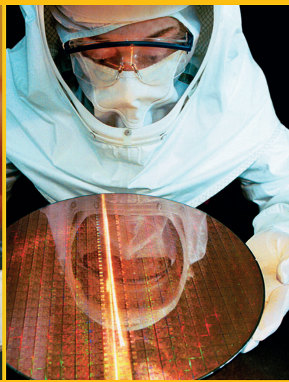
Volume 09 Issue 03 Published, August 3, 2005 ISSN 1535-864X DOI: 10.1535/itj.0903


Intel® Technology Journal

Managing International Supply and Demand at Intel

Modern-day Supply-Chain Management is the e-Commerce of manufacturing. Technology advances have enabled supply chains to provide greater flexibility in planning, sourcing, making, and delivering products globally with greater efficiency and lower costs. This issue of Intel Technology Journal (Volume 9, Issue 3) discusses how we are continuously improving the technology needed to manage product supply and demand in Intel's international network.

Inside you'll find the following articles:

Managing Uncertainty in Planning and Forecasting

Intel's Processes for Capacity Planning Optimization

Optimizing Supply-Chain Planning

Using Capacity Options to Better Enable Our Factory Ramps

Inventory Modeling

Redefining the Test Equipment Supply Chain: The Open Architecture Revolution

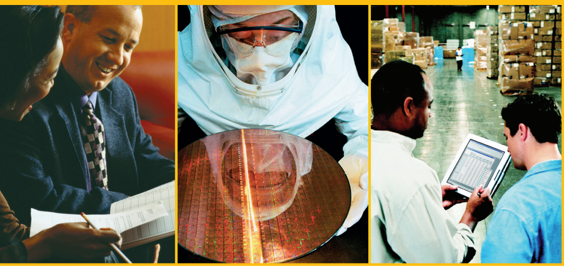
RosettaNet for Intel's Trading Entity Automation

e-Procurement–Strengthening the Indirect Supply Chain Through Technology Globalization

RFID: The Real and Integrated Story

More information, including current and past issues of Intel Technology Journal, can be found at:
http://developer.intel.com/technology/itj/index.htm

Volume 09 Issue 03 Published, August 3, 2005 ISSN 1535-864X DOI: 10.1535/itj.0903



Intel® Technology Journal

Managing International Supply and Demand at Intel

Articles

Preface	iii
Foreword	vii
Technical Reviewers	ix
Managing Uncertainty in Planning and Forecasting	175
Using Capacity Options to Better Enable Our Factory Ramps	185
Redefining the Test Equipment Supply Chain: The Open Architecture Revolution	193
e-Procurement–Strengthening the Indirect Supply Chain Through Technology Globalization	203
Intel's Processes for Capacity Planning Optimization	211
Optimizing Supply-Chain Planning	223
Inventory Modeling	233
RosettaNet for Intel's Trading Entity Automation	239
RFID: The Real and Integrated Story	247

THIS PAGE INTENTIONALLY LEFT BLANK

Preface

Managing International Supply and Demand at Intel Corporation
By Lin Chao
Publisher, *Intel Technology Journal*

From internal enterprise processes to external business transactions with suppliers, transporters, channels and end-users, Supply-Chain Management is the system of purchasing, producing, and delivering product to customers. Traditionally, Supply-Chain Management has meant factories, assembly lines, warehouses, transportation vehicles, and time sheets. Today's Supply-Chain management is a highly complex, multidimensional problem set with hundreds of thousands of variables for optimization. e-Commerce has changed the very foundations of manufacturing in virtually every industry. Modern-day Supply-Chain Management is the e-Commerce of manufacturing. An Internet-enabled supply chain may have just-in-time delivery, precise inventory visibility, and up-to-the-minute distribution-tracking capabilities. Technology advances have enabled supply chains to become strategic weapons that can help avoid disasters, lower costs, and make money.

Engineers at Intel Corporation have been refining Supply-Chain Management processes to provide greater flexibility in planning, sourcing, making, and moving products globally with greater efficiency and lower costs. The nine papers in this issue of Intel Technology Journal (Volume 9, Issue 3) describe in detail the science of how Intel manages product supply and demand globally. This ability, which transcends science, is the magic, art, and heart of Intel core competency.

Orchestrating product transitions across Intel is challenging yet vital. The first paper describes how Intel engineers address product transitions risk and uncertainty based on a planning approach consisting of three methods; the Product Transition Index (PTI), the Transition Playbook, and the Transition Dashboard. Based on case studies, the PTI is a structured and repeatable method for evaluating the state and impact of market, product, and marketing factors. The Playbook then helps the organization identify and determine how to respond to risks in a rapid and coordinated manner. The Dashboard guides navigation through the Playbook.

The next three papers look at procurement. The second paper is on capacity planning for expensive equipment with very long lead times and high costs. Using a model similar to options instruments used by financial markets, Intel purchases options for capital equipment. Options give Intel the right to purchase an equipment tool in a reduced lead time at a certain pre-determined price. These options were originally developed for lithography suppliers since lithography tool lead times are long while they are subject to multiple changes in product demand.

The third paper is on an open architecture standard called OPENSTAR* for semiconductor test equipment. The OPENSTAR architecture is a standardized infrastructure definition used for combining instrumentation from multiple suppliers into a common platform. The goal of this

* Other names and brands may be claimed as the property of others.

effort is to leverage standards at the instrument interface level (power, cooling, communication, and device interfacing).

The fourth paper looks at a plan to create a unified global procurement solution. The program, termed "e-Procurement," focuses on the global end state and targets three focus areas: tools, people, and processes. e-Procurement has a single global Enterprise Resource Planning (ERP) system as its foundation. An Internet negotiations tool was introduced to achieve additional cost savings and negotiation efficiencies.

The next two papers are on manufacturing. At each of its manufacturing plants and across the virtual factory, Intel is constantly adjusting product mix, manufacturing equipment (or tool) requirements, and overall business processes. This has dramatically impacted our ability to meet product demand and capacity utilization within both 200 mm and 300 mm Fab/Sort Manufacturing (FSM) and Assembly/Test Manufacturing (ATM). The fifth paper discusses optimization techniques that help automate different decision-making processes and provide common methodologies to collaborate and discuss optimal solutions. These models have saved Intel a great deal in capital cost over the past five years.

The sixth paper is on new automated data systems and optimization tools based on Linear Programming used to manage multiple divisions and stages of Intel's supply chain. These tools balance requirements to satisfy demand, achieve inventory targets, and remain within production capacity to reduce costs and satisfy demand across Intel's supply chain. They have been developed to evolve the planning process and facilitate continuous improvement while maintaining visibility to the logic and data flow. Planning time has decreased dramatically; supply costs have been reduced; and demand satisfaction has improved.

The next two papers are on logistics. The seventh paper describes a new way of optimizing Intel Corporation's supply chain, from factories to customers. The methodology uses statistical methods to characterize the order distributions of customers and the distribution of times to ship products from different points in the supply chain (factories to customers).

The eighth paper is on the RosettaNet[*] standard based on XML-based protocols to facilitate secure electronic exchange of trading entities over the Internet. Over the past five years, Intel has aggressively pursued utilizing RosettaNet to support its supply chain. The paper reviews the success Intel had in building new business processes using the e-Business infrastructure of RosettaNet. Also, the future of Business-to-Business (B2B) exchanges and the next generation of B2B architecture are discussed.

The final paper looks at using Radio Frequency Identity (RFID) technology in supply-chain operations. Intel's supply network organization formed a unique collaboration with a major Original Equipment Manufacturing (OEM) customer to run a proof-of-concept experiment utilizing RFID tags in the combined supply chain of Intel's Malaysian assembly/test facility and the OEM's Malaysian laptop assembly plant. The paper chronicles this path-finding project from inception to completion and shipping of over 70,000 CPUs to the OEM customer in a four-week period.

---

[*] Other names and brands may be claimed as the property of others.

In only a few years, the very fundamentals of manufacturing have changed dramatically. Modern-day Supply-Chain Management is the e-Commerce of manufacturing. Technology advances have enabled supply chains to provide greater flexibility in planning, sourcing, making, and delivering products globally with greater efficiency and lower costs. Intel is a star at today's supply-chain optimization and this ability is one of Intel's key strategic assets.

THIS PAGE INTENTIONALLY LEFT BLANK

Foreword

Managing Intel's International Network of Supply and Demand

By Karl Kempf
Intel Fellow and Director of Decision Technologies
Technology and Manufacturing Group

The term "supply-chain management" is often used in the industry to describe all activities involved in satisfying customer requests. As a company with revenues of $34.2 billion in 2004, these activities at Intel are extensive to say the least. We manufacture hundreds of different devices in a dozen sites spread around the world supplying the computing and communications industries with chips, boards, and systems that are the "ingredients" of laptop and desktop computers, servers, and networking and communications products. On the supplier-facing side of Intel, we deal with thousands of suppliers of goods and services ranging from simple raw materials to some of the most complex production equipment ever developed. On the customer-facing side, we interact with thousands of customers around the clock, around the world. At Intel, we manage supply and demand as equally important for our continued growth. In addition, given the number and geographic distribution of Intel's suppliers and customers, the relationships form a network that is much more complex than a chain. Through our research and development projects, we are continuously improving the technology needed to manage supply and demand in Intel's international network.

Supply activities in the network are typically operationalized as "*plan*, *source*, *make*, and *deliver*." *Plan* includes integrated planning across source, make, and deliver. While optimizing individual activities frequently improves local performance, such efforts seldom have impact across the entire supply/demand network. It is only through the development of computer information and decision support systems that span all activities that we help minimize costs and maximize revenues for the whole corporation.

Source means procuring goods and services while building win-win relationships and mitigating risk. A variety of technologies can be employed to realize these goals, including financial instruments such as contracts and options, web-enabled communication facilities, and industry-standards setting, to mention a few.

Make spans all facets of production and requires that we employ financially sound operating methods to be successful. Given the long lead time required for building new facilities or modifying existing ones, there are the strategic problems having to do with future capital expenditures. There are also tactical problems involving efficiently utilizing current facilities given long manufacturing lead time. In every case, the goal is minimizing cost while maximizing demand satisfaction.

Deliver encompasses getting goods and services to other businesses (Business-to-Business, or B2B) and end consumers (Business-to-Consumer, or B2C) in a timely and cost-efficient manner. Once again, a number of approaches are required to achieve these goals including positioning warehouses and sizing inventories, contractual arrangements with shipping firms, web-enabled communications, and others.

On the **Demand** side, we strive to continuously improve our technology to forecast and influence the desires of the market. This has historically included a variety of demand forecasting techniques. More recently, we have been placing increasing emphasis on the timing of new product introductions and special offers as well as price moves and other related marketing techniques.

As you study this issue of Intel Technology Journal (ITJ), notice the collaboration among various branches of the materials group, equipment selection and purchasing groups, those involved in factory automation, simulation and optimization experts, product groups, information technology, and, of course, planning and logistics groups. Notice also that inclusion of our external suppliers and customers is a critical component of many of our technical achievements.

Similar to the tip of the proverbial iceberg, this issue of ITJ exposes only a small but important portion of our efforts to continuously improve the performance of Intel’s international supply/demand network to maximize value for our shareholders, satisfaction of our customers, and efficiency of our employees. Through this directed innovation we will continue to deliver Intel’s world-class products with world-class speed, agility, and cost effectiveness in supply/demand network performance.

Technical Reviewers

THIS PAGE INTENTIONALLY LEFT BLANK

Managing Uncertainty in Planning and Forecasting

Jay W. Hopman, Information Services and Technology Group, Intel Corporation

Index words: planning, forecasting, product transition, product introduction

ABSTRACT

Orchestrating product transitions is a challenging yet vital function across Intel's organization. "Demand Generation" research, sponsored by Intel's Customer Fulfillment, Planning, and Logistics Group, has probed historical and present-day transitions to learn how the organization plans, forecasts, and executes transitions and how we might be able to improve in the future. Completing case studies of real Intel® products led to the development of models to comprehend observed dynamics and to the development of new methods to address common challenges across the phases of product lifecycle management.

The paper covers key findings and outcomes from the research, describing observed problems and detailing solutions that have been identified or developed from 2002 to 2005. The solutions focus specifically on the functions of planning and forecasting, highlighting the need to integrate a broad base of information into a stack that includes not only hard data but also strategies, assessments, uncertainties, risks, and contingency plans. The overall solution stack is comprised of a planning system linking business strategies and assessments to a playbook of risks and contingency strategies. Use of these methods helps the organization plan for uncertainty and improves agility by mapping out tactics in response to potential risks. Forecasts also benefit from the use of repeatable, systematic methods and the integration of uncertainty.

Initial findings from pilots with two business groups validate the approaches, but integrating new methods into broadly used tools and processes is not without challenges. Research continues on both the proliferation of these methods and an exciting new capability, the use of market mechanisms to resolve ongoing challenges associated with traditional hierarchical planning and forecasting. Potential applications and advantages of market-based systems are discussed.

INTRODUCTION

Intel's business is one of transitions. The steady stream of technological innovation driven by Moore's Law requires one product and manufacturing process transition after another, each bringing a new generation of capability and computing power to the market. While transitions are ultimately beneficial, delivering value to consumers and shareholders, they also introduce uncertainty and risk to Intel's product management across all demand and supply functions.

Our research into planning and forecasting through periods of product transition was spurred by specific cases where transitions did not turn out as well as they might have. We set out to study the transition management process from a systems perspective. Each team involved in phasing in one generation of product and phasing out another uses processes (in the form of policies, strategies, or models) to manage data (input and output) and interfaces with other functional teams, each driven by various indicators of operational and strategic success. Using hard and soft data we studied these aspects for several products, seeking to identify sources of uncertainty and their adverse impacts on the bottom line. We then developed concepts and methods that would help the business navigate transitions with greater success.

An early lesson of the research was that in high-volume markets the stakes of a product transition are high, and numerous factors can compromise transition success. A substantial miscalculation of the timing of market demand or a technological glitch impairing supply can cost the company $500M in a market segment worth over $10B in annual revenue.

Through the case studies we classified four sources of uncertainty: *market*, which includes economic, business, and seasonal cycles; *product changes*, those of Intel, competitors, and complementors; *marketing actions*, which include pricing, promotions, and advertising; and *systems*, the methods by which we forecast, plan, execute, measure, and monitor our business.

We developed two methods, each intended to be a part of the *systems* used to manage product transitions, to help assess and account for the uncertainties stemming from market, product, and marketing factors. The first method, the Product Transition Index, divides a transition into eight vectors and calculates the energy in each vector, driving the transition's pace and ultimate success. The second is the application of an idea that came out of research at Stanford and Hewlett Packard, i.e., Transition Playbooks. These are designed to coordinate organizational response to risks. Other methods we have considered in our research include range forecasting and the use of market mechanisms. In this paper, we touch on each of these methods, describing our experiences with them to date and considering how the collection of ideas addresses the overall challenge of transition management.

SOURCES OF UNCERTAINTY IN TRANSITION MANAGEMENT

Case studies of past product transitions at Intel and other companies have revealed many failure modes [1]. Among the more common are the following:

- Weaker than expected demand for a new product, causing a surplus of new product and a shortage of old product.
- Stronger than expected demand for a new product, causing a shortage of new product and a surplus of old product.
- Delayed supply for a new product, causing a shortage of old product and potentially depleting inventories of any product to sell.
- Weaker demand for the old product in anticipation of the new product, causing a surplus of old product. If the new product is then delayed, it is known as the "Osborne Effect" and can drive a company out of business for lack of product to sell [2].

Three factors are at work in these scenarios: market risks, which are the demand for old and new product; technology risks, which are the supply of the old and new product; and the alignment of demand and supply. Although weak demand or supply can be damaging, we have observed that a large gap between demand and supply tends to be the most damaging outcome.

Another way to slice the uncertainties in transition planning is a four-layer model consisting of market, product changes, marketing actions, and systems. A description of each layer and key findings from our case studies follows.

Market

The best understood and most documented uncertainties in forecasting are market forces. For that reason this layer is the least interesting in the model from a research perspective. Planning supply hinges on demand, which hinges on many market forces from macroeconomic cycles to industry specific cycles to seasonality. Since these forces are essentially outside of a company's control, the best response is modeling ranges of likely results and comprehending potential outcomes in the planning processes. The solutions we present later comprehend market uncertainty but demonstrate that it is only one of many drivers in transition planning.

Product Changes

Whenever a product changes within a given market, uncertainty results. Sometimes the product is the company's own. Other times it belongs to competitors or complementors. Product roadmaps and tactical plans should be mindful of competitive forces, and the impact of complementary products on a company's own are a key element of planning. However, the most dominant factor in transitions–one that is completely under the company's own control–is changes made to one's own products. The technology and feature gap between two generations of products and across market segments (e.g., high end, mainstream, and value products) is fundamental to sales. Numerous examples within and outside Intel demonstrate that product sales can soar or plummet due solely to competition among a company's own products.

Marketing Actions

Products cannot be brought to market without consideration of pricing and promotion, but interestingly the impact of these policies on demand is not deterministic. Due to the many factors that drive demand, predicting the exact result of a price move or an advertising campaign is improbable at best. Still, product sales can be modeled most simply as a function of capability and price, with the ratio of the two determining customer value. Our critical finding is that price is both a powerful and overused lever. We therefore looked to solutions that encourage the use of other product and marketing levers and considered the lasting repercussions of the levers that are used to manage a transition.

Systems

The most interesting source of uncertainty in transition planning turned out to be the very processes and tools used to manage the business. Forecasts proliferating through our sales and marketing and business planning functions are judged four or more times between customers and the supply network. Each layer of judgment hinges on local knowledge, local policies, and local indicators and incentives. We observed that the propagation of these datasets at best involves lag, judgment, and some loss of context–the strategy and uncertainty behind the data–and at worst may include clear bias and gaming. To clarify, any given team routinely judges forecasts up or down, always based on experience and available information, because the incoming information is deemed too high or low, or because the outgoing forecasts are expected to be judged up or down by a subsequent owner.

Another key finding is that the forecasting systems tend to be noisy. One fundamental cause seems to be the use of point estimates for sales by product SKU, family, or manufacturing start. We observed that a series of updated point estimates conveys uncertainty through the volatility of the signal over time. In this method quantifying the uncertainty requires tracking information over time, which is too broad a view for the busy planner. Instead, planners tend to chase the dominant data point in the forecast, something we call "change from prior." The critical information in an updated forecast is not the actual forecast (typically unit sales) but rather the delta or "change from prior" since the last forecast. The phenomenon of each new forecast or supply network plan reacting to the change from prior propagates noise through the system. Instead of ignoring insignificant volatility, planners often transmit it.

Looking across these sources of uncertainty led to the discovery of an additional source, actions implemented by the organization (product, marketing, or supply changes) intended to manage the transition. We found that actions taken in different groups across the organization were not always planned and executed in synch, so the net impact of these actions sometimes manifested itself in the form of unexpected results.

As our research team entered the solution space we considered methods, some developed internally, some discovered or recommended to us along the way, that we believed would help manage or even reduce the uncertainties affecting our product transitions.

SOLUTIONS TO AID TRANSITION PLANNING AND FORECASTING

The following principles were developed out of the case study work to guide our general approach to improving planning and forecasting.

- Global strategies should drive local actions that support global optimization. Local policies and incentives should be flexible, shifting with global strategy.
- Forecasts should convey a contextual layer above the numbers. Each forecast handoff subjects data to loss of context and a new round of judgment. Context–strategy and uncertainty–should be communicated across internal and external interfaces.
- Processes should be designed to identify and attenuate noise. "Over-nervous" planning reacts too strongly to short-term trends and aggressively closes gaps, sometimes leading to oscillation and amplification.
- Organizational processes should systematically manage uncertainty. Contingency planning, scenario planning, and range forecasting improve positioning and reaction speed.
- Market assessment and response (strategic and tactical) should be as systematic and repeatable as possible, codifying tribal knowledge and enabling new types of analysis. It should capture the past and present sufficiently well to help predict and manage the future.

Product Transition Index

Based on these principles, we developed a planning approach consisting of three methods, each used to encourage collaboration and coordination among functional teams across the organization. The first, Product Transition Index (PTI), is an assessment tool used to gather information about the product transition. PTI is a model containing eight vectors that dictate the pace and success of a transition. A total of 65 factors identified in our research are scored to complete the PTI, and the scoring process requires integrating the knowledge of teams across sales, marketing, planning, manufacturing, and engineering. Table 1 lists the vectors in PTI and provides a brief summary of key factors within each vector.

Table 1: PTI vectors and summary of key factors

PTI Vector	Summary of Factors
Product Capability	performance, features, usability, compatibility; anticipated product longevity, quality, reliability
Product/ Platform Pricing	cost of the product itself, the platform that uses the product, the process of adopting, manufacturing, or integrating the product; historical and expected price stability; costs of competitive products
Timing	time since last product introduction, anticipated time to next product introductions, age of the installed base, timing of competitive introductions
Marketing Indicators	product alignment to market segments, breadth of product applications, potential market size, timing and aggressiveness of promotion, end customer impression of product
Environment	economic conditions, customer demand trends; health of own company, value chain partners, competitors
Competition	performance, features, market perception of competing products, competitors' manufacturing capability and capacity, and alignment between competitors and value chain
Value Chain Alignment	cost and complexity facing value chain, reliance on new standards/technology, reliance on suppliers to deliver, perception of product attractiveness, balance of customer pull versus own push
Internal Execution	manufacturing risks such as design readiness, capacity, process health; clearance of regulatory hurdles, sourcing risk for materials

A scored PTI shows the relative energy imparted to the transition by each vector. Scores range from cold to hot with the center of the range aligned to the typical past transition in that product family. If all vectors are scored down the middle, the product transition should be expected to unfold at a rate on par with the average of past transitions. Hotter scores predict a faster transition, colder scores a slower transition. A faster or slower transition is not necessarily better or worse. Rather, the PTI should be assessed for the overall balance of demand and supply for the old and new generations of product. A scenario of slow demand and slow supply is easier to manage than one of fast demand and slow supply or vice versa. The scoring process should therefore be used to identify risk factors in demand, supply, and demand-supply alignment that could derail the transition.

Transition Playbook

The second method in our planning solution is Transition Playbook, an idea developed in research at Stanford and Hewlett Packard [3]. The intent of the playbook is to enable strategists and managers to map out the tactics the organization will use to respond to risks that might impact the transition. Sports teams develop plays so that in the stress and time constraints of a game tactics can be invoked without delay and the team players can perform with nearly perfect synchronization. A playbook in business likewise encourages advance planning and analysis so that the business functions can respond quickly and in concert to keep the transition on track.

