and Tracey Pilkinton led and managed sense-and-respond pilots within the IBM manufacturing divisions. Kevin McAuliffe, Kumar Bhaskaran, Joachim Frank, and Eric Wayne, all of the IBM Software Group, led and designed the process models and the sense-and-respond framework.

*Trademark or registered trademark of International Business Machines Corporation.

**Trademark or registered trademark of Object Management Group or Sun Microsystems, Inc.

CITED REFERENCES

1. S. Haeckel, "Adaptive Enterprise: Creating and Leading Sense-and-Respond Organizations," *Harvard Business School Press*, Cambridge, MA (1999).
2. S. Kumaran, "Model-Driven Enterprise," *Proceedings of the Global EAI Summit*, Banff, Canada (May 2004), pp. 166–180.
3. OMG Model Driven Architecture, Object Management Group, http://www.omg.org/mda/.
4. R. S. Kaplan and D. P. Norton, "The Balanced Scorecard—Measures that Drive Performance," Harvard Business Review, 71–79 (January–February 1992).
5. R. J. Thierauf and R. C Klekamp, *Decision Making through Operations Research, 2nd Edition*, John Wiley and Sons, New York (1975).
6. P. Herzum and O. Sims, *Business Component Factory: A Comprehensive Overview of Component-Based Development for the Enterprise*, John Wiley and Sons, New York (2000).
7. K. Gottschalk, S. Graham, H. Kreger, and J. Snell, "Introduction to Web Services Architecture," *IBM Systems Journal* **41**, No. 2, 170–177 (2002).
8. E. Christensen, F. Curbera, G. Meredith, and S. Weerawarana, *Web Services Description Language (WSDL) 1.1*, World Wide Web Consortium (2001), http://www.w3.org/TR/wsdl/.
9. Business Performance Management, IBM Corporation, http://www.ibm.com/software/info/topic/perform/resources.html.
10. B. Dietrich, N. Donofrio, G. Lin, and J. Snowdon, "Big Benefits for Big Blue," *OR/MS Today* (June 2000).
11. G. Gallego, K. Katircioglu, and B. Ramachandran, *A Note on the Inventory Management of High Risk Items*, Research Report RC23038, IBM Research Division, Yorktown Heights, NY (July 2002), to appear in Operations Research Letters.
12. G. Gallego, K. Katircioglu, and B. Ramachandran, *Inventory Management with Multiple Grade Parts and Downgrading*, Working Paper, IBM Research Division, Yorktown Heights, NY (September 2002), to appear in Naval Research Logistics Research Letters.
13. K. Katircioglu, B. Ramachandran, and G. Gallego, *An Inventory Optimization and A Decision Support Tool for Semiconductor Manufacturing*, Working paper, IBM Research Division, Yorktown Heights, NY (August 2002).

14. E. A. Silver and R. Peterson, *Decision Systems for Inventory Management and Production Planning, Second Edition*, John Wiley and Sons, New York (1985).
15. *Adaptive e-Business: Enabling the On-Demand Enterprise,* Publication G510-3644-00, IBM Corporation (2003). May be obtained by contacting the first author.
16. *Stage of Maturity Correlates with Performance, Profitability, and Sales Growth*, Supply Chain Council Industry Research Study, http://www.supply-chain.org/public/researchstudies.asp.

Accepted for publication August 26, 2004.
Internet publication January 7, 2005.

Shubir Kapoor
IBM Thomas J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, New York 10598 (shubirk@us.ibm.com). Mr. Kapoor is an advisory engineer in the Analytic Models and Architecture Department, which researches and develops techniques and systems for managing enterprise supply chains. He received his master's degree in computer science in 1997 from Pune University, India, and joined IBM Research in 2000. Mr. Kapoor's seven years experience in the IT industry includes the architecture, design, and implementation of business systems, using analytic techniques for modeling businesses, and developing rule-based systems. His technical interests include service-oriented architectures, server-side enterprise applications, rule-based expert systems, Web-based application development, and diagnostic systems for use in supply chain and e-business applications.

Kamal Bhattacharya
IBM Thomas J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, New York 10598 (kamalb@us.ibm.com). Dr. Bhattacharya, a member of the Business Informatics Department, focuses on the application of model-driven architecture concepts to real-world business scenarios, adaptive enterprises, and business performance management. Prior to joining IBM Research in 2001, Dr. Bhattacharya served as an e-business IT Architect for IBM Global Services in Germany and worked on several large scale e-business projects in the automotive and travel industry. He received a doctoral degree in theoretical and computational physics from Georg-August University, Goettingen, Germany, in 1999.