Playbooks (see Figure 1) consist of a primary transition strategy, transition risks, and contingency transition strategies. The primary strategy is formulated based on the output of the PTI process and the market strategies for the product. We observed that market strategies for new products commonly have a blend of three objectives: profit, market segment share (unit sales), and market or technological leadership. The most critical step in developing the playbook is to understand the weighting of these (and perhaps other) objectives and to understand which results will constitute a successful transition. The PTI results shed light on how readily these objectives will be achieved and identify the risk factors that could prevent a successful transition. The primary strategy implicitly includes both tactics that lead to success and preventive strategies aimed at avoiding the more threatening risks.

Transition risks generally fall into four categories: new product demand greater than or less than supply and old product demand greater than or less than supply. We have found that the mapping of risks identified in the PTI scoring process to these categories is straightforward. It is impractical to account for all risks in a transition, so risks must be prioritized. We relied on a standard Intel definition of risk (probability x severity) to guide prioritization.

Contingency strategies are comprised of both preventive and mitigation strategies. If a risk is seen on the horizon as the transition unfolds, it can potentially be circumvented. But if the risk is already imminent, then the remaining option is to minimize its impact. Some contingency strategies are specific to risks while others are targeted at category of risk. So, the risk of a certain technical glitch likely requires a direct response to the glitch. The risk of demand exceeding expectations for a generation of product could trigger any number of tactics designed to speed product delivery or perhaps shift demand to another product (ideally not the competitor's).

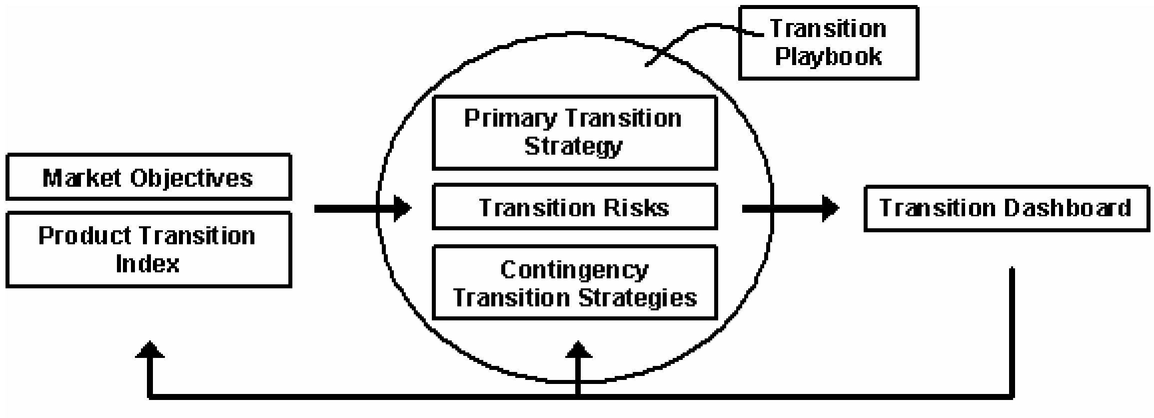


Figure 1:Transition playbook planning system

Defining primary and contingency strategies involves choosing from many potential tactics. As the business develops strategies it is helpful to refer back to the PTI model for guidance. While two of the PTI vectors, environment and competition, are largely outside of the company's control, the remaining vectors contain a number of levers that can be used to influence the transition. Table 2 lists a matrix of control by vector that guides application of levers.

Table 2: Transition control by PTI vector

PTI Vector	Complete	Long Term	Indirect	None
Product Capability	✓	✓		
Product/Platform Pricing	✓			
Introduction Timing	✓	✓		
Marketing Indicators		✓	✓	
Environment			✓	✓
Competition			✓	✓
Value Chain Alignment		✓	✓	
Internal Execution	✓	✓		

Transition Dashboard

The third method in our solution is the Transition Dashboard, which is intended to monitor the risks identified in the playbook. The dashboard tracks the key risks to demand, supply, and demand-supply alignment for both the new and old products and should be used to trigger execution of the playbook. An ideal dashboard is tied to execution of increasingly aggressive tactics within the contingency strategies as the transition moves farther off-track. It also indicates when the existing playbook is no longer able to satisfy the success criteria for the transition. In such a case, it may be necessary to revisit the market objectives or the PTI assessment for the product and then revise the playbook.

One of the clearest benefits of a transition playbook is the definition of measured responses. We observed that Intel employees are highly adept at closing gaps, so adept that in transition management gaps are sometimes closed so far as to open an inverse gap. Within a fairly short span of time inventory shortages can turn into surpluses and vice versa. Using a dashboard with well-defined triggers to invoke appropriate and measured contingency strategies can help keep the supply network out of a state of oscillation or even bullwhip.

Integration with Other Solutions

The research team developed a new approach to forecasting by integrating the PTI scoring system with diffusion modeling. Our research partner, Paulo Goncalves at the University of Miami, developed system dynamics models based on the equations from epidemiology applied to product and technology diffusion by Frank Bass [4]. The first step of this method involves fitting past transitions in similar product families to the diffusion models. The set of parameters (market size, coefficient of innovation, and coefficient of imitation) from past products provides a range of likely parameters for the new product. A range of potential diffusion curves can be calculated directly using the system dynamics model. However, each product transition is unique, having characteristics that make it behave differently from past product transitions. We capture these differences using the results of the PTI assessment for the current product (which implicitly compares the current product to past products). The scores from PTI are used in the model to calculate the attractiveness of the new product relative to the company's old product and to competitive products. The attractiveness is then used to modulate the diffusion curves, and the sensitivities of unit sales to various factors in the model can be tested using Monte Carlo techniques. While this technique holds promise, the lack of readily available price data for historical products has been an obstacle preventing precise calculation of attractiveness as a function of price. Nonetheless, the concept holds promise for future application.

The playbook method is a natural bridge to range forecasting. It was noted earlier that Intel's traditional forecasting systems are based on point estimates (from a statistical perspective the estimates are most analogous to expected value). Playbooks can be thought of not only as planning maps but also as decision trees. With each risk comes a probability of occurrence and a range of potential impact on sales. Similarly, each contingency strategy has a probability of being invoked and an expected range of efficacy. A playbook analysis might reveal that a transition faces an aggregate risk of 60-80% that new product sales will come in below expectations, and a 45-65% chance that supply will come in below expectations. Numbers can also be expressed in pure units. So, the playbook might reveal a 50% probability that Q2 sales of the new product will fall within 6.0m and 8.5m units. Such range forecasts can be used broadly, and a team of demand and supply planners at Intel has begun integrating range planning into some of our systems in the past year.

The benefits of range planning are threefold. First, the organizational mindset is moved from artificial certainty to uncertainty. Rather than building to hit an expected value outcome and then chasing that outcome as it changes, the supply network can build to cover a range of outcomes. The focus of planning shifts from guessing and optimizing the expected outcome to analyzing financial and operational performance across a range of outcomes. The final decision on which parts of the range of outcomes to cover becomes a largely strategic decision based on the results of these analyses. Second, the amount of noise in the forecasts is reduced because the range forecasts can absorb some degree of volatility period to period without adjustment. Less energy is devoted to processing noise. Third, range forecasting encourages portfolio management of capacity and materials. The blend of fixed (lowest cost, dedicated use), fungible (higher cost, use across product lines), and flexible (higher cost, shorter lead time) capacity and materials helps to cover different outcomes with varying degrees of cost and risk.

Another area we have researched for the past few years has involved the use of market mechanisms as substitutes or complements to traditional hierarchical forecasting systems. As PTI and playbook are intended to aggregate and coordinate information from across the organization, market mechanisms may also be used to aggregate knowledge and provide better insight on demand trends. We are preparing to launch a series of market experiments to assist planning for product families that have proved challenging for our traditional forecasting processes. The market forecasts will be evaluated based on their accuracy, volatility, and the speed with which they react to market significant events. Much of our learning in the area of markets has come from the University of Iowa Electronic Markets [5] and the forecasting experiments performed at Hewlett Packard in conjunction with the California Institute of Technology [6].

RESULTS OF INITIAL PILOTS

Application of PTI began in 2004, a few months prior to the release of a new generation of product. Intel's central marketing and planning organization, the team most directly responsible for managing demand and supply alignment, used PTI as a process through which to collect information about the new product and the transition. We organized sessions with several teams from our sales and product marketing organizations, having each team score and provide comments on the factors for which team members had information.

The assessment process revealed several interesting insights. First and foremost, the two marketing teams representing key components of the new platform each felt that the other team's component would be the one to drive sales of the platform. We interpreted that as a bad sign because each team felt that their own product would not be the main driver. Second, the prevailing wisdom expressed both outside and even inside these sessions about product strengths and weaknesses did not match up to factor by factor analysis within PTI. A few areas that were widely considered strengths could not be justified as strengths based on hard data. Third, sales representatives alleviated fears that technical issues or manufacturing challenges might slow adoption of the product, but they had insight that the overall cost of the new product platform might impede sales within certain market channels.

The resulting PTI scores showed the vectors driving the speed of the transition to be environment (hot economic and recent sales trends), internal execution (product ready for moderately fast ramp), and marketing indicators (solid alignment to some market segments). Product capability and competitive factors were also somewhat positive, while timing was neutral. Vectors inhibiting the rate of transition included value chain alignment (typically strong support from some customers but rather weak support from others) and, to a lesser degree, price (platform cost).

Based on the PTI assessment and a comparison to actual sales results from a product released the year prior in the same family, we determined that the consensus forecast was optimistic. If we define the best whisper forecast among central planners for sales over the next two quarters as x, the official forecast being published and used to drive supply was about 1.2x. The estimates coming in from the sales organization were fairly volatile from month to month but ranged from 0.65x to 0.9x. Based on the sales organization's past forecasting patterns, the central planning group felt that these figures were pessimistic. After completing the PTI assessment, we published a report about six weeks prior to launch stating that sales were unlikely to exceed 0.93x and would probably come in lower. Given all available information, we stated that only an improbably large second quarter after release could result in a higher sales total. Within about six weeks after launch the official forecast dropped to approximately 0.9x and continued to decline. By the beginning of the second quarter after launch the forecast accurately called the final result of 0.79x.

In hindsight, the PTI assessment enabled the pilot team to identify the strongest drivers and inhibitors of the transition. Considering all factors affecting the transition and comparing it to a recent transition in the same product family we were able to generate a prediction that was accurate enough to benefit the bottom line through better allocation of factory capacity and sound inventory planning. The participants in the process from the central planning team felt that in comparison to past transitions the PTI process brought better insight and enabled better forecasting. As work on this transition began to slow, the team promptly began discussing application of the method for the next major transition.

In 2005, we began applying the playbook method with a different Intel business unit. The senior management team of this unit requested an assessment of the product and technology roadmap against the direction of the overall market and the strategies being employed by competitors. To tackle this problem we combined the playbook approach with a scenario planning process that has been applied at Intel for the past five years. Scenario planning considers long-term business strategies and product roadmaps against potential future states of the market. Representatives from across functional teams work together to envision potential future market states, which are then used to script possible story lines for Intel's businesses. We felt that combining this approach with the playbook approach would bring a complete vision of how the entire business fits together. A product roadmap is a series of transitions, and analyzing each transition as a standalone event and as part of a five-year business plan seemed a sensible approach. The scenario planning piece helped define market objectives, primary strategies, and risks for individual transition playbooks. In return, the playbook enabled more actionable output from the scenario planning process. Indicators of important market shifts can be included in the dashboard and used to trigger contingency plans within the playbook.

The output of the scenario planning process emphasized the importance of the upcoming product transition in the greater context of the business. Everyone left the room with a clearer definition of success for the overall business and for the product transition. A playbook for the upcoming transition is now in development. We have analysis covering eight dimensions in the playbook, including the impacts of these factors: qualifying various SKUs for production and sale, design wins with various customers, timing of product launch, and manufacturing process health. The best and worst potential outcomes (in unit sales for the new product) have a ratio of 4:1, which at face value makes for difficult supply network planning. But, within that range the business now understands the influence critical drivers will have on demand and supply and can begin pulling levers months to quarters ahead of product launch to drive a successful transition and keep demand and supply aligned.

CHALLENGES

The greatest challenge to developing and proliferating new planning and forecasting methods is getting the methods piloted in an operating and bandwidth constrained organization. Everyone involved in operational planning has a full workload and is already using a suite of applications and processes to do their job. We encourage grass roots participation and work our way to organization-level pilots by starting within the organization and working up to senior management. Our partners in business groups take the methods to their own managers as potential solutions to recognized problems, encouraging employees to participate and fit R&D into their otherwise operationally focused schedule. Finding organizations willing to partner on a pilot takes time and quite a bit of selling, but an initial success in one pilot starts to open other doors.

As we near the piloting of market solutions we face more specific challenges. A participant base of at least 20-30 individuals is required for good results. Recruiting these participants, who will be expected to remain involved for more than six months, requires a blend of marketing and incentives. We will then need to demonstrate the exact benefits of their participation to the company in order to retain interest.

Another challenge is identifying suitable metrics for testing solutions. Obvious choices include operational metrics such as inventory levels and return on invested capital, but it is difficult to isolate the effects of the new methods among all the other factors in the environment. We are also looking at forecast signals to see improvements in accuracy, volatility, and timeliness (response to events). Direct feedback from partners and participants is highly valuable. If they state that using the process brought higher confidence, enabled better judgment, enabled anticipation of risks and responses, or reduced the workload to produce a forecast, then a clear benefit has been achieved even if it is not purely quantitative. The ultimate indicator of benefit is whether we achieve successful transitions and whether we keep transitions on track using the methods, and we will certainly be tracking that indicator through all future activities.

A final challenge is porting new concepts, methods, and processes to next-generation tools. Research in information technology has a limitation in that product and application roadmaps are largely vendor driven. The choice of building home-grown solutions or sticking to vendor roadmaps always exists. Our current approach is to build simple tools for the purpose of piloting new methods while leading our business partners to develop new requirements for vendor-developed tools based on their experience with the pilots. Pulling vendors directly into research is another option, but the pros and cons are many. In some cases we will likely choose to engage vendors directly, but thus far it seems even layering simple tools above our more robust operational systems can yield good results.

SUMMARY

Planning and forecasting have become exercises of data sets and spreadsheets. The numbers themselves, judged and translated three or four times between customer and supply network, lose the business context of strategy and uncertainty. Along with hard data, the entire chain of customer fulfillment, from sales to marketing to planning to distribution, needs to grasp this context in order to manage the transition to the right global indicators and results.

Revenue, profit, and market position are optimized only when the right products can be sold at the right time at the right price. The uncertainties posed by markets, product changes, marketing actions, and the systems used to manage the business make this outcome largely unattainable.

Based on our case studies we developed PTI as a structured and repeatable method for evaluating the state and impact of market, product, and marketing factors. PTI helps aggregate, document, and communicate information from around the organization. Playbook then helps the organization identify and determine how to respond to risks in a rapid and coordinated manner, with the dashboard guiding navigation through the playbook.

Range forecasting complements statistical and Monte Carlo methods and produces more stable forecast signals with embedded uncertainty. Playbooks can be used to produce range forecasts, as can market mechanisms. A range forecast encourages more intelligent and strategic positioning of capacity, materials, and inventory and discourages chasing best guesses of demand.

Market mechanisms speed the transmission of demand signals and more often than not beat the accuracy of traditional forecasting systems. We plan to test market-based systems and compare the accuracy, volatility, and timeliness of their output to our standard forecasting systems.

Each of these solutions has been or will soon be piloted within Intel as our many business units seek to manage uncertainty more effectively. The results of our pilots with PTI and playbook have been encouraging, and the application of market and range forecasting methods outside Intel (and within Intel to the limited extent we have tried them) has shown considerable promise. In combination the methods form an arsenal of tools to drive a more profitable business and a better positioned and strategically and financially more valuable supply network.

CONCLUSION

The strong focus of this work on transitions begs the question of planning and forecasting in the steady state. In reality, steady state does not exist very long in high-tech industries. Shorter product lifecycles have resulted in rather dynamic markets; managing product lifecycles is now less relevant than managing transitions from peak to peak. A product's ramp up is followed by a ramp down, and the perspective of balancing the ramp down of each generation with the ramp up of the next focuses organizational energy toward the dynamic and uncertain reality of the transition.

During our research we encountered several methodologies that use mathematical models and historical data to forecast transitions and optimize supply. While these are sound approaches they fall short if they are blind to the factors in PTI. Managing supply without regard to the particulars of demand is optimizing the wrong problem, sweating the "ones" digit while hoping the "tens" digit comes in as expected.

Similarly, planners immersed in a world of spreadsheets and point estimates may not have insight into how a transition is unfolding or how the company will react. Without the formal mechanism of a playbook to convey the risks on the horizon and who will take what action to counter them, tribal knowledge, hallway conversations, and other informal networks are used to convey context and guide policy. Managers may work together to develop strategy across the organization, but unless mechanisms are in place to coordinate the execution of those strategies and to drive strategy into local policies and decisions, the organization is not achieving its potential level of synchronization. Individual planners need not be able to articulate the complete management strategy for the new product, but they should certainly know what to expect next and which actions they themselves should take if the transition starts to go off track.

One capacity planner told us that at the end of the day models developed within the supply network are as accurate and only a fraction as volatile as the signals from the demand side of the organization. In other words, ignoring the demand signal until the time to build product draws near works just as well. If demand information is to be used to advantage, the supply network must perceive it to be a credible source of information. The PTI and playbook processes, in combination with range planning and market mechanisms, provide opportunities to make demand forecasting more structured, stable, honest, repeatable, and timely. The playbook also offers marketing, planning, and manufacturing teams a means to more effective coordination through advance planning.

Intel's Customer Fulfillment, Planning, and Logistics organization has articulated an objective of shortening the distance between the customer and the supply network. The methods described in this paper are among the options available to do exactly that.

ACKNOWLEDGMENTS

David Tennenhouse and Mary Murphy-Hoye for the insight to create new research in this area.

Adam J. King and Carole Edwards for strong support of case study activities and for taking advantage of the findings.

Mary Murphy-Hoye (again), Jim Kellso, Dan McKeon, Keith Reese, John Vicente, and Doug Busch for consultation and for building connections that enabled this research to grow in scope and influence.

David McCloskey, Troy Clark, Patricia Nealon, and Alan Young, among many others, for sponsoring and supporting pilot activities around Intel.

Hau Lee, Feryal Erhun, Blake Johnson, Paresh Rajwat, and Xiaoshu Shao (Stanford University), Paulo Goncalves (University of Miami), Jim Hines and Jim Rice (MIT) for your creativity and many hours invested in this research.

REFERENCES

[1] Erhun, F., Hopman, J., Lee, H.L., Murphy-Hoye, M., and Rajwat, P., "Intel Corporation: Product Transitions and Demand Generation," *Stanford Global Supply Chain Management Forum*, Case No: GS43, 2005.

[2] Osborne, A. and Dvorak, J.C., *Hypergrowth: The Rise and Fall of Osborne Computer Corporation*, Avon Books, 1985.

[3] Billington, C., Lee, H.L., and Tang, C.S., "Successful Strategies for Product Rollovers," *Sloan Management Review*, 39(3), Spring 1998, pp. 23-30.

[4] Bass, F.M., "A New Product Growth Model for Consumer Durables," *Management Science*, 15, 1969, pp. 215-227.

[5] Berg, J. E. and Rietz, T. A., "Prediction Markets as Decision Support Systems," *Information Systems Frontiers*, 5(1), 2003, pp. 79-93.

[6] Chen, K.-Y. and Plott, C. R., *Prediction Markets and Information Aggregation Mechanism: Experiments and Application*, California Institute of Technology, 1998.

AUTHOR'S BIOGRAPHY

Jay Hopman has been with Intel IT since 1993 and IT Research since 2001. Jay graduated from Purdue University in 1992 with a B.S. degree in Computer and Electrical Engineering and completed an MBA degree (concentration in Strategic Analysis) at the University of California, Davis in 2000. His focus areas in research and development at Intel have included distributed systems performance, economics of IT investment, applications of market mechanisms, and planning and forecasting systems. His e-mail is jay.hopman at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

Using Capacity Options to Better Enable Our Factory Ramps

Viswanath Vaidyanathan, Technology and Manufacturing Group, Intel Corporation
Dave Metcalf, Technology and Manufacturing Group, Intel Corporation
Douglas Martin, Technology and Manufacturing Group, Intel Corporation

Index words: capacity, options, forecasting

ABSTRACT

Supply-chain management at Intel Corporation is more correctly defined as managing a supply demand network, since supply and demand are treated as equally important in a complex network. Cyclical industry trends, steep ramp curves, and small changes in the electronics industry can drive significant changes to individual semiconductor equipment suppliers as can be seen in Figure 1.

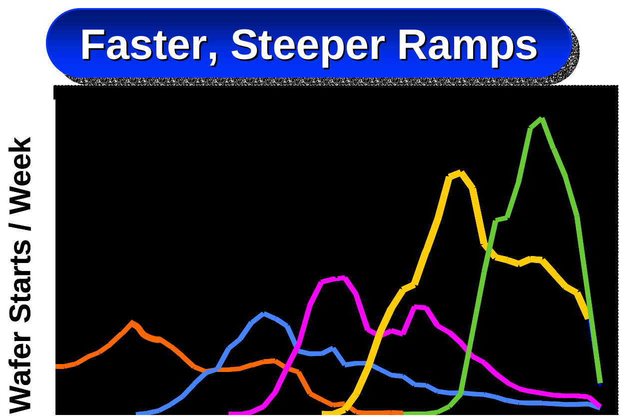


Figure 1: Typical semiconductor manufacturing ramps

Capacity planning within the Intel supply demand network is a complex process. Demand forecasts for Intel capital equipment tool sets, especially in lithography, are complicated by the extremely long lead times, expensive tool costs, and high contractual cancellation fees. In addition, changes in quarterly Intel manufacturing factory roadmaps cause considerable changes in the lithography exposure tool requirements. Lithography exposure tool requirements are extremely sensitive to changes in market demand, corporate strategy, equipment productivity, die size, field size, re-use, and product performance-related issues. These changes typically result in the overall lithography exposure tool requirements going up or going down, thereby potentially putting Intel at risk for cancellation fees with lithography equipment suppliers. In this paper, we provide an overview of options that signal a breakthrough for Intel in this field. Options ensure Intel's flexibility to demand changes while at the same time limit Intel's cancellation risk exposure.

INTRODUCTION

In an ideal world, Intel would maximize revenues while minimizing equipment costs by bringing equipment up to production just in time to support the demand as shown in Figure 2.

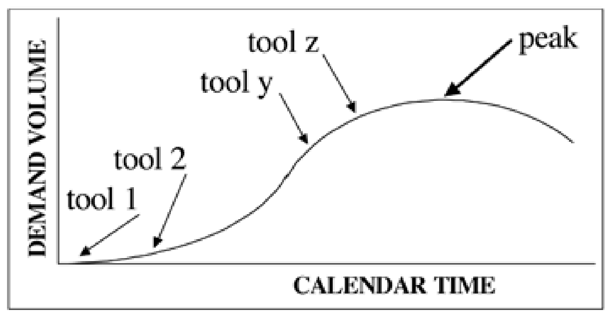


Figure 2: Just in Time production equipment

There are at least three key issues affecting this solution.

The first issue is the demand projection. It is difficult to predict what the peak volume of a particular product family will be or precisely when that peak volume will be realized. The risk of purchasing the first few tools of a lithography exposure tool model is small since Intel can control when and at what level production will start. However, purchasing the last few lithography exposure tools carries a higher risk since the market controls the peak height and the position over time. Delivering a lithography exposure tool early or having one that is not needed wastes Intel capital and increases Intel's costs. But on the other hand, not having a tool when it is needed, results in lost or delayed sales. Both cases can result in decreased profits.

The second issue has to do with equipment performance, and, again, predictability is the problem. For tools that have previously been run in manufacturing at Intel, there is a historical basis upon which to project performance parameters needed to calculate the number of tools required for a specific production volume. With new models there are estimated parameters, but confirmation of performance will not occur until well after the majority of orders are placed. Buying the first few lithography exposure tools carries a low risk since these tools will be needed regardless of their ultimate performance. However, refined tool parameters based on actual performance of the first few tools can easily translate into needing a higher or lower number of tools than originally predicted during the ramp.

The third issue concerns equipment suppliers. Given that some of the lithography production equipment is among the most complex and costly ever built, it is not surprising that long lead times and significant cancellation penalties are involved. In addition, once the tool has been built, it must be shipped, installed, and qualified for production which can take several months. Each of these steps has an associated uncertainty in duration, and these uncertainties stack up. Standard payment terms for suppliers in the semiconductor industry are x% of the equipment price 30 days after tool delivery and the remaining y% after the tool is satisfactorily installed at the Intel manufacturing factory. Long before actual payment is due, lithography suppliers are required to make a substantial investment in both research and development and in the pre-purchase of materials to deliver lithography exposure equipment in high volume. Technology Manufacturing Engineering (TME), a group within Intel that primarily deals with equipment development and capital procurement, created a program called "options" with the lithography equipment suppliers. This program provides the suppliers the incentive to pre-purchase high-cost materials and risk-build equipment resulting in shorter lead times for Intel. Shorter lead times and the flexibility of options have helped Intel as well as the supplier to favorably react more quickly to changing market conditions.

Problem Statement

The semiconductor industry downturn in 2000 left capital equipment suppliers with huge amounts of excess inventory that they had to either write off or sell at a loss. With net profits squeezed the equipment suppliers did not want to take more inventory risks. However, market conditions dictate factory roadmaps that in turn dictate the exact amount of lithography exposure tools needed in the Intel factories. The Intel challenge was to derive innovative solutions to order the right amount of equipment at the right time in an environment of volatile demand and tool performance. The supplier challenge was to work with Intel on alternative capacity-risk-sharing methodologies to enable faster response to market changes and reductions in cycle time. Both Intel and suppliers needed innovative solutions to address these problems. Figure 3 shows the variability in the semiconductor industry.

Due to the inherent uncertainties of forecasting, Intel tends to be conservative in estimating tool capacity parameters. This introduces a buffer into the system at the start of a ramp. However, as the ramp matures and more knowledge is gained on the tool, both the overall requirement forecast and the tool capacity parameters change, which can affect the quantity needed during the ramp and at the end of the ramp. This exposes Intel to higher cancellation fees with the suppliers, which in turn increases pressure on Intel to accurately forecast lithography exposure tool requirements and also motivates Intel to look for innovative ways to reduce overall risk.

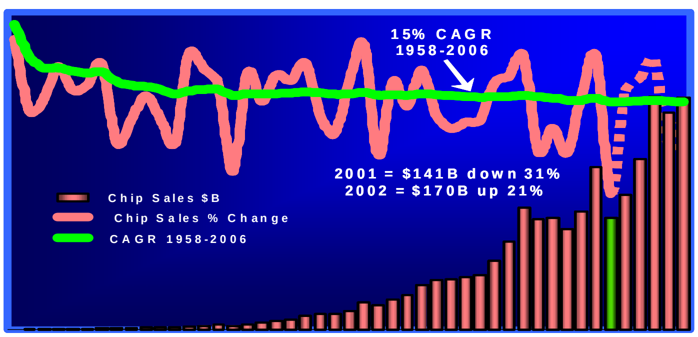


Figure 3: Forecast variability

Problem Resolution Today

Manufacturing equipment capacity needs at Intel are primarily driven by three broad factors:

Tool performance: changes to the Model Of Record (MOR) parameters of run rate, utilization, and availability.

Product requirements: product performance, die size or process changes which require a different lithography exposure tool type for a given step of the process.

Total capacity changes: ramps are increased or decreased resulting in changes to the number of lithography exposure tools required.

By the time that Intel is able to confirm tool performance, product requirements and total capacity requirements, we typically are in one of two scenarios: Intel has too many lithography exposure tools or too few. Until the last few years the market economics were such that if Intel had extra lithography exposure tools, it didn't affect them adversely. New cost pressures have changed capital spending expectations, driving Intel to seek more precise methods to meet our customer needs without spending too much too soon.

Intel's capacity problems are currently resolved through two tactics:

1. Over-forecasting: Subsequent cancellation.
2. Under-forecasting: Tool allocation.

Over-forecasting sets up both the supplier and Intel for excess capacity and cancellation costs. Intel business processes and systems were not proactive enough to prevent this from happening. However, our business realities are such that this has occurred. This situation leads to order cancellations with suppliers, which costs both Intel and the suppliers.

Allocation is the process used by Intel to allocate lithography exposure tool deliveries to the most important requirements based upon process priorities, process margins, and other factors. When forecasts occur inside supplier lead time, lithography exposure tool deliveries are allocated among the requests. This process leaves some needs unmet (tools are too late to meet needs) and requires intense work from Intel and the supplier in order to move tool orders around to best fill the new needs.

These two solutions (cancellations and allocation) are obviously undesirable and they required TME to come up with an innovative way to manage lithography exposure equipment capacity.

OPTIONS: AN INNOVATIVE CONCEPT

Options give Intel the right to purchase a tool in a reduced lead time at a certain pre-determined price. Intel purchases the options at a certain price from a supplier and must exercise or transfer the option to another tool prior to the expiration date.

Options provide Intel and the supplier with purchase order lead time, cancellation, and payment terms that are different from the standard terms and conditions of the standard corporate purchase agreement. Options provide a strategic approach to manage lithography exposure tool demand changes that are responsive to varying ramp needs but still limit Intel's cancellation liabilities with suppliers.

Options were originally developed for lithography suppliers since lithography tool lead times are long while their demand is subject to multiple changes in product demand. Due to their cost and lead time, lithography exposure tools are the primary ramp constraints at Intel's wafer manufacturing factories.

Conventional lithography tools have a long forecasted lead time plus long Purchase Order (PO) lead times as shown in Figure 4. These long lead times do not allow much flexibility to changes in demand for Intel or the supplier.

Additionally, since lithography suppliers have to invest upfront on the material and labor to build a lithography exposure tool, they inherently have high cancellation fees that increase over time up until the delivery date. Purchasing and exercising options provided Intel the right to procure un-forecasted lithography exposure tools in lead times much shorter than the contractual lead times.

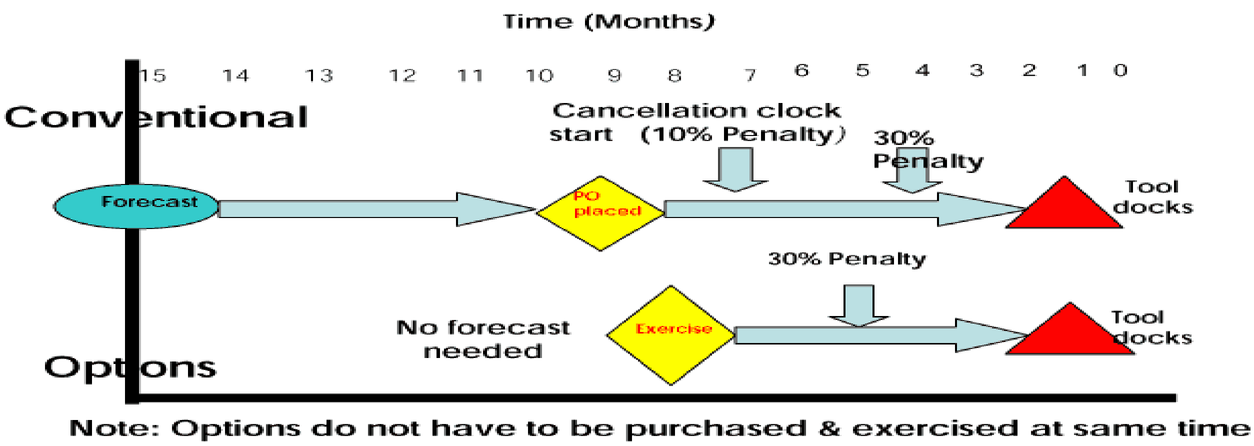


Figure 4: Illustration of options

An Intel team, with representation from Capital purchasing and Finance, developed a model to determine a fair option price that Intel would be willing to pay the supplier based on their actual cost of purchasing long lead material.

OPTIONS: BENEFITS AND RISKS

Options, being the first of its kind in both the semiconductor industry and for Intel capital equipment, has a unique set of benefits and risks that are listed below:

Intel Benefits

- Reduced lithography exposure tool lead time to react more quickly to upside demand.
- Additional time to place lithography exposure tool orders while limiting cancellation liability.
- Better management of lithography exposure tool forecast fluctuations due to Long Range Plan (LRP) or MOR changes.
- Limited risk in market downturns due to flexibility of down-payment transfer.

Supplier Benefits

- Sharing of long-lead material costs.
- High probability of tool purchase since Intel is motivated not to let the options expire.
- Incentive to risk building tool to meet short lead time.
- Options purchase (down payment) provides cash in hand earlier compared to conventional tool sale (time value of money).
- Competitive advantage (especially where Intel uses dual suppliers).

Program Risks to Intel

Options expiration will result in Intel losing the down payment. This could be due to end-of-life of the tool or because of poor management of the options by Intel.

The potential benefits of the options concept made it extremely attractive to one of Intel's new 300 mm manufacturing facilities to pilot and pursue options for its ramp. The Intel wafer manufacturing facility needed to place the POs for the final lithography exposure tools; however, there were several pending changes, which could drive down tool requirements:

- **MOR improvements**: The tool model was new to Intel and a first of its kind for the supplier. All MOR data were based on performance specs against the earlier generation tools; however, the actual performance of the new lithography exposure tools was unknown.
- **Process flow revisions**: The potential removal of two existing process layers would free up existing tools for re-use within Intel.
- **Process capability changes**: Pending process capability changes that impact equipment run rate would have had an impact on the total lithography exposure tool requirements.
- **Long Range Plan (LPR) changes to peak:** An increase or decrease in the LRP and Wafer Starts Per Week (WSPW) would change the quantity of the new tool requirements.

- **Product mix changes:** When a product's die size is shrunk, and changes are made to the die per field, the result is a change in the number of lithography exposure tools required. If the LRP product mix changes to incorporate mostly shrink products, more lithography exposure tools will be required.

Obviously, any of the above could go in the other direction: MOR degradations, tool capability degradations, or a higher volume ramp peak could create uncertainty around placement of the equipment POs. The current capital process dictates that we plan to the official Plan of Record (POR). However, given these uncertainties, Intel could spend millions of dollars on capital that would sit idle (and aging) for up to two quarters until it could be converted for use on the next Fab process startup. Figure 5 shows a typical ramp cycle and where options are being used in the ramp life cycle for lithography exposure tools.

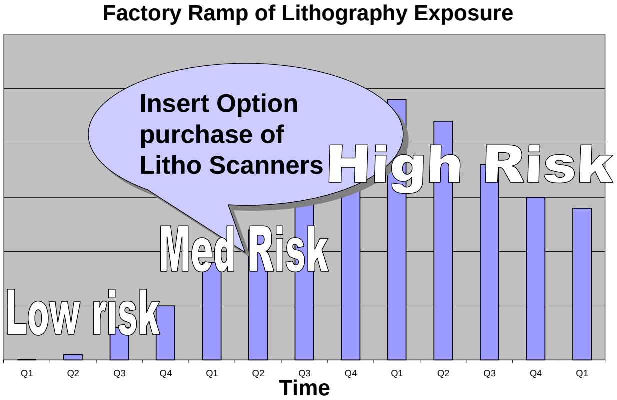


Figure 5: Intel use of options for ramp peak

OPTIONS: INTEL RESULTS

By purchasing options for three lithography exposure tools, Intel was able to delay the decision to place POs as well as delay the onset of cancellation fees by four months. In the semiconductor industry, where the equipment life cycle is 18 months for each technology generation, a four-month savings is significant. The four-month window allowed Intel to have better information on MOR, process flow, process capability, and the LRP and thus make a better purchase decision. The cost of this delay is the time value of money of the options price. The value of this delay is the ability to delay depreciation expense plus avoid installing a lithography exposure tool that is aging and idle.

In real terms, Intel estimated an 80% probability that at least one tool would not be needed and a 40% probability that a second tool would not be needed for the ramp peak at one of their 300 mm factories. On a wafer cost basis, if Intel could avoid the purchase of high-dollar value lithography exposure tools, it would mean a reduction in wafer cost due to reduced capital depreciation. In the past, taking any risk on the purchase of a tool that constitutes a wafer manufacturing facility constraint has not been feasible at Intel. The options program helped the Intel wafer manufacturing facility pursue a cost-savings opportunity for four months without any increased risk to output.

A cross-functional team within Intel, including representation from the Purchasing, Factory Planning, and Finance organizations developed an options management process that is shown in Figure 6. Options are pre-paid assets that have to be very closely monitored through various LPR cycles and are accounted for within our forecasting and accounting systems.

The team recommended purchasing options only if either of the two following conditions exist:

1. If an Intel wafer manufacturing facility has a lithography exposure tool requirement that is within the supplier's forecasted lead time and there is no alternative to get the lithography exposure tool delivered to Intel when needed.
2. If the Intel wafer manufacturing facility expects a decision within lead time that would reduce the number of lithography exposure tools required.

It became apparent that if Intel had a requirement for a new lithography exposure tool that was within the suppliers forecasted lead time and there was no alternative to getting a tool on time, the option's purchase would be a valuable tool.

Managing options is also very important as an expired option results in unforeseen write-offs to Intel.

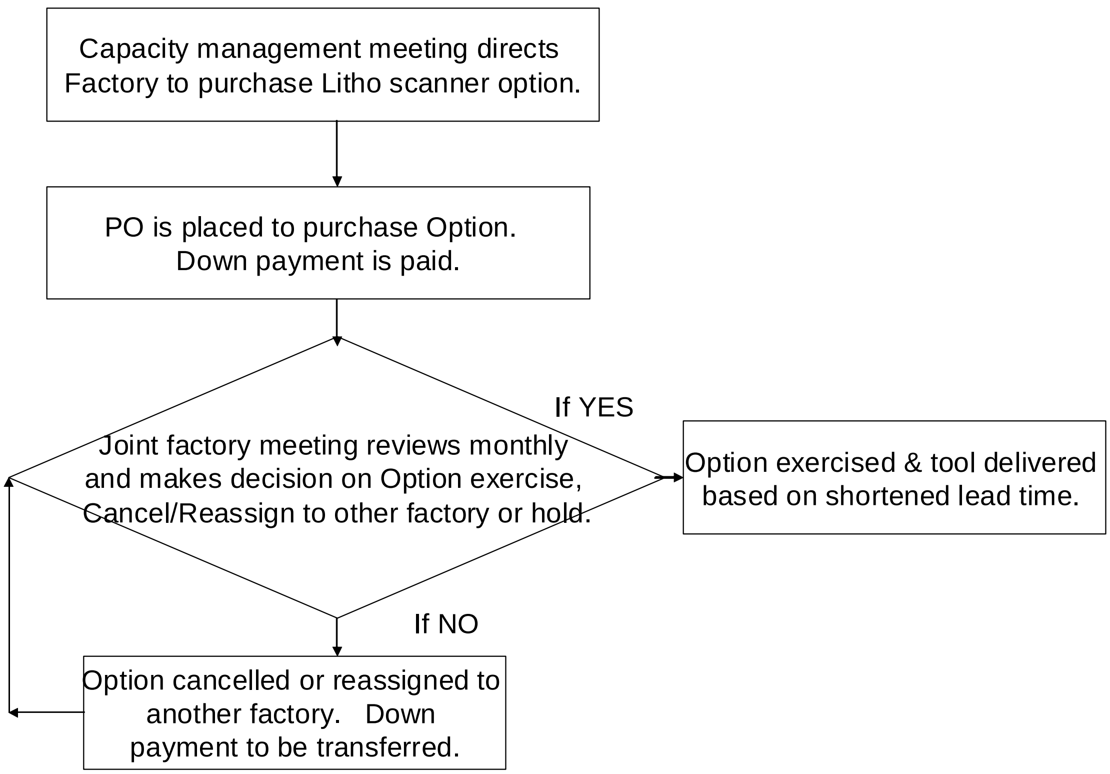


Figure 6: Options management process

CONCLUSION

Managing forecast variation in the Capital Supply Chain can be very complex. Sources of forecast variability are numerous. Tool needs are determined months in advance of peak requirements. Tool performance matures over time. Market drivers change within forecasted windows. Options have given Intel a significant competitive advantage and provided the flexibility to react to future upside or downside market fluctuations while limiting cancellation liability with suppliers.

The focus of this paper has been the use of options for lithography equipment, but the techniques described here can be used for any equipment used in Intel's manufacturing lines. All of the arguments made here are also applicable to the materials purchased to enable manufacturing–spares, piece parts, reticles, to name but a few. Extensions into contracting for transportation, software, and so on will come with time. But perhaps the most important reason to practice and continuously improve these techniques is so that Intel can continually improve in meeting our customers requests.

ACKNOWLEDGMENTS

The authors acknowledge many valuable discussions with and feedback from Beth Adkison, Bob Bruck, Joan Duemler, Barbara Haney, Karl Kempf, Janice Golda, and Peter Silverman, all of whom have enabled making the options program a success.

AUTHORS' BIOGRAPHIES

Viswanath (Vish) Vaidyanathan has been with Intel since 1994 and is currently a training manager within TMG-Training. He has worked in various capacities at Intel factories and most recently as the capacity manager in the Lithography Capital Equipment Development (LCED) group. He has also been the recipient of an Intel Achievement Award for his work related to ergonomics and safety during his tenure at Fab 9 in New Mexico. Vish holds an MBA degree from the University of New Mexico and an M.S. degree in Industrial Engineering from Wichita State University. His e-mail is viswanath.vaidyanathan at intel.com.

Dave Metcalf has been with Intel since 1987 with the majority of his time spent managing numerous capital equipment suppliers across a variety of factory functional areas. Dave's current position is the supplier and business manager within the Components Automation Systems group at Intel. His primary focus is to engage closely with suppliers in helping them perform to expectations and to help enable Intel's technology and cost roadmap. Dave was awarded the Intel Achievement award in 2002 for his work with a supplier to accelerate a critical technology. Dave graduated from Brigham Young University in 1986 with a B.S. degree in Accounting. He received his MBA degree from Arizona State University in 1987. His e-mail is david.r.metcalf at intel.com.

Douglas V. Martin Jr. has been with Intel since 1984 and has been a lithography capacity engineer in the LCED group for the past five years. Douglas has had experience in Fab, Sort, Assembly, and Test capacity planning within Intel over the past 20 years. He graduated from San Jose State University in 1983 with a B.S. degree in Industrial Management within the SJSU Industrial Engineering Department. His e-mail is douglas.v.martin.jr at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

Redefining the Test Equipment Supply Chain: The Open Architecture Revolution

Jason Katz, Technology and Manufacturing Group, Intel Corporation
Don Edenfeld, Technology and Manufacturing Group, Intel Corporation

Index words: test, open architecture, Semiconductor Test Consortium

ABSTRACT

As the automated test industry begins embracing the open architecture environment, equipment suppliers and their customers will need to evolve equipment development methodologies to fully benefit from the emerging business model.

The development of standards in the equipment industry allows suppliers to share in the cost of establishing basic infrastructural framework elements, releasing valuable resources to focus on development of distinctive, value-added technologies and services. Already a very competitive and financially unhealthy industry, the test equipment market will benefit from lowering the cost to deliver solutions to their customers. As a result, end users will benefit from increased innovation, more valuable capital assets, and reduced re-engineering.

Open architecture has evolved from vision to reality with the release of the Semiconductor Test Consortium's OPENSTAR* specification to the industry. With open standards, end users now have the ability to strategically manage the sustainability and extendibility of their fleet through a pipeline of module developments with traditional equipment suppliers and third-party developers for hardware and software solutions. With a more stable capital equipment fleet, end users can eliminate the cost and resource investments related to re-engineering and maintaining multiple solutions for similar problems and concentrate on improving their test processes, developing strategic supplier relationships, and innovating breakthrough technologies.

This paper illustrates the transformation of the supply chain to leverage the benefits of an open architecture. We focus on the structural challenges faced by the test equipment industry, demonstrate why the steps that have been taken are insufficient, and how open architecture can benefit the supply base as well as the customers.

INTRODUCTION

Semiconductor devices are among the most complicated structures designed and manufactured by humans and are becoming more complex with each passing moment. Regardless of this complexity, customer requirements demand that device incoming failure rates be measured in the 100s of defects per million or less. In the semiconductor manufacturing process, the test step is critical to this demand; it is pivotal to containment of defects and the product quality seen by the end customer.

Test is accomplished using highly automated test equipment that is designed to achieve highly accurate and repeatable results with high defect coverage and extremely high throughput. The fundamental test challenge is to execute the smallest number of measurements that cover the largest number of potential manufacturing defects in the shortest time possible. Typical test times are measured in the low seconds for devices of well over 10 million transistors.