Steve Buckley
IBM Research Division, Thomas J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, New York 10598 (sbuckley@us.ibm.com). Dr. Buckley, a research staff member since 1987 and a manager since 1995, is currently manager of the Analytic Models and Architecture Department, Mathematical Sciences. His current activities are focused on sense-and-respond systems, an area in which his team has implemented prototypes for several IBM lines of business. Dr. Buckley received his Ph.D. degree in computer science from Massachusetts Institute of Technology (MIT) in 1987. He also received an M.S. degree in computer science from Pennsylvania State University in 1978, and a B.S. degree in applied mathematics and computer science from Florida State University in 1977.

Pawan Chowdhary
IBM Thomas J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, New York 10598 (chowdhar@us.ibm.com). Mr. Chowdhary has eight years of experience in the architecture, design, and implementation of high-performance distributed applications using object-oriented methods. As a member of IBM Global Services, which he joined six years ago, he led a technical team to winning and implementing a project for American Express. He also made contributions to IBM alphaWorks®, the IBM channel for delivering emerging technologies to early adopters. For the next two years he worked as an independent consultant advising clients on developing insurance applications. Two years ago he joined the IBM Research Division as an advisory software engineer in the Analytic Models and Architecture Department working on sense-and-respond systems. He received his bachelor's degree in electronics engineering from Nagpur University, India, in 1996.

Markus Ettl
IBM Thomas J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, New York 10598 (msettl@us.ibm.com). Dr. Ettl is a research staff member in the Analytic Models and Architecture Department, which researches and develops techniques and systems for managing enterprise supply chains. After receiving his doctoral degree in computer science from Friedrich-Alexander University in Erlangen, Germany, in 1995, Dr. Ettl joined the IBM Research Division and worked on the IBM Supply Chain Analyzer, a consulting tool that was used in the analysis of enterprise supply chains. He contributed to the work on extended-enterprise supply chain management, which was awarded the INFORMS Franz Edelman Award in 1999. Other awards and commendations for inventions and technical accomplishments include an IBM Outstanding Technical Achievement Award and several IBM Research Division awards.

Kaan Katircioglu
IBM Thomas J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, New York 10598 (kaan@us.ibm.com). Dr. Katircioglu is one of the leading researchers and a consultant in supply-chain management at the IBM Corporation. With more than ten years of experience in the field of operations research, management science, and logistics, his expertise covers the areas of inventory optimization, distribution and manufacturing operations, e-business, and supply chain management. He has a bachelor's degree in industrial engineering and a master's degree in statistics. He completed his Ph.D. in management science, transportation, and logistics at the University of British Columbia in 1996. Since then, he has worked on several projects for various divisions of IBM and its customers, published papers, and made several conference presentations. He is a member of INFORMS and IEEE.

Erik Mauch
IBM Personal Systems Group, 3039 Cornwallis Road, PO Box 12195, Building 002/Office MM304, Research Triangle Park, Raleigh, North Carolina 27709-2195 (emauch@us.ibm.com).
Erik Mauch is a Program Manager in the Integrated Supply Chain organization. His primary responsibility is leading the day-to-day execution of the Personal Computing Division (PCD) demand/supply conditioning activities and supporting other fulfillment transformation activities. He also supports the ongoing development and process improvement activities for the PCD and xSeries® demand conditioning processes. He joined IBM in 1999 and worked on improving existing and designing new supply-chain-distribution networks. Previous to his career with IBM, Mr. Mauch was a Principal Consultant in the Supply Chain Practice of PriceWaterhouseCoopers Consulting. He has a bachelor's degree in industrial engineering from Lehigh University and an M.B.A. degree from Carnegie Mellon University.

Larry Phillips
IBM Personal Systems Group, 3039 Cornwallis Road, P.O. Box 12195, Building 201/Office C104B, Research Triangle Park,

Raleigh, North Carolina 27709-2195 (plarry@us.ibm.com).
Mr. Phillips, a Senior Project Manager and IBM Certified Executive Project Manager in the Personal Computer Division, works in product planning and project management, and specifically manages a demand/supply conditioning project. He joined IBM in 1977 after receiving his master's degree in electrical engineering from the University of Tennessee at Knoxville. His experience in IBM includes Series/1 processor chip design, PC graphics and display development, market research, market development, and PC options and building block processes. For his significant contributions to the business he has received three Division Excellence awards. ■