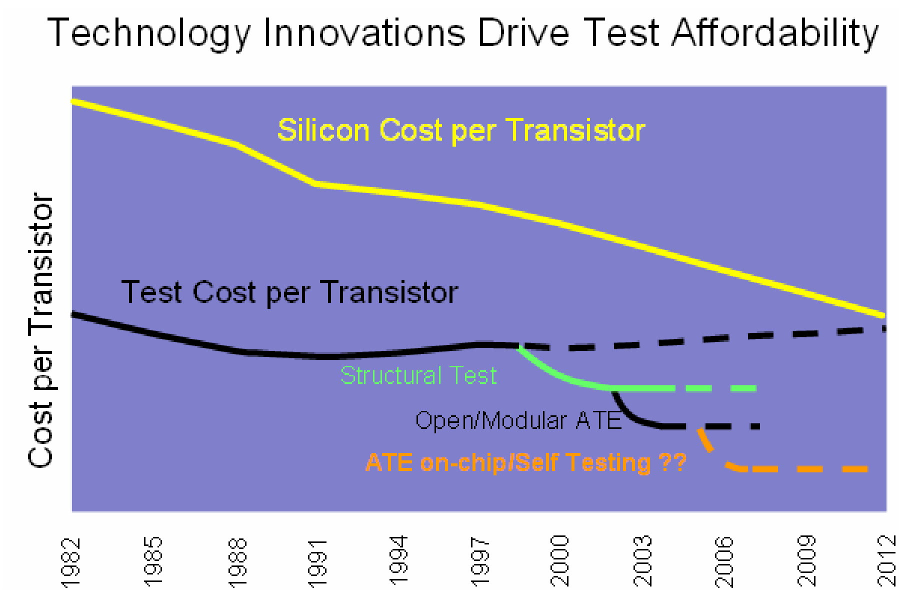


Figure 1: Historical cost per transistor [10]

Test equipment cost versus performance has been an on-going debate between semiconductor manufacturers and equipment suppliers. Semiconductor manufactures face an environment of shrinking device Average Selling Prices (ASPs) and time-to-market windows. This drives the need for just enough capability to test a particular device, with low capital and sustaining costs, available early enough to learn how to use the equipment effectively ahead of initial device silicon. The cost focus of manufacturing has driven development of a variety of low-cost equipment solutions over the last five years. This, combined with the fact that manufacturing tools represent the majority of the total equipment sold, has caused dramatic changes in the Total Available Market (TAM) and Return on Investment (ROI) of the equipment industry. Equipment suppliers face an environment of revenue constraints that has resulted in poor balance sheets and high research and development costs.

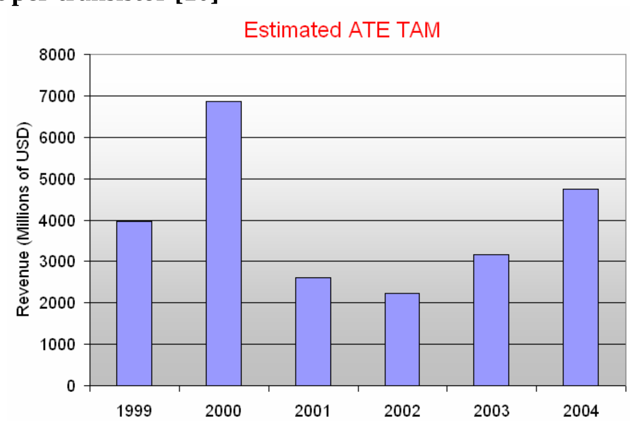


Figure 2: Estimated automated test market size [8,9]

In 2004, the test equipment industry represented a $4.8B industry where six major suppliers, each representing at least 5% Market Segment Share (MSS) garnered approximately 93% of the market. The TAM has been flat to down in four of the past five years and is 33% smaller than in 2000 as shown in Figure 1. Over this same period, Research and Development (R&D) investment has remained relatively flat, which has caused a disproportionate and unsustainable ratio of R&D as a percentage of revenue. These fundamental trends are not expected to change over the next several years which implies that business as usual may be catastrophic for one or more of the major equipment suppliers. Suppliers need to identify new methods to leverage development investment across many devices and customers and recover that investment as quickly as possible.

In this paper, we address a fundamental paradigm change that is emerging within the industry, a shift toward open architecture test equipment. Open architecture has the potential to reduce R&D costs while protecting Intellectual Property (IP) and innovation, and to increase productivity by targeting investment in new capabilities rather than re-engineering. The trends that enable this transition are described and the future landscape of the industry is discussed.

EVOLUTION OF THE TEST EQUIPMENT DEVELOPMENT MODEL

There are many different methodologies that are utilized to identify defects within a device. At a high level, these methodologies fall into one of two categories: functional and structural test. Functional test emulates the end use environment that would be seen by the device in the final application; tests replicate the actual function of the device, such as a system "boot" cycle in the case of a microprocessor. Structural test utilizes special structures that are included in the design to provide enhanced controllability and observability of internal device nodes; structural tests are specially written to disturb specific fault locations in the design and bear no relevance to actual device function. Each method has important implications for the capability of the test equipment: a functional tester typically needs to match the device performance while a structural tester may have significantly lower performance than the device.

Test plays three specific roles in the life of any device:

1. Product development uses test equipment to verify and guarantee device design functionality and performance.
2. Wafer probe is a test step that is done immediately following fabrication while the devices are still in wafer form. The wafer probe process step is driven by business decisions rather than product quality (with the exception of known good die requirements). The value of wafer probe is in the reduction of scrap costs from two sources: rapid data feedback to reduce misprocessing due to fabrication excursions and early identification of defects to reduce downstream processing costs of defective material. These savings opportunities must be balanced by careful management of the cost of the wafer probe process. Process cost can be reduced through many techniques, some of which may result in a reduction of test coverage at this step, by identifying fabrication excursions as early as possible.
3. Final test is done after the device has been packaged, typically as far downstream in the process as possible to minimize the risk of introducing defects after test. Final test is responsible for all remaining defects to ensure end-customer quality.

Product development is highly dependent on functional- and specification-based test methods, demanding the highest performance and typically most expensive equipment. Wafer probe and final test may contain a combination of structural and functional tests; the selection and implementation of these tests determines the complexity and cost of the required equipment. The investment structure of the industry faces the fundamental challenge that the highest cost and investment intensive equipment has a very small market potential (product development). The lowest cost equipment serves the largest market, but lower margins starve the R&D requirements of leading-edge technology. Further, semiconductor manufacturers typically demand that the same platform service all purposes in order to increase engineering productivity.

The rapid pace of advancement for the Device Under Test (DUT) has meant that the equipment designer was constantly faced with providing a capability that actually processed information faster than the DUT, but had to be constructed out of older generation technology. The significant performance disadvantages of the available components meant performance would need to be derived by architectural innovation. These architectural innovations typically resulted in sharp increases in equipment complexity.

In the 1990s, a typical new platform design required more than 100 hardware and software development engineers, an investment of $50-100M, and a 24-36 month cycle time. This investment and time-to-money scenario resulted in a tool capital cost of several million dollars and the need to sell several hundred tools to generate reasonable profit margins. In this generation of equipment, individual platform designs were targeted to match customer market segments to partition the test problem and reduce equipment design complexity. This resulted in the traditional memory, mixed signal, and logic test platform delineation.

Equipment design was approached from the system level, with little to no consideration given to feature growth. Design tradeoffs driven by practical cost, resource, and manufacturability considerations resulted in the selection of custom ASIC design for critical circuits and off-the-shelf components for basic functions. The ability to encapsulate tester functions was limited by the low integration density of available components and the practical limitations of Printed Circuit Board (PCB) sizes.

The complex and interwoven nature of the equipment design generated highly proprietary systems, requiring an intimate knowledge of the circuits and interconnects (Figure 3). The equipment was targeted at a very narrow capability window and delivered just-in-time. In the end, the customer spent several million dollars for each tool and typically experienced poor reliability, highly complex diagnosis, with limited extendibility to meet future requirements.

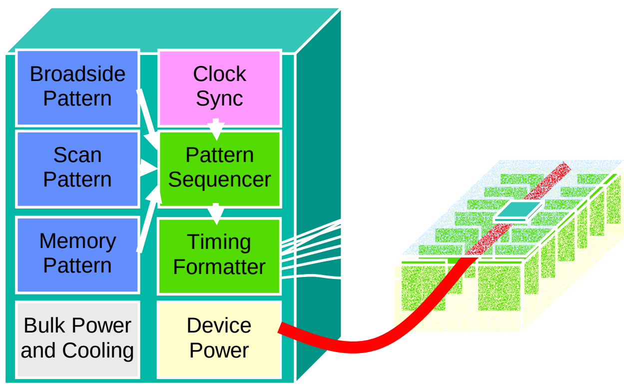


Figure 3: Traditional test equipment architecture

Moore's Law stipulates that the number of transistors available to design on a chip will double with every new process node. At the same time, the transistor switching speed increases. This rapid change in device design complexity and performance when combined with the nature of test equipment development represents a capability gap for design verification and product manufacturing. Further, the rapid obsolescence of the equipment creates a significant cost barrier due to the need to replace the entire manufacturing fleet with each new device design; a cost barrier significant enough to make or break the product.

In today's design and manufacturing environment the traditional equipment development model breaks for several fundamental reasons.

1. Device design time is continuing to shrink while equipment design cycle time for a new platform has remained essentially flat.

2. Device design complexity is increasing and product segments are collapsing, making the traditional device-type-based test partitioning obsolete–a single device now requires all of the capabilities that have traditionally been partitioned between distinct test platforms.

3. Device performance is increasing at a pace that makes new test equipment obsolete almost before it can be delivered, challenging the ability to achieve a reasonable return on invested capital.

4. Platform conversion costs are a significant portion of any equipment selection decision and represent a barrier to entry for new suppliers.

5. No single platform is capable of meeting all of the needs of the market, or in many cases even a single customer. Further, most customers are unwilling to align with a single supplier due to concerns over the business impacts of eliminating competition in a highly proprietary market.

A paradigm change was needed in equipment development to enable cost-effective engineering and production test without sacrificing leading-edge capability. The key to achieving this change was the significant advancements in circuit integration levels to provide encapsulation of equipment function into a practical physical space. This has enabled a transition to test instruments and the concept of a universal slot equipment architecture. This architecture creates a generic slot definition: all test instrument functions can then be designed to fit within that slot. The result is a platform infrastructure that may remain fixed over an extended period of time while significant new capabilities are introduced in the form of new instrumentation.

Current-generation test equipment is based on this universal slot architecture. The infrastructure has been reduced to power distribution, cooling, and communication, based on fixed, generic budgets on a per-slot basis (Figure 4). Instrumentation can be populated as needed, plugged into any slot, and integrated into the existing software environment. The result is a highly modular, configurable tester with minimal retraining to add new capabilities.

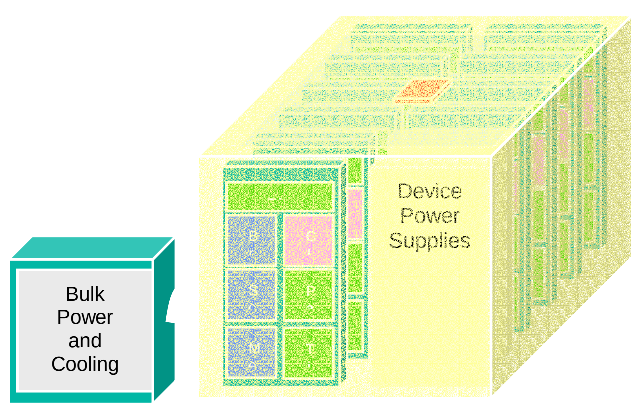


Figure 4: Universal Slot Test Equipment architecture

This architecture enables the user to purchase just what is needed, driving significant increases in equipment re-use and associated capital cost savings while also minimizing the training and transition costs associated with a new platform. The supplier need only develop the incremental capabilities, focusing resources and investment while reducing new capability design cycle times.

The transition to the universal slot architecture has been a significant step for the industry, necessary but insufficient to address many fundamental issues that remain.

1. Where does the customer go to find instrumentation to support specific device test requirements when the supplier is unwilling or unable to provide it at a competitive cost?
2. Most universal slot architectures lack fundamental market differentiation outside of the breadth of instrumentation available to support them. Continued investment in several competing slot definitions is inefficient and unproductive.
3. Few universal slot architectures have been defined sufficiently to stand the test of time. Keeping the shape and color the same while requiring full replacement of all of the infrastructure and instruments misses the target, even if the software environment remains somewhat stable.
4. Each individual platform has unique facilities requirements that can cost factories millions of dollars in retrofit costs when changes are required to add capacity or adapt to changes in product mix.
5. Each individual platform still has a unique programming interface, user model, and maintenance model that carries a significant investment in training, core competency development, and management business systems.

The industry is in transition. Current platforms that are based on this architecture are being marketed as sufficient to sustain customers for many years to come. If this is the case, one would theorize that the market for platform sales must eventually saturate and reach some steady state size (or at least as close as can be expected for such a cyclical industry). If platform sales saturate, and this is the value proposition of the supply base, then there is a significant business model challenge looming that will challenge the economic structure of the test equipment industry beyond what it already faces today.

THE OPEN ARCHITECTURE OPPORTUNITY

Coming out of the largest downturn in semiconductor history, the major test equipment suppliers are generally laden with poor balance sheets, unhealthy R&D ratios, and gloomy growth forecasts. Despite the high ratio of R&D spending, even the largest ATE vendors cannot be "everything to everyone" as demonstrated by the MSS disparity shown in Figure 5.

	MSS		
Test Segment	1st	2nd	3rd
Memory IC	Advantest 71%	Yokogawa 17%	Agilent 6%
Digital / Logic IC	Advantest 26%	Yokogawa 17%	Credence 13%
Mixed Signal IC	Advantest 46%	Teradyne 23%	Credence 8%
Analog / Linear IC	Credence 64%	Agilent 15%	All Others 20.6%
RF / Microwave	Teradyne 51%	LTX 18%	Credence 13%
SoC Test	Teradyne 52%	Agilent 17%	Advantest 8%

Figure 5: ATE market segment share [9]

In the past year many industry experts, like Gartner Dataquest, have predicted further market consolidation beyond the 2004 acquisition of NPTest by Credence, due in part to the disparity between the R&D required to develop new systems and the total available market.

In the early eighties, Dan Hutcheson of VLSI Research developed an equation that theorized how many suppliers a given market can sustain. Conceptually, any given market can only support a certain number of suppliers depending on the R&D required to develop a product and the total available market. The hypothesis provides valuable insight into the alarming health of the test equipment industry as it helps illustrate one of the fundamental problems in the marketplace: redundant and unjustifiable R&D.

Applying this equation, we estimate that the ATE industry can sustain three or four major suppliers (depending on assumptions) *without government or industry consortia intervention.* Currently, the test equipment industry consists of six major suppliers that have lost a combined total of approximately $4.2B since 2002. Although most of these companies returned to profitability in 2004, due in part to the 50% market growth, most analysts are predicting a steep TAM decline over the next two years. If this decline occurs, further consolidation is inevitable, if the fundamental cost structure of the industry cannot be reduced substantially.

The Open Architecture Initiative, begun in 2002 by Intel and currently represented by the Semiconductor Test Consortium (STC), put forth the concept of a standardized infrastructure architectural definition as a basis for combining instrumentation from multiple suppliers into a common platform. The goal of this effort is to leverage standards at the instrument interface level (power, cooling, communication, and device interfacing) to focus R&D on what customers actually pay for–the ability to test their devices.

The STC has turned this concept into reality through definition of the OPENSTAR architecture and has published a set of related instrument standards (available at http://semitest.org/site/About/Specifications* that are available to the industry. OPENSTAR leverages a universal slot architecture, focused on defining the interfaces to allow interoperability without stifling innovation or increasing the risk of intellectual property exposure. Focusing precious R&D resources on intellectual property development and sharing infrastructural development costs across the industry will lower the cost to develop new products and allow for a more healthy industry balance sheet.

Despite the current significant (and unsustainable) level of R&D funding, no single supplier has been able to provide the entire spectrum of test capability. Open architecture enables suppliers to focus on their areas of core competency to deliver value while enabling the customer to minimize the platform diversity that their engineering teams and factories must manage. Of critical importance is the realization that the infrastructure standards enable a diverse industry environment as shown in Figure 6. Such an environment comprises several vertically oriented system Original Equipment Manufacturers (OEMs) that provide complete development, integration, and sustaining services; as well as more horizontally oriented services suppliers who are focused on instrument development, qualification, integration, logistics, and field support.

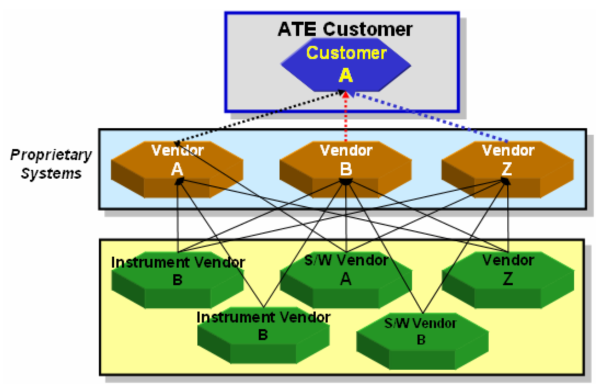


Figure 6: Traditional ATE supply chain

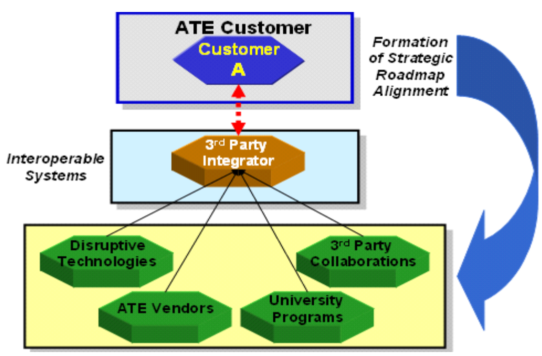


Figure 7: Open architecture supply chain

As depicted in Figure 7, the ATE user can strategically utilize and collaborate with Tier II suppliers to acquire optimal factory solutions. This mixture of vertically oriented solutions with cost-effective and efficient development of lower-volume, customer-specific solutions is a key enabler to increased innovation with decreased R&D costs. Customers can leverage a wide array of suppliers including university research, individual instrument suppliers, as well as traditional test equipment supplier offerings while providing a consistent infrastructure and training requirement. Open architecture permits customers to develop strategic technology pipelines while maintaining the ability to incorporate disruptive technologies into the existing test infrastructure.

At the heart of a horizontally oriented market is the need for instrument qualification and system integration services. Customer requirements for well integrated and sustainable test equipment have not changed. The ability to deliver this type of solution has been the strength of the traditional test equipment suppliers, but at the cost of proprietary, closed architectures. The ability to provide well-integrated systems containing instrumentation from multiple suppliers that can be efficiently maintained and serviced in the field is a core requirement of a successful open architecture.

The system integrator function offers several specific value-added services to the equipment as listed below. Note, however, that this is far from a complete list as the highly scalable and configurable nature of open architecture equipment, while solving many problems, will produce new challenges and exacerbate existing ones:

1. Design support for new instrument developers to lower the barrier to entry and simplify the learning curve.
2. Development services to simplify instrument software integration and check-out in the full system environment.
3. Qualification and certification services to verify that instruments conform to the hardware mechanical and interface standards.
4. Confirmation of interoperability with other instrumentation.
5. Electro-mechanical infrastructure sourcing and integration.
6. Specific system configuration integration, verification, and sales.
7. Worldwide support for complete systems post sales, including logistics, spares, service, and applications.

In this model the integrator could be either one of the vertically integrated OEMs (modeled after the traditional test equipment supplier) or a non-traditional horizontally focused third party. These services differentiate the integrator or OEM and enable higher operating margins than acting as a pure play distributor, where the customer is required to integrate each of the individual system components. Due to the steep and critical nature of product ramps and relatively limited user expertise, customers will typically not accept the additional risk associated with integration in house.

Beyond the opportunities evident in equipment development, standardized slots allow the equipment infrastructure to further embed itself within the factory facilities. Customers will find that asset management becomes focused on the instruments rather than at the system level as it is today. This represents an order of magnitude increase in business system complexity and opens the door for many value-added services including configuration management, instrument reliability and maintenance history tracking, spares depots, and field support for applications and maintenance.

Additional opportunities will also emerge for fundamental changes in how instrumentation is valued and paid for. Rental or leasing options will be more cost effective and a lower risk for the customer as well as the capital owner. This enables customers to rapidly and cost effectively flex equipment capability to adapt to changing market dynamics and natural shifts in requirements from customer to customer, as technologies are phased out of one company and brought up in others. The interoperability of the modules will allow individual instrument designs to appeal to a broader customer base over a longer period of time, thereby deriving greater revenue per design.

Open architecture is the logical next step for an industry that is already converging towards proprietary implementations of fundamentally similar architectures. The traditional test equipment supply base is already facing difficulties differentiating their product based on architecture. The forecast longevity of these systems forces a business model change to focus on deriving revenue from incremental capability sales based on instrumentation. Open architecture is the logical end state where the platform and infrastructure are based on standards, and the supply base is focusing R&D investment on what customers are willing to pay for: value-added technology development and service delivery.

BENEFITS OF OPEN ARCHITECTURE

Open architecture creates many opportunities for the test equipment supply base and the customers, but the benefits need to be clearly defined. Transitioning into an open architecture marketplace radically changes how the supply chain is managed and the relationships between suppliers and customers. How can this be justified?

Historically, the ATE industry has been mired in an adversarial seller to buyer relationship. Customers requiring test solutions carefully canvas the industry to find the most optimized equipment to meet product cost of ownership and technical requirements while suppliers scramble to profitably meet cost and technical targets set by the buyer. Customers attempt to drive the cost-learning curve of their product environment into the supply base while suppliers struggle to justify the investment in new development. Neither side believes the positions taken by the other are reasonable, and in the end reach a stalemate of dissatisfaction where there are no obvious choices. Within the industry, pockets of “strategic” agreements have been put in place between customers and their suppliers, but customers constantly drive competition to minimize exposure.

The open environment allows such strategic alliances to take hold and provide the valuable ROI that they are intended to produce by driving competition to the instrument level. No longer does a customer need to hesitate over the selection of the platform based on concerns over whether it will be positioned to meet the requirements after several years of careful investment and deployment.

Admittedly, open architecture testers will struggle to show a dramatic cost of ownership improvement over competing current-generation proprietary solutions when an initial ramp of capacity is occurring. In this scenario, open architecture testers will provide a marginal cost-savings benefit (at best). The true value of open architecture is evident during the follow-on technology ramps and product mix transitions as investment becomes incremental with a high degree of confidence versus replacement.

Open architecture opens the door to many optimizations in which both the supply and customer base can benefit:

- ***Acquisition costs:*** Suppliers need only invest in the development of new technologies or compelling value-added extensions of existing solutions. R&D investment is lower for the supplier, and capital acquisition costs are lower for the customer, reducing time to money for both parties. Customers need only purchase the instrumentation needed to adapt to specific device requirements.
- ***Capacity management costs:*** By extending the life of the tester infrastructure and pushing the infrastructure into the factory, customers can efficiently flex capacity by shipping instruments instead of one ton testers. Testers require expensive and careful logistics

planning to locate the specific transport method able to deal with the size and weight of the complete system.

- ***Utilization:*** As product test plans are developed to take advantage of this environment, the factory is able to rapidly adapt existing capacity to meet ever-changing volume mix requirements. Costly and time-intensive platform conversion steps can be eliminated. High utilization also frees up valuable factory space to alleviate existing bottlenecks and improve factory output per square foot.

- ***Upgrade costs:*** Most equipment purchases are driven by incremental testing requirements. Upgrades can be a fraction of the cost of a full system, but must be available within the platform. Open architecture lowers the risk that capability will not be or can not be made available to meet the need at the same time that it lowers the supplier investment in providing new capabilities.

- ***Factory efficiency:*** The cost of maintaining multiple test platforms arises from many sources: it ranges from the ability to guarantee that a given spare part is available to the ability of a particular operator to drive the equipment. The fewer the number of platforms, the more operationally efficient the factory becomes in terms of headcount and inventory expenses.

- ***Cost of spare line items:*** Due to the consolidated equipment base, the number and breadth of spare line items can be vastly reduced. No longer must duplicate instruments be stocked to provide essentially the same functionality simply because the platform they plug into is different.

- ***Training costs:*** Utilization of common platforms allows engineers and technicians to focus their training on new technology instead of the entire programming, maintenance, and operating procedures of a new platform. More important than the reduction in training cost is the fundamental improvement in engineer and technician expertise as they focus on fewer variables.

- ***Opportunity costs:*** The projects, developments, and opportunities that are lost due to the limitations of the existing test equipment model are substantial. Moving to a standardized platform infrastructure allows the customer to integrate the "slots" into the factory and make engineering and manufacturing test decisions based on how those slots are populated with specific instrumentation. This enables unprecedented flexibility to adapt the equipment to ever-changing device requirements without the need for costly and resource-intensive platform conversions.

- ***Maximizing ROI:*** Equipment suppliers will be able to focus their resources on the areas that customers truly covet: innovation, IP creation, and capability development. Liberating supplier resources from mundane platform tasks enables them to provide more value, delivering more services but at a lower cost, thereby increasing profit margins.

- ***Strategic relationships:*** Open architecture can eliminate the adversarial customer versus supplier relationship by lowering investment risk and enabling third-party support. These strategic alliances allow companies to define the key development areas, decide which technologies to pursue, and mature the process of transferring new technology to the factory. Where a customer roadmap diverges from its suppliers', the companies can strategically choose the opportunities in which to engage with alternate suppliers without severing or damaging the relationship (Figure 8).

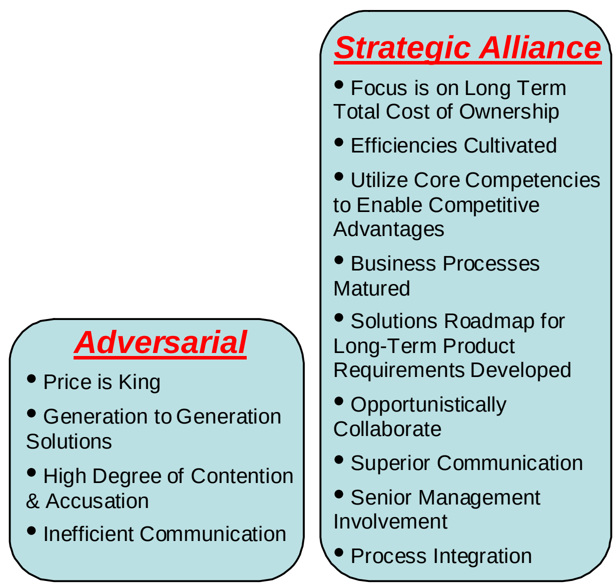


Figure 8: Supplier management shift

CONCLUSION

Economic indicators are beginning to challenge the traditional business model of the test equipment industry. The disparity emerging between the increasing operating and R&D costs of the supply base and the flat or decreasing total available market is not sustainable. There is little indication that there will be a significant increase in the size of the overall market in the near term; as a result the industry must look for opportunities to reduce cost while continuing to deliver cost-effective test solutions.

Initially with Intel's Open Architecture Initiative, and now in the Semiconductor Test Consortium OPENSTAR specifications, a shift from proprietary, monolithic equipment toward modular, scalable architectures and interchangeable instrumentation has been taken from vision to reality. OPENSTAR compliant equipment is now available on the market and a significant number of tools have been deployed in production.

Open architecture enables the supply base to focus investments on value-added services, intellectual property innovation, and product development. Further, it provides an environment where suppliers can focus on their areas of core competency to develop best-of-breed capability without being distracted by other portions of the test requirement that are necessary but outside of their expertise. No longer must every supplier be able to be everything to everyone.

The fundamental challenge in open architecture lies in the restructuring of the industry that a change of this significance entails. There are many opportunities for new value-added services as well as for traditional equipment suppliers. The industry is beginning to embrace change and make real progress in establishing new business models; this is a long-term strategic direction that will take many years to achieve.

REFERENCES

[1] Hutcheson, Dan, "The number of competitors that can be sustained in any given market is bounded by R&D requirements and market size," Chip Insider, September 25, 2003.

[2] Advantest 2004 Annual Report (Ending March 31st, 2004), http://www.advantest.co.jp/investors/annual-report/annual-report-2004/pdf/overview.pdf *

[3] Agilent 2004 Financial Reports (Ending October 31st, 2004),http://www.investor.agilent.com/phoenix.zhtml?c=103274&p=irol-fundIncomeA*

[4] LTX 2004 Annual Report (Ending July 31st, 2004), http://library.corporate-ir.net/library/10/108/108944/items/145940/87691_LTX_AR_04.pdf*

[5] Credence 10-K Report (Ending October 31st, 2004), http://media.corporate-ir.net/media_files/irol/68/68634/reports/10K103104.pdf *

[6] Teradyne 2004 Annual Report (Ending Dec 31, 2004) http://ccbn.tenkwizard.com/filing.php?repo=tenk&ipage=3337260&doc=1&total=&attach=ON&TK=TER&CK=97210&FG=0&CK2=97210&FC=000000&BK=FFFFFF&SC=ON&TC=FFFFFF&TC1=FFFFFF&TC2=FFFFFF*

[7] Yokogawa 2004 Annual Report (Ending March 31st, 2004), http://www.euroland.com/arinhtml/jp-yke/2004/ar_eng/mainPage.asp?CompanyCode=jp-yke&aryear=2004&artype=ar&replang=eng*

[8] "Long-Term IC Manufacturing Equipment History and Forecast, v5.05," *VLSI Research.*

[9] "Market Share: Semiconductor Test Equipment, Worldwide, 2004," Mark Stromberg and Laurie Balch, May 2005.

[10] SEMI SBC Presentation, Sunlin Chou, Intel Corporation, May 2005.

[11] [10] "STC Presentation," Semiconductor Test Consortium, October 2004.

AUTHORS' BIOGRAPHIES

Don Edenfeld is an engineering manager in the Intel Test Operation group in Hillsboro, Oregon. He received a Masters degree in Electrical Engineering from the University of Virginia in 1996. His team is responsible for equipment strategies for engineering and production test. His e-mail is donald.e.edenfeld at intel.com.

Jason Katz is a commodity manager in the Intel Test Operation group in Chandler, Arizona. He received his Bachelor's degree in Finance from Indiana University in 1999 and is currently pursuing his Masters degree in Business Administration. He is responsible for developing commercial strategies that enable engineering and production test. His e-mail is jason.r.katz at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

e-Procurement–Strengthening the Indirect Supply Chain Through Technology Globalization

Kinnar Ghiya, Technology and Manufacturing Group, Intel Corporation
Marci Powers, Technology and Manufacturing Group, Intel Corporation

Index words: supply chain, technology, ERP, indirect materials, cost savings

ABSTRACT

Indirect Materials (IDM) can be defined as goods and services that are not directly used in the production of Intel products. IDM spending accounts for 60% of Intel's procurement spending with a significant impact on Intel's supply-chain and bottom-line profitability.

In 2002, indirect procurement at Intel was sub-optimal with 60+ ways to buy. There were purchase order delays impacting internal customers, limited aggregation of Intel's spending power, and no standardization of global systems. To address these gaps, the Materials organization initiated a comprehensive transition plan aimed at unifying the organization and creating a world-class global procurement solution. The program termed "e-Procurement" kept a keen focus on the global end state and targeted three focus areas: tools, people, and processes. This report focuses on the technological advancements made to strengthen the IDM supply chain.

e-Procurement focuses on several innovative solutions with a single global Enterprise Resource Planning (ERP) system as a foundation. An Internet negotiations tool was introduced to achieve additional cost savings and negotiation efficiencies. Now, suppliers participate in live on-line reverse Internet Negotiations to win Intel's business. New online "e-Catalogs" directly connect requisitioners to the supply base and provide efficiencies through touchless transactions and contract compliance. Several data models were improved and a reporting system was introduced giving visibility into global spending by supplier, commodity, and country.

INTRODUCTION

The manufacturing industry has traditionally focused its procurement resources on optimizing procurement practices in Direct Materials (DM), defined as materials needed to make the product. In the past several years, organizations have realized that they spend 60% of their procurement spending on Indirect Materials (IDM), goods and services that are not directly related to the making of a product (Figure 1). Indirect procurement provides the "next big opportunity" for organizations to optimize the supply chain and save money.

Figure 1: What constitutes Indirect Materials (IDM)

Intel's Indirect Materials (IDM) organization launched an initiative called "e-Procurement" back in 2002 that focused on strengthening the IDM supply chain. The major focus of the program was technology enhancements and additions. The program also focused on improving business processes and people skill sets.

This paper outlines the challenges, improvements, and results for the e-Procurement program that transformed the Intel IDM organization into a world-class supply-chain benchmark.

CHALLENGES

It is worth noting that the solution to the IDM challenge was not strictly a technical one. Technology-only solutions have previously ended in failure. Success began with business process analyses that leveraged the available technology. The starting point for our technology success was to focus on data. The team spent considerable time understanding the "as is" process and then used a rigid Total Quality Data Management (TQDM) process to ensure that the "to-be" process established solid data quality. Figure 2, also known as "the mess," depicts the "as is" process.

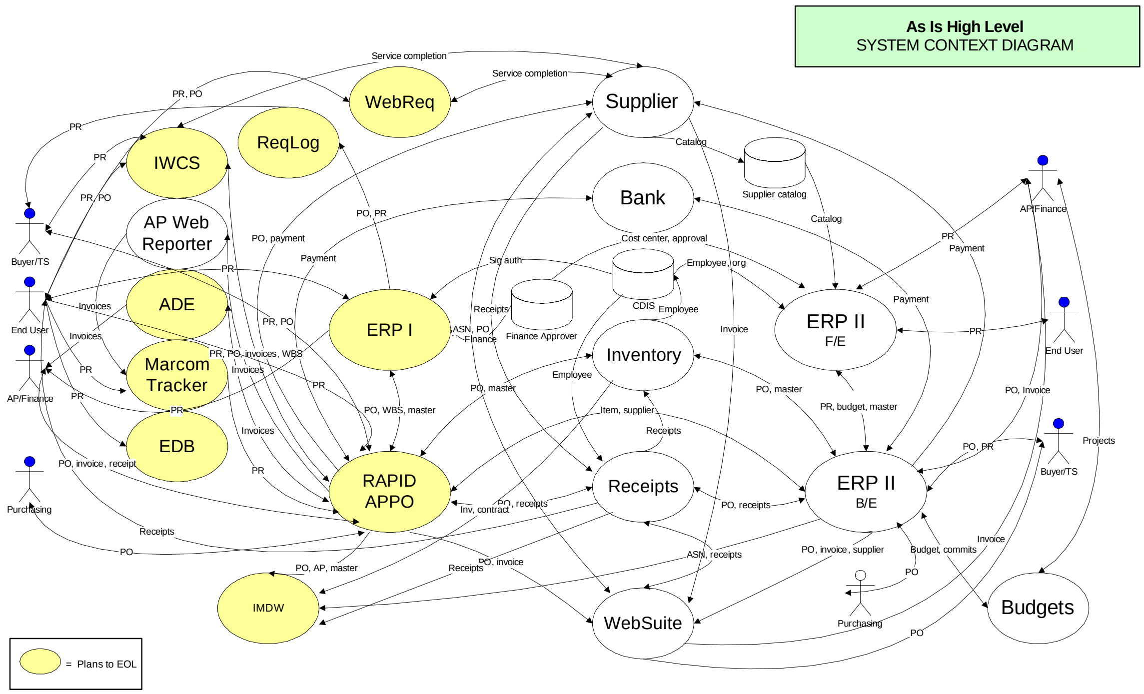


Figure 2: 2002 "As-Is" system

After the data analysis was completed several conclusions were drawn. The solution relied heavily on home-grown and highly customized applications. Based on the amount of data flowing among systems, it is difficult to link data back to the originating system of record with audit requirements. The overall architecture had been developed without regard to data visibility at an enterprise level. The whole "as is" process was extremely high maintenance: the architecture did not meet all of the business needs of the organization, high cost of ownership was associated with the solution architecture as deployed and maintained, and the system was built incrementally over time, resulting in a complex solution.

Drivers

There were several triggers that intensified the need for a significant transformation for IDM and services.

Sub-Optimal Procurement Solution

In 2002, Intel employees who needed IDM or services had 60+ ways to buy and pay for them. Methods ranged from manual to automated solutions. There were two issues with this sub-optimal solution. First, the requisitioner was unsure of the correct method to buy or pay leading to incorrect method utilization. Second, it required significant maintenance on procurement and IT organizations to maintain these methods. There was a need for standardization and automation.

Lack of Spending Visibility

In 2002, IDM did not have any reporting solution that was able to measure Intel's global IDM spending accurately. The data visibility was limited. There was a lack of standard material schema codes to enable spending aggregation. In a market that was getting increasingly global, it was important for Intel to accurately measure spending in order to maintain greater control over spending and have greater leverage when it came to negotiating terms. Finally, there wasn't any systematic solution that measured the spending by contract, by supplier, or by geography.

Revenue vs. Consumption Growth

The e-Procurement team compared the IDM spending trend with Intel's revenue trend in 2002. The findings were concerning: the consumption (IDM spending) was growing at a faster rate than revenue (see Figure 3). This was directly impacting Intel's bottom-line profitability. It became important to put solutions in place that provided global spending visibility. It also became important for the IDM team to influence internal customers to reduce spending.

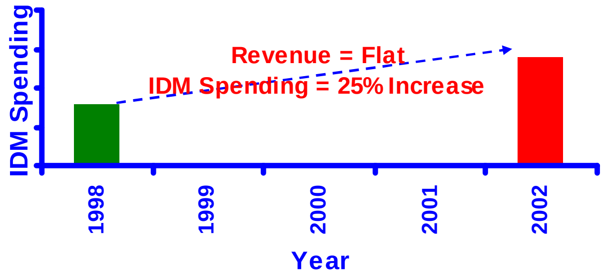


Figure 3: IDM spending growth

Percentage of Spending in IDM vs. DM

Traditionally, many companies, including Intel, focus mainly on reducing the cost of DMs; these costs are more visible as they relate to product or services costs. However, the e-Procurement team found that similar to several other companies, Intel spent 60% of their procurement dollars on IDM and services (see Figure 4). The Materials organization clearly needed to look closely at the IDM supply chain.

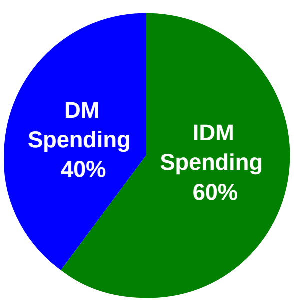


Figure 4: Intel spending pattern

Intel Revenue Pattern

As personal computing becomes prevalent across the globe, a higher and higher percentage of Intel's revenue comes from global markets. This global trend in revenue also has implications for procurement systems and tools. To support the global trend in revenue, many of Intel's business units are now locating globally. The e-Procurement team needed to look at standardization of purchasing processes and tools across the globe to support this trend.

Increase in Maverick Spending and Supply Base

If IDM continues to work in "regional" silos and Intel continues to expand into the global workforce, the resulting effect will be an increase in the supply base and an increase in spending with non-preferred suppliers (also called "maverick spends"). For this reason alone, Intel must have an IDM global spending policy. This also provides an opportunity to optimize Intel's supply base, thus reducing supplier management expense.

Industry Benchmarking

Research was conducted on several companies to find an optimal solution to a strong global supply chain. The IDM organization investigated best-known methods and several companies shared knowledge. Among the practices that needed to be observed were the following:

- Utilize e-Tools.
- Have a standard global procurement process.
- Have a standard global sourcing process.
- Have global procurement teams.
- Engage business partners.
- Have global spends aggregation.
- Enforce "correct ways to buy."
- Put consumption reduction programs in place.

MANAGEMENT VISION

Senior management can play an important role in changing the IDM process by creating a vision and setting the stage for change. A large-scale improvement program without vision and buy-in from management is one of the key reasons for failure. The vision should identify areas of focus, key metrics, and an overall timeline. This is exactly what the Materials organization management and IT organization management did. Their vision is known as the World-Class Indirect Procurement vision (WCIP) (see Figure 5).

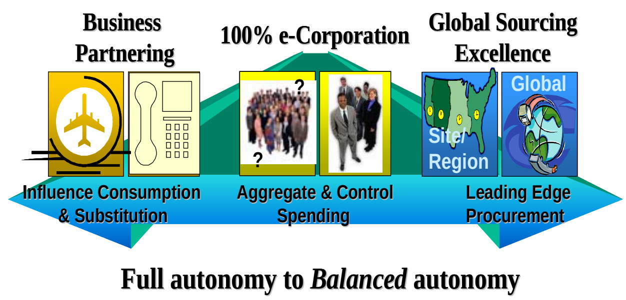


Figure 5: World-Class Indirect Procurement (WCIP)–IDM vision

The vision helped the organization align to common goals and focus areas and provided a way to communicate and check progress.

Solutions

Our industry benchmarking revealed that a "technology only" approach would have limited success in strengthening the IDM supply chain. The e-Procurement program had to be more than just a new set of tools. The structure of the organization, existing processes, people skill sets, and behaviors also needed attention. As a result, while the overall program focus remained on technology improvements, significant improvements in process and people systems were made. A combination of these three factors drove the level of improvement we needed to realize the IDM vision (Figure 6).

Figure 6: WCIP enablers

Tools (Technology)

The "to be" solution (Figure 7) offered the following benefits.

- Reduced system architecture complexity.
- Centralized procurement systems attached to the global Intel communications backbone.
- Improved and standardized data models.
- Proven data architecture in use globally.
- Near 100% match against requirements.
- Reduced IT operations management overhead.
- Leveraging of existing, off-the-shelf software.
- Scalability of core system components.
- Extensibility to other modules and third-party apps.
- Visibility of all data regulated through the data model.

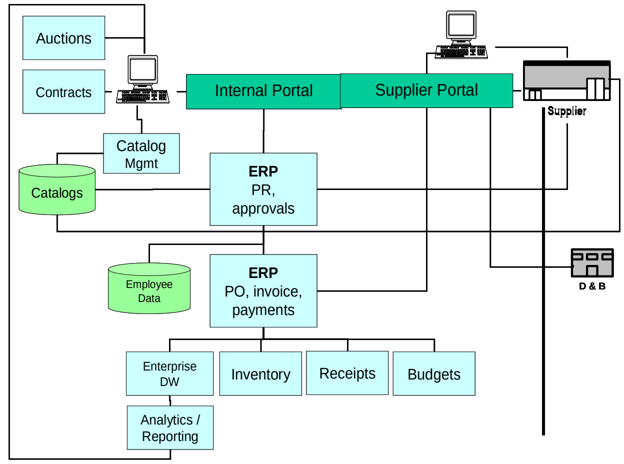


Figure 7: "To-Be" solution

Another key aspect of technology was to leverage our systems to support new processes. Key areas to highlight are global deployment, minimal system modification, limited interfaces, and parallel end-of-life programs for retiring systems.

Examples of key programs that were implemented were Internet negotiations, e-Catalog, and global procurement reporting.

e-Catalogs

Using e-Catalog, a requisitioner can directly access a supplier's list of items with minimal procurement involvement. This application improves the Throughput Time (TPT) and provides an excellent Graphical User Interface (GUI). Items that have high transaction volume can be described easily with or without pictures, or come directly off the shelf, such as office products, were a good fit for e-Catalog technology. During the first full year, thirty-nine catalogs in nine countries were deployed. Deployment of new catalogs will likely continue through mid-2008.

There are three primary benefits to the use of e-Catalogs: The first is the ability to get data by supplier, requisitioner, commodity, or by specific items purchased. The contracted purchase price is consistently guaranteed via the catalog. Aside from the contracted price, the catalog offers only items that have been contracted. Lastly, there is about a one-week reduction in processing time.

Internet Negotiations

The Materials organization was utilizing traditional negotiations with no technology that involved a significant amount of face-to-face contact, telephone calls, e-mail, faxes, etc. The traditional process is iterative and time intensive. The e-Procurement team, through benchmarking, found an application capable of performing negotiation activities on-line.

Internet Negotiations, an on-line negotiation capability, allows repetitive and real-time bidding by multiple suppliers, within a singular negotiation forum. This capability utilizes the Internet to reduce cycle time and provide a total cost evaluation for suppliers.

During the last three years, the Materials organization has been highly successful in using this tool in many of its commodity negotiations and has achieved an average of 10% in additional savings.

Global Procurement Reporting

"Spends visibility" was one of the key drivers for the e-Procurement program. The consolidation of business processes and tools aimed to create an environment where the data layer was integrated so a buyer and commodity manager had comprehensive visibility to spends data in the relevant and desired cuts. The next step was the creation of a data model that produces answers to the relevant business questions–this is the point where reporting technology steps in. A Business Intelligence (BI) layer with multi-dimensional On-Line Analytical Processing (OLAP) technology utilizing Enterprise Data Warehouse was identified as the required infrastructure.

Within the requirements gathering and design process it was identified that some of the dimensions are either non-existent or not in the correct structure. The critical ones were supplier, commodity, and contract (see Figure 8). A major effort was invested in creating the correct supplier structure (ability to identify 100 instances of Supplier A and a legal relationship among the suppliers). Similarly, a commodity hierarchy was created from scratch representing the "non-part numbered items" purchases. This commodity hierarchy is aligned with industry-standard coding.

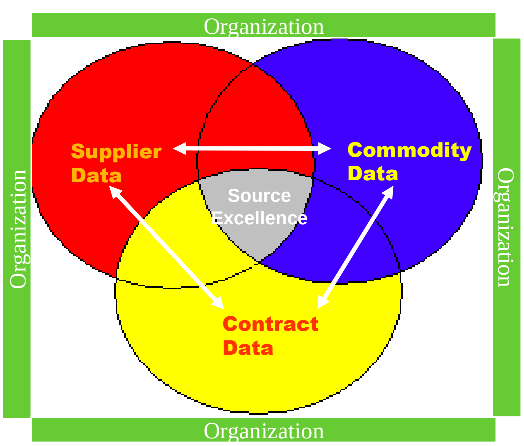


Figure 8: Data model

People

Several improvements were made in CP's People Systems. First, while the resources remained aligned organizationally to regional management, they became matrixed functionally to central teams. For example, a sourcing specialist for IT products in Europe still reported to European management; however, his or her functional responsibilities aligned to a global IT products sourcing team. In addition, a transformation was needed to transition from "procurement only" regional teams to global teams. The concept of a global source team (Figure 9) was introduced in mid-2003. Global source teams not only included IDM employees, but also key business partners and the finance organization as equal members on the team.

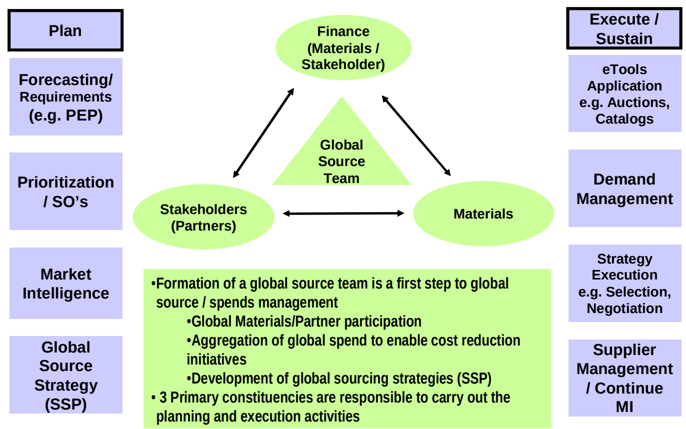


Figure 9: Global source teams

In addition to the team development, several new training modules in business partnering, market intelligence, sourcing plans, and diversity training were introduced for employees to enhance their skill set.

Process

A 5-step source process was introduced in 2003 (Figure 10) as a common global framework for IDM's sourcing and fulfillment professionals. This 5-step process provided a consistent framework for IDM employees worldwide to operate within. A standardized framework enabled a common approach and activities. Global source teams now had a process to follow that develops and enables optimal sourcing strategies for Intel.

Figure 10: 5-step source process

Another key process change was how materials and services were coded in our systems. A new coding scheme was introduced (called "Commodity Schema") and aligned with the industry standards. A common global coding process provided an opportunity for spending aggregation and data accuracy.

RESULTS

The e-Procurement program has been highly successful for Intel in the fields of technology, business processes, and people systems for indirect materials and services procurement. The program continues to deliver great savings and is expected to pass a cumulative savings of over $300M by the end of 2005. Figure 11 shows before and after progress.

BEFORE	AFTER
• 25% spending off system • 70% spends visibility • <50% supplier data accuracy • 65+ requisition and pay processes • Multiple legacy systems • Various sourcing processes • No standard commodity code information • Customer selects suppliers • Environmental scan limited in IDM space • Catalog usage limited and only in US • Traditional negotiations	•< 3% spending off system • 95% visibility by end of '05 • 80% of spends with supplier hierarchy • <5 requisition and pay processes • >97% spends on ERP • global procurement policy • strategic sourcing plans with stakeholders • Global 5-step source process • 500 commodity codes – 80% accuracy • Preferred supplier list implemented • > 30 eCatalogs globally • Internet negotiations used in >10% of spending

•Intel e-Procurement transformation delivered in 3 years, while the industry average is 5-7 years
•Intel has moved from lagging to a leading capability for process standardization, ERP integration and data

Figure 11: e-Procurement results summary

Solid progress has been made in technology deployment and enhancements. Over 95% of IDM spending is now channeled through a single global ERP system. A global reporting capability allows for a multi-dimensional view of IDM spending. Intel's IDM spending analysis progress was recognized by *Purchasing Magazine* in an article in December 2004. Intel's 60+ ways of purchasing methods and tools has been streamlined to five ways to buy. An Internet negotiations tool revolutionized supplier negotiations methodology. A catalog application is now in use in nine different countries.

The introduction of a 5-step source process helped the IDM organization drive a consistent, standardized framework for sourcing across the globe. The 5-step source process also enabled much larger engagement and influence with internal stakeholders than ever before. The success of this process was also recognized by *Purchasing Magazine* in its April 7, 2005 issue.

The IDM team has also evolved significantly. The organization was transformed from "regional procurement silos" to a global workforce. The organization formalized their relationship and their procurement expertise with the internal stakeholders through global source teams.

While the journey continues, it is estimated that the e-Procurement program delivered changes in the IDM procurement in three years, beating the industry average of five to seven years.

Finally, a couple of quotes from senior mangers at Intel: Ann Marie Kenitzer-Director, Requisition to Settlement Capability Management, ISTG says, "Intel's e-Procurement initiative has transformed indirect materials into a global, strategic purchasing capability in one half the time of industry benchmarks and has delivered over $300M in bottom-line savings to date. The implementation of standard global business processes, integrated and innovative information solutions, and emphasis on quality data is creating a new paradigm for global spend visibility and control of indirect materials that will deliver increased business value for Intel into the future."

Craig Brown, VP, TMG, Director-Materials, "e-Procurement program has delivered leadership results in spend management, sourcing excellence, stakeholder alignment, and controls. It has resulted in far more efficient indirect spending over prior years."

ACKNOWLEDGMENTS

We thank Eric Housh from Materials Communications for his significant help in making our thoughts more presentable. We also thank Craig Brown, Ann Marie Kenitzer, Doug Haughton, Travis Johnson, Chris Kloeppel, Robby Muller, Mike Millane, Mitchi Haight, and Judy Wente for their review of this paper and their contributions towards its content.

We also thank the ISTG/Materials e-Procurement team, without whose hard work, our vision would not have become a reality.

Above all, we acknowledge the tremendous drive and desire from the IDM employee base to improve current practices, enhance technology, and become a world-class indirect procurement organization.

REFERENCES

[1] Avery, S., "Intel goes global with indirect buying strategy," *Purchasing Magazine*, April 2005.

[2] Carter, P, S Beall et al., "Indirect Spend" *Critical Issues Report*, www.capsresearch.org*, September 2003.

[3] Minahan, T. A., "Global Sourcing: What You Need to Know to Make It Work," www.aberdeen.com*, August 2003.

[4] Mitchell, P. and R. Shah, "Indirect Procurement Field Review: Setting the Record Straight," *AMR Research, Inc.*, August 2001.

[5] Monczka, R. M. and R. J. Trent, "Cross-Functional Sourcing Team Effectiveness," Center for Advanced Purchasing Studies, *Focus Study*, 1993. www.capsresearch.org*.

[6] Verespej, M., "An elite club: Are you or your company one of the best in Spend Analysis?," *Purchasing Magazine*, December 2004.

[7] Yurday, E., "Business Value of Information Systems at Intel," *Case: OIT-40, Stanford Graduate School of Business*, May 2004.

AUTHORS' BIOGRAPHIES

Kinnar Ghiya has been with Intel for the last 8 years in a variety of supply-chain roles: materials engineer, sourcing specialist, purchasing manager, e-Procurement sourcing program manager, and in his latest role as a co-manager of Asia procurement operations. Outside Intel, Kinnar has worked in two other companies in supply-chain roles. He holds a B.S. degree in Electrical Engineering, an M.S. degree in Physics from Birla Institute of Technology and Science, India, an M.S. degree in Systems Engineering from the University of Arizona, Tucson, and an MBA in Technology Management from Arizona State University in Tempe. His e-mail is kinnar.k.ghiya at intel.com.

Marci Powers, CPIM, has been with Intel for 17 years in a number of different positions in both Materials and Production Planning. Her roles have included Fab Planning Manager and Build Plan Manager while in Production Planning. In Materials, her roles have included Site Purchasing Manager, Business Operations Manager, and various Program Managers. Marci received a B.S. degree in Business from Arizona State University in Tempe. She currently serves on the board of Junior Achievement of New Mexico and has held two other board positions including one on the Rio Grande Minority Purchasing Council. She was a graduate of Leadership Albuquerque and has volunteered for Youth Leadership Albuquerque. Her e-mail is marci.powers at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

Intel's Processes for Capacity Planning Optimization

Anwar Ali, Technology and Manufacturing Group, Intel Corporation
William J. Campbell, Technology and Manufacturing Group, Intel Corporation
Lance I. Solomon, Technology and Manufacturing Group, Intel Corporation
Megan A. Walsh, Technology and Manufacturing Group, Intel Corporation
James R. Wuerfel, Technology and Manufacturing Group, Intel Corporation

Index words: optimization, linear programming, planning, capacity

ABSTRACT

Customer demands for Intel® products are rarely consistent or predictable but must be fulfilled to the best of Intel's ability. Intel also regularly increases product differentiation and provides additional platform offerings. As a result, product mix, manufacturing equipment (or tool) requirements, and overall business processes at each of Intel's manufacturing plants and across the virtual factory are constantly being updated and adjusted. These practices dramatically impact how demand can be met and how capacity is utilized within both 200 mm and 300 mm Fab/Sort Manufacturing (FSM) and Assembly/Test Manufacturing (ATM).

There are a number of modeling challenges: working with an installed tool base while planning new purchases, the requirement to distribute volume requirements across sites and toolsets, and the ability to re-use tools across sites and between manufacturing processes. These constraints require interaction between multiple groups and separate capacity planning methods, and they have become increasingly difficult to manage. A more systematic and automated approach is called for.

Mathematical optimization models have been groundbreaking in their ability to gather key stakeholders around a repeatable approach. Not only have the optimization models been used to generate solutions to complex tools, they have also been used to foster collaboration between different business organizations at Intel. This has, in turn, greatly increased communication between stakeholder groups and reduced the cycle time required to produce business-ready solutions. The cost savings that resulted from using each of these tools individually as well as cumulatively has been dramatic. The use of these tools has reduced response time remarkably and aided in decisions resulting in over $1.5B of cost avoidance over the last five years.

In this paper, we reveal how an optimization approach provided powerful solutions within the FSM and ATM spaces both strategically and tactically. We also review each of the individual solutions and describe how they work together within Intel's virtual factory network.

INTRODUCTION

Intel's capacity planning process is done at different levels of detail for different time horizons. Decisions for equipment purchase or re-use are made based on target capacity, with protective capacity to support demand variability. Production planning is done for multi-year horizons split across the virtual factory network. Rough-cut capacity planning is done for a multi-month horizon using the split volume for each factory. Here the resource requirement planning is reviewed, and adjustments are made to the production plan, labor, collaterals, and material. Finally, production control is done for multi-week horizons.

The Fab/Sort Manufacturing (FSM) and Assembly/Test Manufacturing (ATM) Industrial Engineers (IEs) and Strategic Capacity Planning (SCP) co-own the multi-year capacity planning, with the IEs owning the capital purchase and re-use process. The ATM IEs also own the multi-month individual factory capacity planning, while manufacturing and planning owns production control.

Since each process has different data needs and business rules, we developed a family of applications for each one, as follows.

(1) Fab Routing Optimizer–performs product splitting and capacity balancing across Fabs.

(2) Sort Volume Optimizer–helps explore wafer sort opportunities to save additional testers.

(3) Capacity Roadmap Optimizer–performs equipped capacity split among multiple ATM factories to minimize tool purchases.

(4) Capacity Model Optimizer–performs detailed individual ATM factory capacity checks and optimizes tool allocation to product.

(5) Re-use Optimization–optimizes inter-factory (Fab, Sort, Assembly, and Test) capital equipment re-use and conversion opportunities.

Partnerships were created between Intel's Operational Decision Support Technology (ODST) group, the Re-use Teams, ATM Industrial Engineering, Strategic Capacity Planning, and TME Strategic Programs and Planning to aggressively address these issues by upgrading existing spreadsheet models to a more automated, faster, and more accurate suite of optimization models. The suite includes the Capital/Capacity Planning System that handles volume splitting, the Capacity Model Optimizer that performs site-based capacity checks, a Re-use Optimizer, an inter-factory capital equipment re-use opportunity optimization tool, the Fab Routings Optimizer that performs product splitting and capacity balancing across Fabs, the Sort Volume Optimizer that helps explore opportunities to save additional testers, and the Sort Allocation Tool that works on a shift level to meet outs and minimize setups. These models span Fab and Sort, Assembly and Test, on through equipment re-use.

FAB ROUTINGS OPTIMIZER

A major component of the periodic planning cycle is the routing of products in each process to individual wafer Fabs. These routings are important, since they impact toolset requirements. Strategic Capacity Planning (SCP) performs these routings with the cooperation of the individual factory capacity planning groups. There are a number of guidelines that need to be followed in performing the routings including under-loading and balancing virtual factory loadings. Prior to the implementation of the routings optimizer, these routings had been performed manually with the aid of ad-hoc spreadsheets without analytically comprehending the impact of product-mix implications on factory lithography capacity. Operational Decision Support Technology (ODST) partnered with SCP to produce an optimization model to semi-automate and aid in routings, comprehend product-mix, and to provide a consistent routings methodology used by all analysts and manufacturing processes. In the following sections, we discuss the routings problem, our approach to it, the distinguishing features of the optimization model, and our next steps.

The Routings Problem

Fab product routing is a manually intensive, quarterly activity that requires multiple revisions each cycle. This process requires the SCP Fab process coordinator to juggle yield vehicle direction, New Product Introduction (NPI) and product site alignments, and to finance driven constraints in a unit demand routing system while comparing over/under loads to wafer start capacities. SCP coordinators receive feedback and approval from the Fab planning/sort/yield managers, and then do their best to route products by percentages of the unit demand in response to the stated wafer start capacities.

Routing is a tedious, time-consuming process that requires uninterrupted concentration and constant checks by the analyst.

If photolithography limits are known for certain sites, routers try to accommodate this as well. These limits are commonly expressed as a maximum number of wafer starts, or percentage of product mix, that a factory can produce of a particular product. For example, in many technology nodes, factories have often needed to place limits on products that they could produce. Frequently, the initial routing results did not take into account the impact on lithography of the product mix capacity. Since these routings occur manually, time constraints do not allow this in-depth analysis to occur. The initial routing results are distributed to all involved for feedback within one week. Fab manufacturing engineers evaluated the impact of the routings on their specific sites by reviewing the Rev0 product mix and providing feedback. This feedback was subsequently incorporated into routings that drove further manual manipulation within the routing system to rebalance the Fab loadings while simultaneously maintaining all other constraints. This can be viewed as restarting the routing process. This routing process is completed for the final routings that are then used for the SCP coordinators quarterly publication.

Becoming proficient as an SCP analyst takes a long time, due to the many details and issues involved. With the ODST model the analyst is able to make changes based on feedback and to let the model incorporate them while rebalancing other products and maintaining routing constraints. The model provides a consistent solution strategy and allows analysts to complete their routing tasks more quickly.

Modeling Approach

We implemented a linear programming optimization model in two phases. In the first phase, we utilized the Frontline Systems Premium Solver Platform and Large-Scale LP Solver plug-ins to Excel*. This provided us with maximum flexibility in developing the optimization model. In this environment, we had rapid prototyping capabilities and were able to quickly experiment with different model formulations in order to best meet the SCP requirements. In the second phase, we converted the optimization model From the Premium Large-Scale LP Solver to an ILOG OPL Studio* [2] model and moved the model data from Excel to a SQL2000* database. The OPL/SQL solution was quicker and worked better with larger data sets than the model based on Excel. The model data are housed in a centralized planning database and are connected to other data sources for factory capacity information. Moving the data to SQL2000 provided enhanced data integrity and manipulation capabilities.

Model Features

There are a number of different conflicting objectives that are included in this model. Weights are added to balance the different units of measure and to indicate relative priorities. The model attempts to maximize the minimum loadings of all factories, "smooth" product loadings within a factory over time, and minimize the overloading/underloading of a site, based on input factory goal loadings.

There are a number of fixed constraints that must be met by the model. All product demand must be routed. Additionally, the planner may specify a minimum or maximum number of wafers that need to be routed to individual factories.

This optimization model uses the following input data: individual Fab process capacities, engineering requests, product demand, known factory routings on low-volume products, the previous quarter loadings, Fab overall loadings targets, and product mix information.

The model outputs include individual Fab percentage loadings and the product wafer starts and percentage loadings by Fab in each time period.

In addition to Fab routings, ODST and SCP have been investigating ATM routings. ATM and Fab routings have many differences, and they have become separate projects. One large difference is in how subcontracted routings are handled. One similarity, though, between ATM and Fab routings is in the separation of routings. While Fab routings are process specific, ATM routings are based on platform/package combinations. ATM routings are constructed as separate Assembly and Test models; the output of the Assembly model becomes part of the input to the Test model.

* Other brands and names are the property of their respective owners.

SORT VOLUME OPTIMIZER

As product and process breakthroughs increase the total number of memory bits per wafer on memory processes, the time to test these bits (even at a basic read/write level) grows at nearly the same rate. This requires more testers to support a given wafer start level than previous memory generations have required. The primary way to address this reality is to increase test parallelism (the number of die tested simultaneously). In the first 8" Flash Sort process, multiple die were tested in parallel. Our current testers can test a greater number of die in parallel. Additionally, testers may be able to test even larger numbers of die in parallel in the future. For each of these parallelisms there have been enabling tester/prober (cell) upgrades. In all, there are many unique cell/parallelism (platform) combinations. A single product may be eligible for multiple platforms at each of its three sort test operations (flows). The number of product/flow/platform combinations is practically endless. Given that the cycle time to enable a product on a new platform is fairly long, poor product allocation decisions could waste significant tester resources. It is critical to find product to platform allocations (for each of the multiple test steps), tool purchases (testers and probe cards), and a product conversion plan that ensures maximum sort capability at the minimum cost (capital, expense, and labor). Complicating matters further, these decisions must be made for all factories as well as for each individual site simultaneously.

Sort Industrial Engineers (IEs) from all sites would periodically meet and attempt to solve these issues. However, the amount of data and the possible solution sets were too complex to continue to solve manually.

Methodology and System Architecture

The tool itself, the Sort Volume Optimizer (SVO), is comprised of a mixed integer optimization model developed in OPL Studio. The Mixed Integer Program (MIP) is similar to a traditional Linear Program (LP), but contains decisions that must be integer values, adding complexity and, therefore, solve time to the solution space. The SVO was built using Microsoft Access* as the front-end Graphical User Interface (GUI). Visual Basic

* Other brands and names are the property of their respective owners.

for Applications (VBA*) codes in Access and Excel perform data transformation and SQL operations, importing flat files from the capacity model to the SVO database and preparing data for MIP solve. This architecture is shown in Figure 1.

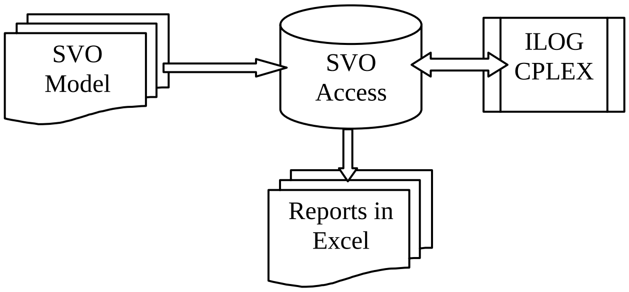


Figure 1: SVO architecture

The Access VBA code also generates data fallout reports to assist end users in debugging data. The VBA code then loads and executes the ILOG OPL Studio compiled model using component object model objects. The optimization results are then written back to the database. Finally, the Product Allocation report, Tool Purchase and Transfer reports, and other detailed reports are created in Excel (see Figure 2).

Model Formulation

The optimization model is comprised of five sets of main decision variables to determine the optimal product platform allocation roadmap, tester and probe card purchase and transfer schedules, product to platform conversion (qualification) schedules, and product cross shipping requirements/schedules. The model can also determine when it is more cost effective to miss volume than it would be to perform product platform conversions and/or to make additional tester and probe card purchases. The decisions are made in order to optimally meet the constraints and rules of the system while minimizing total cost. There are approximately 50 global constraints that must be followed in the decision-making process.

The model requires a great deal of input data in order to be able to optimally perform its decision-making routine. The Sort IEs must provide the cost and penalty information that is required in the objective function, i.e., information regarding the test operations such as run rates, utilization, tool limitations, and site space constraints. In addition, starting inventory and demand information, conversion details, and any resource and site limitations are also required inputs.

* Other brands and names are the property of their respective owners.

Advantages

The new SVO models capacity and capability and solves for the strategic product-to-platform development planning in this complex environment more thoroughly and optimally than previous manual methods allowed. As a result, the SVO has enabled new business processes (such as tool re-use strategies) and has the primary benefit of reducing the decision-making process by six weeks each quarter, thereby enabling a greater than 17% reduction in product-platform development throughput while enabling full utilization of legacy sort test platforms.

CAPACITY ROADMAP OPTIMIZERS

SCP and ATM IEs develop the CPU capacity roadmap over both the short-term and long-term horizons. Previously, allocating product capacity between multiple factories was a very manual and time-intensive process that produced sub-optimal results due to partial information at each stage. SCP generated a roadmap, IEs responded with major tool and space constraints for a revision, and finance checked revenue concentration at the end. The net result was a multi-week turnaround to create a joint SCP/IE roadmap that contributed to capital estimations with relatively inaccurate data. The need for a faster, more agile roadmap and accurate capital estimation process required an overhaul of the complete system. This work is similar in concept to that of Berman and Hood [1].

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	Product	Site	Tester	Parallelisr	Flow	1	2	3	4	5	6	7	8	9	10	11
2	1316 Eng_	SF11	1316	x16	Sort 1	733	733	733	508	508	508	508	503	503	503	
3	1316 Eng_	SF11	1316	x16	Sort 2		733	733	733	508	508	508	508	503	503	
4	1316 Eng_	SF11	1316	x16	Sort 2a	733	733	733	508	508	508	508	503	503	503	
5	1316 Eng_	SF14	1316	x16	Sort 1	73	73	192	192	192	192	192	192	192	192	
6	1316 Eng_	SF14	1316	x16	Sort 2		73	73	192	192	192	192	192	192	192	
7	1316 Eng_	SF14	1316	x16	Sort 2a	73	73	192	192	192	192	192	192	192	192	
8	1316 Eng_	SF23	1316	x16	Sort 1	14	14	14	14	14	14	14	14	14	14	
9	1316 Eng_	SF23	1316	x16	Sort 2		14	14	14	14	14	14	14	14	14	
10	1316 Eng_	SF23	1316	x16	Sort 2a	14	14	14	14	14	14	14	14	14	14	
11	4400 Eng_	SF11	4400	x36	Sort 1	501	501	501	485	485	485.51	485	481	481.51	481	
12	4400 Eng_	SF11	4400	x36	Sort 2		501	501	501	485	485	485.51	485	481	481.51	
13	4400 Eng_	SF11	4400	x36	Sort 2a	501	501	501	485	485	485.51	485	481	481.51	481	
14	4400 Eng_	SF14	4400	x36	Sort 1	232.49	232.51	232.51	232.51	232.51	232	232.51	196.51	196	196.51	
15	4400 Eng_	SF14	4400	x36	Sort 2		232.49	232.51	232.51	232.51	232.51	232	232.51	196.51	196	
16	4400 Eng_	SF14	4400	x36	Sort 2a	232.49	232.51	232.51	232.51	232.51	232	232.51	196.51	196	196.51	
17	4400 Eng_	SF18	4400	x36	Sort 1	149.51	149.49	149.49	149.49	149.49	149.49	149.49	149.49	149.49	149.49	
18	4400 Eng_	SF18	4400	x36	Sort 2		149.51	149.49	149.49	149.49	149.49	149.49	149.49	149.49	149.49	
19	4400 Eng_	SF18	4400	x36	Sort 2a	149.51	149.49	149.49	149.49	149.49	149.49	149.49	149.49	149.49	149.49	
20	4400 Eng_	SFD2	4400	x36	Sort 1	1212	1212	1212	1137	1137	1137	1137	1137	1137	1137	
21	4400 Eng_	SFD2	4400	x36	Sort 2		1212	1212	1212	1137	1137	1137	1137	1137	1137	
22	4400 Eng_	SFD2	4400	x36	Sort 2a	1212	1212	1212	1137	1137	1137	1137	1137	1137	1137	
23	ARMAGO:	SF11	1316	x16	Sort 1			36	282.92	266.78	290.97	485.14	485.08	485.08	460.44	
24	ARMAGO:	SF11	1316	x16	Sort 2				36	282.92	266.78	290.97	485.14	485.08	485.08	
25	ARMAGO:	SF11	1316	x16	Sort 2a			36	282.92	266.78	290.97	485.14	485.08	485.08	460.44	
26	ARMAGO:	SF11	1316	x16	Sort 1	40										
27	ARMAGO:	SF11	1316	x16	Sort 2		40									
28	ARMAGO:	SF11	1316	x16	Sort 2a	40										
29	ARMAGO:	SF14	1316	x16	Sort 1		80	80	130	130	130	130	130	130	130	
30	ARMAGO:	SF14	1316	x16	Sort 2			80	80	130	130	130	130	130	130	
31	ARMAGO:	SF14	1316	x16	Sort 2a		80	80	130	130	130	130	130	130	130	
32	CRYSTAL	SF11	4400	x36	Sort 1	62.39	16.21	117.66	160.13	132.1	131.6	139.74	159.63	150.63	141.13	
33	CRYSTAL	SF11	4400	x36	Sort 2		62.39	28.37	180.15	233.42	203.24	203.17	217.55	230.67	223.12	
34	CRYSTAL	SF11	4400	x36	Sort 2a	62.39	24.32	159.32	208.99	179.52	179.31	191.61	206.99	198.96	192.77	
35	CRYSTAL	SFD2	4400	x36	Sort 1	62.39	16.21	83.33	97.73	94.85	95.43	103.74	94.71	96.65	103.27	

ProductAllocation / PlatformReqt / TransTo / TransFrom / TotalTesterReqt / ProdPlatRe

Ready NUM

Figure 2: Product platform allocation report in Excel

Integrated Approach

To achieve this, SCP, ATM IE, and ODST partnered to develop the CPU Capacity Roadmap Optimizer (CRO). This mathematical optimization model integrates the key rules and information to produce a solution that addresses all of the critical requirements during the first pass. With CRO, capacity roadmap allocation is now a joint effort with combined information that generates better roadmaps faster. These roadmaps adhere to key constraints for space, revenue, factory capability, and product and site ramp guidelines with conflicts clearly visible. For direct dollar savings, CRO looks to minimize key tool purchases with re-use up front (versus waiting for IE response) as long as other criteria are satisfied.

The business cycle for updating the roadmap model is a constantly evolving understanding of the inputs and outputs of the process, where each cycle builds on the previous one, and the impact of adjustments is quickly understood. The other key aspect is the feedback loop incorporated into the process. In the initial stages of entering the data, data quality is usually an issue. Whether it be a routine data entry error or significant change in an input that has impacted the solve outcome, the business user needs to be able to identify and resolve data problems quickly. The CRO provides multiple data validation checks to catch routine errors. Also, maximum solve time is an input parameter.

Business Rules (Constraints) and Criteria

The CRO considers many business rules in its optimization: which factories can run each product, factory space utilization, revenue targets, factory capacity goals, worker headcount, and new tool purchases.

The Architecture of the CRO

The CRO captures business rules in an MIP developed in ILOG OPL Studio that runs on the end-users' computers (in three countries) linked to a common SQL database. The CRO architecture enables ease of use and understanding by exporting both input and output data to Excel, displaying objective function components (penalties) to highlight constraints by factory and quarter, running pre-solve data error checks, and managing multiple scenarios.

Figure 3 shows, at a high level, the inputs and outputs of the model and what business group is responsible for maintaining those data.

Key Features

The CRO is adaptable to a changing business environment because of its user-defined penalties, ability to enable smoothing and rounding constraints, and its ability to make tradeoffs between runtime and business rules. The architecture has proven to be easy to maintain and is extendable to advancing business requirements. The user can smooth to a previous roadmap to reduce churn between cycles or to calculate the penalties on a manual roadmap. The ability to "what-if" provides a richer understanding of the relative impact of such things as space, tool costs, and revenue. The CRO provides a graph of high-level roadmap results to speed up understanding for both users and management.

With these features in place, a joint business user and technical team analysis was conducted to determine the correct penalty levels. This was done with a simple run of the model looking at the percentage each penalty consumed. By having the percentages, the joint team was able to dial the individual penalty values up and down to determine the correct level, based on business drivers and technical accuracy. This process is repeated periodically to ensure that the levels still accurately reflect current business priorities.

Since the detailed information about tool level consumption per product is available, the IEs are able to verify and validate the numbers with manual calculations. In cases where the model and the manual calculations do not match up, continuous improvement is possible.

Benefits

The CRO provides "better" roadmaps by consistent enforcement of constraints, and by eliminating sub-optimization with partial information. It also provides a build plan to plan continuity for major product transitions and factory ramps. Roadmaps can now be created for SCP/IE in 40% less time, supporting capital procurement requirements quickly. With this new tool, a richer set of "what-ifs" can be considered in the same time that one analysis was done previously. The quality of information has greatly improved through integrated information and models: the data on tool productivity metrics and space utilization are checked for quality earlier in the process. Since CRO supplies constraints and management summary data, users can visualize solve quality, and management can get more complete data upfront to facilitate explicit management discussion of constraints and tradeoff options. The net dollar impact of better data, integrated processes, and additional "what-ifs" is estimated at ~$13M, and these also identify opportunities to save multiple testers in the mid-range time horizon.

As a result the business units have a more effective business process by employing a joint SCP/IE roadmap early in the process, allowing the IE's analysis to focus on clean roadmaps with no glaring space/capital gaps and ensuring consistent rules for each version; in other words, each cycle builds on previous rules. This translates into increased productivity. SCP and IE work-hours per quarter were reduced by 11%. Moreover, the workload has shifted from overloaded site resources to ATM IE/SCP for data-population and solves.

RE-USE OPTIMIZER: OPTIMIZING SUPPLIER DOCK DATES AND TOOL RE-USE

Optimizing the delivery of Fab capital to meet process priorities is a critical aspect of meeting the production ramp. A key component in keeping wafer costs down is the re-use and conversion of existing capital equipment to meet the needs of new generations of processes across a worldwide network of manufacturing facilities. There are a large number of variables when making allocation decisions: release and required dock dates, supplier dock dates for new tool purchases, conversion costs, conversion time, new purchase prices, tool handedness and draft requirements, grant and lease considerations, transportation costs, etc. Allocation is not only complex, it is large (involving thousands of tools worth several billion dollars).

Originally, there was a manual process that attempted to facilitate equipment re-use periodically across the different factories at Intel. This manual process has now been replaced with the Re-use Optimizer.

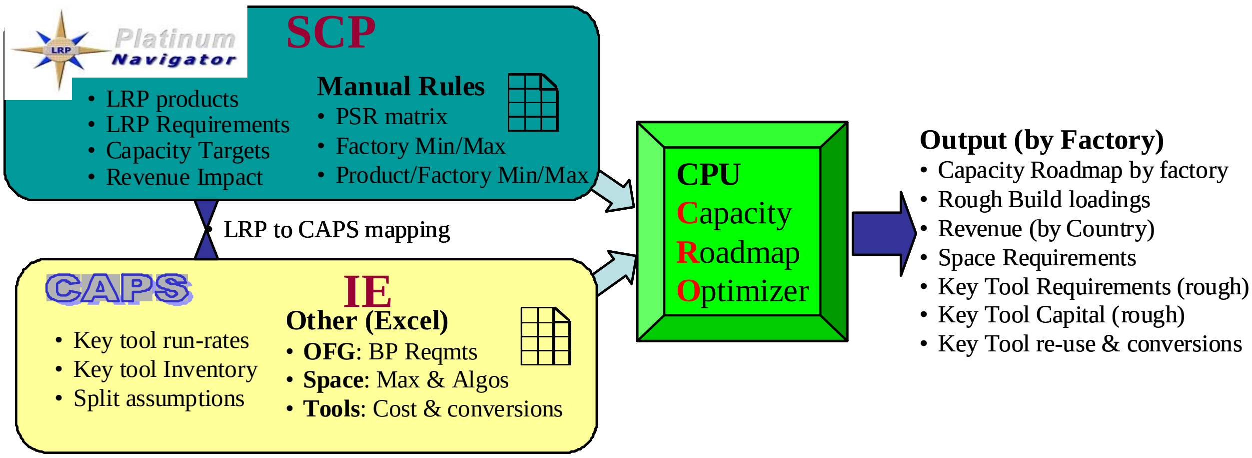


Figure 3: Capacity roadmap inputs/outputs

Re-use Optimization Goals

The Re-use and Allocation team partnered with ODST to change from the manual methods to an automated system using optimization to simultaneously minimize the impacts of supplier tool lateness and to maximize re-use tool opportunities. This optimization model automates the current manual process of allocating supplier dock dates and re-use allocations in a single step.

Methodology

The optimization methodology can be represented by an LP. The model can be described using an assignment problem, a special class of LP. Many other types of model formulations can be found in Winston [3]. Assignment problems can be characterized by a cost matrix composed of the cost of assigning each supply point to each demand point. Equation 1 displays the general form of the balanced assignment linear programming problem. The first line of the equation indicates that the objective function is to minimize the total cost of the assignments. The cost of assigning the i^{th} supply point to the j^{th} demand is denoted by c_{ij}. The decision variables (assignments) are denoted as x_{ij}. The second and third lines are the constraints on the decision variables, requiring that the total supply and demand requirements be met. Lastly, the fourth line indicates that no fractional allocations are allowed. That is, each supply must fill exactly one demand, and that each demand is filled by exactly one supply.

$$\min \sum_{i=1}^{i=m} \sum_{j=1}^{i=j=n} c_{ij} x_{ij}$$

$$\text{s.t.} \sum_{j=1}^{j=n} x_{ij} = 1 \; (i = 1,2,\ldots,m) \quad \text{(Supply Constraints)}$$

$$\sum_{i=1}^{i=m} x_{ij} = 1 \; (j = 1,2,\ldots,n) \quad \text{(Demand Constraints)}$$

$$x_{ij} = 0 \text{ or } x_{ij} = 1 \; (i = 1,2,\ldots m;\; j = 1,2,\ldots n)$$

Equation 1: LP representation of assignment problem

In re-use, supply points are synonymous with existing supplier dock dates or excess tools; demand points are forecast tools, and the costs are a function of the lateness of a supplier dock date allocation and the cost of re-using and converting a tool. The constraints in this system are that each excess tool must be assigned to no more than one required tool, and that each required tool must be allocated exactly one excess tool or supplier delivery slot for a new tool purchase. It is possible, of course, to place more than one tool on hold and to also purchase more than one tool.

Model Specifications

In designing a solution to the complex task of performing the allocation of all Fab capital equipment including re-use opportunities, we use the following inputs to the model:

- Capital equipment forecast quantities, excess tool lists, and equipment costs.
- Tool conversion costs and re-use kit lead-times.
- Factory customer-determined process priorities.
- Special tool circumstances and requirements.

We have made the following assumptions in the determination of the Re-use Optimization cost matrix:

- Non-lithography forecast tools must be outside of new purchase lead-time plus a fixed duration to accept re-use tools.
- Demo time on factory installed re-use tools.
- Re-use tool shipping duration.
- Existing forecast tool allocations within a specified allocation start-date do not change.
- Current allocation lateness less than or equal to a user-specified parameter is ignored.
- Current allocations are given a small weighting factor to reduce "churn"–large changes in factory allocations for small gains.
- Excess purchase orders are given priority allocation in order to minimize cancellation costs (even if cancellation cost is zero).
- Additional rules depending upon geographical locations where the Fabs operate.

SuperSTARS* has been built using the MS SQL 2000 database and Microsoft Visual Basic 6* to generate cost matrix, the ILOG OPL Studio algebraic modeling language, CPLEX* for the optimization engine, and Excel for formatting and providing a user interface to the automated allocation.

Benefits

To date, over $1.5B worth of re-use has been managed through the system. The Re-use Optimizer has been able to effectively manage more detail than human planners could: the people using the system bring their expert knowledge about information that cannot be stored in the current system. The Optimizer presents a first-cut at the allocations, and the users are allowed to override and change allocations as needed. This process has allowed the re-use planning cycle to occur in a few days by using computer networking tools. In the past, the re-use process took multiple weeks of face-to-face meetings, with individuals traveling from around the world to manage this process.

CAPACITY MODEL OPTIMIZER

Periodically a new Build Plan (BP) request is received. Since ATM consists of multiple factories worldwide, the demand is split between various factories by using high-level capacity analysis. To ensure the BP is supportable, a more detailed analysis is performed for each factory (Figure 4) before adjustments are made for an unsupportable BP.

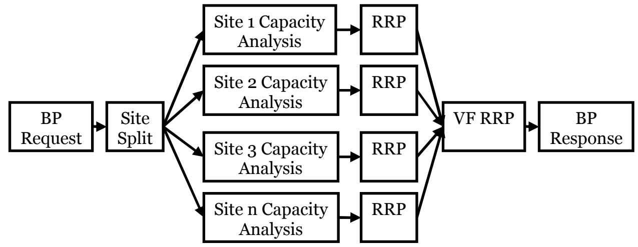


Figure 4: Weekly Build Plan process

The detailed capacity checks are performed for all steps in each factory, including steps with alternate tools. The route and required tool at each step is product specific; therefore, optimizing tool allocation to the right product is important to maximizing capacity. We applied a LP to replace the iterative manual tool allocation process done previously.

Methodology

ATM IEs use a metric model, based on Microsoft Excel, to calculate aggregated capacity for each tool when running a specific product. The capacity per tool is expressed as a runrate in thousands of units per week. The BP is also expressed in the same units, thus all capacity analysis is performed in weekly buckets for a multi-week time horizon.

The total requirement for tool *t* is defined as

$$T_t = USD_t + NSM_t + TR_t$$

Equation 2: Tool requirements

where USD_t and NSM_t are the Unique Scheduled Downtime and Non-Standard Material tool requirement for tool *t*, respectively, and

$$TR_t = \sum_{p,s} \frac{BP_p * Alloc_{pts} / 100}{RR_{pts} * Yield_p / 100}$$

Equation 3: Revenue requirements

is the revenue requirement for tool *t*. BP_p=Build Plan volume for product *p*, $Alloc_{pts}$=Allocation for product *p*, segment *s* and tool *t*, RR_{pts}=Runrate for product *p*, segment *s* and tool *t*, and $Yield_p$=yield for product *p*.

Total tool requirement, T_t, must not exceed the available tool inventory, I_t.

$$T_t < I_t$$

Equation 4: Tool inventory constraints

When the BP has been met and excess tool capacity is available, the unused tools are expressed as available

---

* Other brands and names are the property of their respective owners.

upside capacity for each step. Excess capacity allocation is based on product priority. Low-priority products will be allocated capacity to meet BP only, medium-priority products to installed capacity, and high-priority products to meet or exceed installed capacity. When multiple products have similar priorities, the excess capacity ratio is made equal among those products.

The supportable capacity for each product is determined by selecting the lowest step capacity. When IEs present the capacity statement to the resource requirement planning forum, the most impacting limiting steps are shown. From this, the forum members have a good idea of what they need to focus on, if more upsides are required.

Model Formulation

Since the problem has multiple conflicting objectives, goal programming is used. Slack variables are defined for each goal and penalized according to the importance of the goal. Initially a prototype LP model was developed and demonstrated to the customers using a solver for Microsoft Excel. The model solves one time period for a limited number of products and steps. The penalty values are tuned for various scenarios until the model meets customers' requirements. The model is then migrated to ILOG OPL Studio and expanded to comprehend all products and steps in a factory.

Optimization Model Architecture

CMO was built using Microsoft Access as the front-end user interface. Visual Basic for Application codes in Access and Excel perform data transformation and SQL operations, importing flat files from the capacity model to the CMO database and preparing data for LP solve (Figure 5).

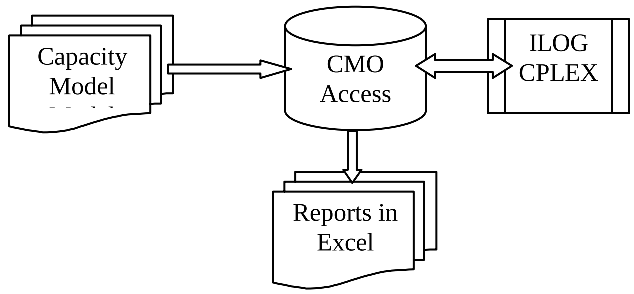


Figure 5: CMO architecture

The Access VBA code also generates data fallout reports to assist the end users in debugging their data. The VBA code then loads and executes the ILOG CPLEX compiled model using COM object. The CPLEX LP results are then written back to the database. Finally, the Limiter Chart and other detailed reports are created in Excel (Figure 6).

CM1 Optimizer Version 1.35 (July 30th, 2003)

FC, FC-BGA2, TUALATIN, MBL, 512, KB, C (FE)

Workweek	**WW2103**	**May'03(Par**	**WW2203**	**WW2303**	**WW2403**	**WW2503**	**WW2603**	**Jun'03**	**Q2'03(Part**
Nominal WW	168.00	168.00	168.00	168.00	144.00	156.00	168.00	804.00	972.00
Target	19.84	19.84	5.62	5.62	4.82	5.22	5.62	26.90	46.74
BP	15.45	15.45	1.74	1.74	1.50	1.62	1.74	8.35	23.80
Limiter 1	EPOXY		TEST	TEST	EPOXY	TEST	DIE PLATE		
Limiter 1 Capacity	20.60	20.60	6.98	7.00	5.96	6.52	7.09	33.56	54.15
Delta	5.15		5.23	5.26	4.47	4.90	5.35		
% Burst	33.30%		299.79%	301.23%	298.73%	302.37%	306.56%		
Limiter 2	CTL		EPOXY	EPOXY	DIE PLATE	EPOXY	CTL		
Limiter 2 Capacity	22.08	22.08	6.98	7.00	6.08	6.52	7.46	34.03	56.11
Delta	6.63		5.23	5.26	4.58	4.90	5.71		
% Burst	42.94%		299.79%	301.23%	306.37%	302.37%	327.24%		
Limiter 3	APL		DIE PLATE	DIE PLATE	CTL	DIE PLATE	EPOXY		
Limiter 3 Capacity	22.62	22.62	7.06	7.11	6.39	6.62	7.55	34.72	57.34
Delta	7.17		5.31	5.36	4.89	5.00	5.80		
% Burst	46.38%		304.55%	307.37%	326.96%	308.52%	332.52%		
Limiter 4	DIE PLATE		CTL	CTL	TEST	CTL	APL		
Limiter 4 Capacity	22.90	22.90	7.45	7.48	6.42	6.96	7.63	35.95	58.85
Delta	7.45		5.71	5.73	4.93	5.34	5.89		
% Burst	48.21%		327.02%	328.42%	329.30%	329.81%	337.45%		
Limiter 5	TEST		APL	APL	APL	APL	CURE		
Limiter 5 Capacity	24.00	24.00	7.63	7.65	6.54	7.14	9.13	38.09	62.09
Delta	8.55		5.89	5.91	5.05	5.52	7.38		
% Burst	55.35%		337.30%	338.41%	337.54%	340.66%	423.01%		

Figure 6: Capacity statement showing the top five limiters

Results

CMO reduced the time needed to prepare capacity statements by 25% while greatly improving the quality and consistency of the process. This comes from reducing the number of iterations required to determine the allocation of tools to products. With CMO, the optimal values are determined automatically. CMO also converts complex business rules into mathematical models and ensures all factories are using consistent methods to declare their capacity statement. The data fallout reports enforce data quality checks in the IE capacity model, improving data quality for use by other capacity solvers.

SUMMARY OF RESULTS

We have shown how this small subset of optimization techniques has enabled cross-site and cross-organizational teams to produce better solutions in a smaller amount of time. The models not only help to automate these different decision-making processes, but provide people common methodologies to collaborate and discuss different solutions and to produce the best results for Intel. The optimization team has also produced many other optimization tools to help in other areas such as wafer purchasing and individual tool-level improvement models. These models have saved Intel a great deal in capital cost avoidance over the past five years, and they have also reduced the time it takes to produce solutions to these difficult problems.

ACKNOWLEDGMENTS

We thank Devadas Pillai and Edward J. Yellig for their vision and belief in optimization modeling. They have allowed and encouraged us to take the risks necessary to deploy new technologies in new areas. We especially thank our peer reviewers. We also thank the many people who worked in making these different projects successful:

Fab Routings Optimizer: Kim Johnson, Jennifer McNeill, Jeff Zino, and Gina Gerave.

Sort Volume Optimizer: Jake Aranda, Roger Pharr, Kosta Rozhenko, and Janna Smith.

CRO: Marcelo Castro, Roberta Roberts, Arlene Say, and Ed Zawacki.

Reuse Optimizer: Iris Rivera, Siroos Sokhan-Sanj, Beth Adkison, Barbara Haney, Leo Monford, Matt Walters, Bill Baldinger, Kalpesh Shah, Viswas Nataraj, and Neeta Pathrabe.

Capacity Model Optimizer: Gizelle Constantino, Chris Birely, Yeoh Hooi Yeam, Tan Teong Wah, and Contessa Preclaro.

REFERENCES

[1] Bermon, S., and Hood, S. J., “Capacity Optimization Planning System (CAPS),” *Interfaces*, vol. 29, pp. 31-50, September-October 1999.

[2] *ILOG OPL Studio User’s Manual*, ILOG, France.

[3] Winston, W., *Operations Research: Applications and Algorithms, Second Edition*, PWS-KENT, Boston, (1991).

AUTHORS’ BIOGRAPHIES

Anwar Ali is a staff engineer in ODST Malaysia. He leads a team of O.R. practitioners in applying discrete event simulation and mathematical modeling techniques to improve manufacturing productivity. Anwar also collaborates with Malaysia universities in initiatives to apply science to manufacturing. He holds a Masters degree in Decision Science. His e-mail is anwar.ali at intel.com.

William Campbell is a senior manufacturing operations modeling engineer. Bill is currently involved in optimization and simulation primarily in the capacity planning and scheduling space. Bill holds a Masters degree in Industrial Engineering from Purdue University and an MBA from Arizona State University. His e-mail is william.j.campbell at intel.com.

Lance Solomon is the Assembly/Test program coordinator for ODST. Lance’s current focus areas include full factory simulation, data integration, and mathematical optimization supporting capacity planning within Assembly/Test factories. Lance holds a B.S. degree in Mathematics from the Pennsylvania State University and an M.S. degree in Industrial Engineering/Operations Research from the University of Texas at Austin. His e-mail is lance.i.solomon at intel.com.

Megan Walsh is a senior manufacturing operations modeling engineer and manager in ODST. She received her Master’s degree from Arizona State University in the area of Operations Research in 2000 with a focus on Stochastic Linear Programming. She has been working at Intel in the areas of optimization and simulation since then. Areas of interest include product allocation assignment, quality and reliability modeling, and mathematical optimization to support high-volume manufacturing. Her e-mail is megan.a.walsh at intel.com.

James Wuerfel is an Intel technologist and ODST’s optimization program coordinator. Jim’s current modeling activities include mathematical optimization and scheduling. Areas of interest include equipment re-use, tool-level optimization modeling, and lot-to-tool scheduling. Jim holds a Masters degree in Industrial Engineering from Purdue University and is currently

pursuing a Ph.D. degree at Arizona State University. His e-mail is james.r.wuerfel at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

Optimizing Supply-Chain Planning

John W. Bean, Technology and Manufacturing Group, Intel Corporation
Amit Devpura, Technology and Manufacturing Group, Intel Corporation
Michael O'Brien, Technology and Manufacturing Group, Intel Corporation
Shamin Shirodkar, Technology and Manufacturing Group, Intel Corporation

Index words: supply chain, optimization, linear programming, production planning

ABSTRACT

Semiconductor manufacturing is a very capital-intensive endeavor that can return substantial revenues. The production planning process must deliver a build schedule that makes efficient use of Intel's capital resources while satisfying as much demand as possible. This schedule should comprehend the flexibility of production resources, the dynamic nature of supply and demand within Intel's supply chain, as well as the timing of new product releases and production facility improvements.

Previous planning processes relied on spreadsheets for heuristic manual decision making with localized data. With the growing complexity of Intel's products and manufacturing processes, these methods had become inadequate and unsustainable. Upgrading the planning process required better decision algorithms, improved data management, as well as more automated and integrated planning processes.

New tools based on Mathematical Programming were implemented in multiple divisions and stages of Intel's supply chain. The development team worked closely with the users to understand their business and capture their operating logic to create automated decision systems. These tools balance requirements to satisfy demand, achieve inventory targets, and remain within production capacity to reduce costs and satisfy demand across Intel's supply chain. They have been developed to evolve the planning process while maintaining visibility to the logic and data flow to facilitate continuous improvement.

Advances in data management were required to complement decision algorithm improvements. The new tools integrate directly into source data systems while providing planning- and optimization-specific functionality, including mechanisms to track parameter changes and supply dynamic reporting capabilities. These advances allow planners to more easily identify data issues and to better understand the planning recommendations from the tools. The robust data management infrastructure enables tighter integration of organizations, increased scalability, and more consistent implementation of solutions across business units.

Advances in decision algorithms, data management, and system automation led to improvements in solution quality, data health, and productivity. The new applications allow planners to rapidly perform analyses on multiple business scenarios to produce better solutions and improve collaboration with other organizations. While results reported by the business users over the past four years have proven the stability and value of this decision support technology, there is still work to be done. Plans for extensions and continuous improvement are provided in the last section of this paper.

INTRODUCTION

Semiconductor manufacturing typically proceeds through three major manufacturing stages as shown in Figure 1.

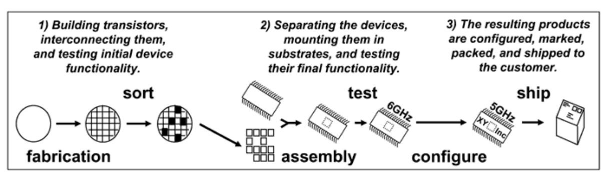


Figure 1: The basic flow in semiconductor manufacturing

First, hundreds of complex devices, each containing millions of transistors, are fabricated on silicon wafers. End-of-line testing sorts functional devices from ones that have manufacturing defects. Second, the wafers are sawn to yield individual devices (called die) that are assembled with substrates that supply physical protection and electrical connectivity. Final testing establishes the detailed performance characteristics of the semi-finished product. Third, after configuration and marking, the final product is packed and shipped to the customer. Basic planning for these stages includes deciding the timing and quantity of wafer releases into fabrication/sort facilities (F/S), die and substrate releases into assembly/test factories (A/T), and semi-finished goods releases into configure/pack facilities (C/P), as well as assuring the availability of substrates to support the A/T plan. These planning decisions are made more complex by Intel's risk management method of having multiple F/S, A/T, and C/P facilities. Each facility manufactures multiple products, and each product is produced in multiple factories.

Product differentiation must be comprehended in the planning process. As shown in Figure 2, there is a stochastic but measurable distribution of performance among the die that exit F/S. This information along with demand is used to decide the appropriate substrates for specific die. Consider microprocessors for example. The higher speed die should normally be placed in server substrates, the die that consume the least power in laptop substrates, and some of both in desktop substrates. Products exiting the A/T process also exhibit a stochastic but measurable distribution of final speed performance. Faster and slower semi-finished products are segregated. The planning system needs to consider these distributions.

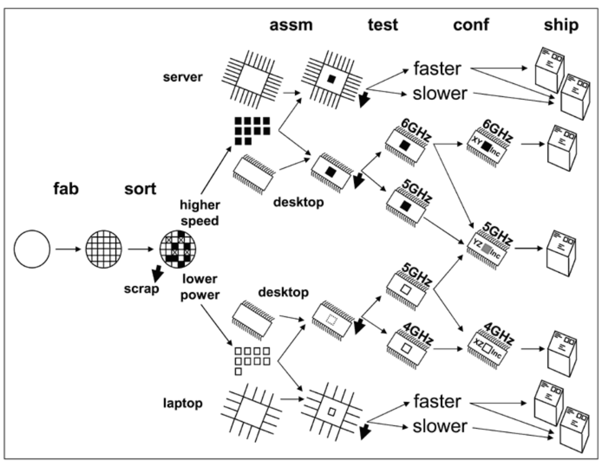


Figure 2: Product differentiation in the production flow

Additional planning decisions must be made for the configuration step as illustrated for desktop products in Figure 2. One feature of semiconductor production is that faster semi-finished products can be configured to run more slowly (6 GHz to 5 GHz and 5 GHz to 4 GHz in Figure 2) depending on demand. Unfortunately, slower products cannot be made to run faster. Furthermore, there are multiple flows that make the same product. Notice in Figure 2 that the slowest of the fast die and the fastest of the low-power die that were put into desktop packages converge to both make the 5 GHz final product. The planning system must comprehend that these two phenomena combine for desktop products to provide many ways to make the 4 GHz and 5 GHz products. Server and laptop products exhibit similar behavior.

These planning complexities are encountered across Intel's microprocessor, memory, and communications businesses, and over strategic and tactical planning problems. The business problem faced by all is to minimize costs and maximize revenues now and in the future. A number of factors make this a very difficult problem for Intel. On the supply side, there is considerable variability in manufacturing throughput time (TPT), line yields (the proportion of product that survives the manufacturing process), and product performance (of the surviving product, there is a distribution of operating speed, for example). There is also considerable variability on the demand side with customers placing orders that they later alter in quantity, due date, and/or delivery location, or cancel altogether. Manufacturing TPTs are much longer than the time it takes for our customers to change their minds and their orders. Product differentiation, complex and interrelated Bill-Of-Materials (BOMs), and shared capacity among product lines make planning even more complex.

Poor plans can waste valuable manufacturing capacity by either letting it sit idle or by using it to make products that the market decides it doesn't want. A slow planning system can eventually produce a plan that, by the time it is issued to manufacturing, is based on stale data making it unexecuteable or sub-optimal. The dual goals of minimizing costs and maximizing demand satisfaction depend on making the right volumes of the right products at the right times. Building efficient plans in a timely fashion is critical to our business success.

This paper begins by describing previous planning approaches and then outlines our development approach. We describe a suite of successful applications across Intel's businesses, quantify the realized benefits, and discuss some of the potential future extensions and improvements.

THE PREVIOUS PLANNING APPROACH

Intel's planning systems and business processes have evolved from the simpler processes that supported limited product segmentation and less complex manufacturing routes in the past. Evolving business conditions create planning requirements that need to consider the high-mix environment where differences in packaging, test parameters, wafer type, product lifecycle, and processing drive different planning decisions. Embedded in the existing systems and processes is business logic referred to as "rules" that has evolved to support operation of Intel's complex supply chain. Many of these rules were held in planners' heads and were never formalized. Extracting these rules and forging them into one coherent consistent set that all planners would use had never been attempted.

Historically, this complex problem has been attacked with spreadsheets and massive amounts of human capital. Previous planning processes relied on localized data management and heuristic decision making that were not able to consider all relevant factors or respond fast enough to support Intel's planning needs. The spreadsheets supported the heuristic methods that planners developed over the years and handed down through generations of new planners. Long hours were spent, including evenings and weekends, in collecting and repairing data and in executing a cumbersome process.

In the best case, a locally good solution might be found, but just as often time ran out and a partially refined plan was issued. The limitations of an extended analysis time and inadequate accuracy prevented planners from exploring all decision options to develop an optimum build plan. While this might have been adequate with simpler processes and products and a less sophisticated marketplace in the past, given today's conditions, it was clear that this process and its associated tools were inadequate and unsustainable.

OUR DEVELOPMENT APPROACH

Not all development approaches would be successful in transforming planning systems from those described above to solutions that can support both current and future business needs. As with most transformations, the difficulty is as much about moving the business personnel to a new process as it is about accomplishing this transformation while the business continues to move, adapt, improve, and respond in a dynamic market. Experience has shown that a successful approach must have the following characteristics.

- Collaborative: Frequent interactions with "fingers-on-the-keyboard" users to validate design options and verify priorities.
- Incremental: Multiple smaller projects that move forward a step a time, often helping to clarify the overall problem and success metrics.
- Iterative: Frequent development cycles with a validation checkpoint after each development effort and before the next requirements refinement step.

Such an approach starts by seeking to understand the current approach and by developing a solution with limited changes to rules and data feeds. It is only in taking this first step that we can begin to understand the health of the data and the strength of the rules that are used. Getting the data and rules documented provides a starting point for further improvements and brings the advantage of standardizing how this problem is solved. More benefit will be realized from subsequent efforts when the current process is really challenged, but the automation of the current process is a necessary predecessor for any other improvements. It is not possible to make changes until we understand the details of the current process. This effort follows the steps shown below, which have been refined during many years of decision algorithm pathfinding.

Step 1. Shadow the end-users in the business. Motivated by the understanding that "the devil is in the details," the first step in decision algorithm pathfinding is to shadow the end-users to understand how the current process works, even to the extent that pathfinders would be trained on existing tools as if they were a new planner. Included in this initial analysis is understanding what data are used to make planning decisions and which business rules are used around each set of data. An important stage in pathfinding occurs when pathfinders believe that they understand the user's algorithm. A more important stage is reached when the user believes that the pathfinders understand the algorithm.

Step 2. Develop a prototype and validate it against the current process. The best and perhaps only way to validate that the decision algorithm has been appropriately captured is to transform it into a working prototype and put it into the hands of the users. The expectation is that, when used in close temporal proximity to actual problem solving with the same production data, the user will quickly point out parts of the prototype that are missing or wrong. Rapid iteration is crucial here to hold the users' interest and confidence. Only when the user proclaims that the prototype is producing plans that are as good as (or better than) the current method and producing them as quickly (or quicker) can pathfinding move on to the next step.

Step 3. Make the solution production-worthy. After the prototype is accepted by the planners, it is stabilized to support regular production use. This included code refinement for both the model and the interfaces to the database as well as extensive module and integrated testing. Moving to this step too soon will slow and misdirect prototyping efforts, but never moving to this step will unnecessarily increase business risk and retard continuous improvement.

The desired properties of the tools resulting from this development approach include improvements in both data and decision algorithms delivering benefits in productivity and solution quality. Improved data management including automated data loads would result in fewer errors from manual data input, easier recognition of issues with input data, and fewer planner hours to correct errors in the data. It would also enable rapid evaluation of the resulting plan, support understanding of the sensitivity of business scenarios to changes in the input data, and streamline dissemination of the results. Decision algorithm automation would guarantee rapid and consistent application of the business rules requiring fewer planner and total hours to execute. It also would provide a foundation for continuous improvement by documenting all the standard (and exception) rules and making them easy to extend and test. Automated decision algorithms would allow planners to explore various business scenarios and drive their recommendations based on that understanding.

Since the purpose of planning is to align supply decisions with forecasted demand, selecting the appropriate algorithmic approach was simple. This type of problem has long been solved in academia and at some other companies with a mathematical technique called mathematical programming, or optimization (Hopp and Spearman 1996, Chopra and Meindl 2001). In a mathematical program, business rules are translated into constraining equations that limit the values of the decisions that the solver is making, and objective functions that quantify the total value of the solution against objectives of the business.

A Linear Program (LP) is a subset of mathematical programming where business rules can be represented as linear equations. An LP solver will make constrained decisions to maximize (or minimize) a linear objective function. Examples of decisions an LP could make are how many wafers to start at a Fab in each week and how much product to allocate to different packaging types and configurations as shown in Figure 2.

Once business rules are represented in linear equations, as either the objective or constraint, an LP can be used to solve very large problems in a short amount of time. For example, a typical LP problem with 150,000 variables can be solved in less than a minute. In addition to rapidly generating an answer, the LP solver generates an "optimum" answer. This means that given the objective function used in the LP, one can be certain that this is the best possible answer that is not influenced by alternate starting points or the order of business rules. The speed and quality of the solution produced by the LP allows planners to explore different input data scenarios (different demand, product priorities, or capacity statements) to understand plan dependency on these factors. An LP also allows evolution of the planning systems through modification of existing rules and the addition of new rules as business needs change.

In this paper, we describe the successful implementation of LP tools in Fab-Sort Manufacturing (FSM), Assembly Test Manufacturing (ATM), and Materials Procurement. This includes work with various product divisions including the Intel Architecture Group (IAG) that produces microprocessors (CPUs) and support chips (Chipsets), the Flash Products Group (FPG) that manufactures a range of flash memory chips, and the Intel Communications Group (ICG) with their range of networking and communication products.

FAB/SORT MANUFACTURING

The monthly FSM Planning Reset process must deliver a build schedule for the next nine months that makes the best use of Intel's fungible and expensive capital resources. This plan should comprehend the dynamic nature of supply and demand within the multi-month TPT of Intel's supply chain, the inherent momentum of Intel's supply including in-process wafers and die, product and process roadmaps, and manufacturing performance improvement projects. The FSM plan must also consider the product differentiation as shown in Figure 2 to better align planned supply to forecasted demand through modification of Fab wafer start schedules.

The FSM Solver for FPG

The first improvement occurred in the Flash midrange planning process in Q4'01. This project was small in terms of scope and resource involvement to determine the value of this approach and identify any potential roadblocks. The project focused on the automation of the decision process without any significant changes to the way data were entered or the way results were evaluated.

The planner worked with an optimization expert to transform the business rules used in the previous heuristic, manual process into automated decision algorithms. The current business rules that guided the manual process were translated into mathematical programming within the LP solver. Business rules such as "remain within limits of each Fab's capacity" were translated into constraints (see Equation 1). Strategies such as "minimize missed demand (DemandMiss) and missed inventory targets (InvOver & InvUnder) utilizing relative penalties (MissPen & InvPen)" are translated into objective functions for the LP solver (see Equation 2). Of course these penalties should be translated into true dollar costs and this is on our continuous improvement plan.

f = fab; p = product; r = resource / process; t = time

$$\sum_{p \in r} FabStarts_{f,p,t} \leq Capacity_{f,r,t}$$

Equation 1: Capacity constraint

$$\min \left\{ \begin{aligned} &\sum_{p,t} MissPen_{p,t} \cdot DemandMiss_{p,t} \\ &+ \sum_{p,t} InvPen_{p,t} \cdot \left(InvOver_{p,t} + InvUnder_{p,t} \right) \end{aligned} \right\}$$

Equation 2: Objective function for Demand and Inventory target misses

The overall planning process using the new optimizing tool requires only 10% of the time required by the old process and achieves 100% of the metrics for a good plan set by manufacturing personnel. The old process scored only 85% on these good plan metrics necessitating a follow-up meeting to negotiate changes. The success of the new tool eliminated the need for this meeting and indicated that it was possible to capture all the decision rules used in solving FPG's FSM planning problem. It demonstrated that these rules could be represented in a set of linear equations and that the solver could be controlled with priorities and penalties that would make sense to planners and support business requirements. The amount of improvement in this project, both in terms of productivity and plan quality, indicated the potential for this approach to improving FSM planning in other areas of Intel's supply network.

There was a lateral implementation of this same planning tool in Q2'02 to the IAG Chipset division. The same data management system, which allows for loading of the input data from existing spreadsheets and inserting the results back into that same spreadsheet, was migrated to Chipsets.

FSM Solver for IAG CPUs

Building from the success in FPG and IAG Chipsets, the decision was made in Q1'02 to develop a similar tool within IAG CPU. The CPU planning problem had the additional complexities of product binning, product configuration, and more complex product mapping as shown in Figure 2.

The wider scope of this project inherited additional complexities but also provided more insights into the data management side of planning solutions. Through this project, we were able to evaluate the current health of planning data and characterize the nature of planning data problems. These problems included the complexity of distributed manual data ownership, integrating data from multiple systems, evaluating the results of an automated decision algorithm, managing data fallout, and maintaining traceability to changes to the data.

Utilizing the CPU solution as a starting point, a similar solution was implemented to support planning for some parts of ICG. The database/solver architecture will allow portions of ICG to modify and add business rules in the form of additional constraints, objectives, and their supporting data feeds to supplement the existing application. The ability to translate the CPU solution for use in ICG again demonstrates the scope of the solution and the ability of the tool to respond to the changing needs of different divisions. Performance results for IAG and ICG in both planner productivity and plan quality were similar to those achieved for the initial FSM FPG tool.

FSM SOLVERS on SQL with .Net

Taking the next step in IAG CPU planning in the area of data management to enable further and more complex decision automation, it was recognized in Q1'04 that there was a need to transfer the application to a more robust SQL platform. SQL provides a data management platform that allows storage of more data and faster, more complex data manipulations. This data management improvement allows storage and comparison of multiple versions and any manual modifications to the data. The multi-user SQL environment allows parallel examination and manipulation of the data to speed up the reset process and enable more thorough investigation of the solution space through multiple solves.

As part of this transition, the team also took advantage of the benefits of .Net through development of interfaces and architecture in .Net. The .Net architecture allows more rapid prototypes and development going forward. It facilitates rapid changes to different data sources if business decisions require it. The .Net components allow for rapid development of reports and interfaces as well as a more robust interaction with the solver application including faster and better data transfer from and to solver applications. Overall, the time required for data management was decreased to 15% of its previous level.

Building off the success of the SQL/.Net transition into CPU, efforts began to migrate this solution to other product divisions. These migrations included modifications to both decision and data components of the application to comprehend division-specific requirements. Each division may have a different set of business rules based on the value of inventory and accuracy of forecasts. These efforts verify that the FSM wafer start planning process is similar enough between divisions to share the same application platform. This widespread implementation enables knowledge and process sharing between divisions.

Figure 3 provides an overview of the extent and focus of this sequence of FSM projects. Each project is shown in terms of what was gained in the area of both data management and decision automation. These efforts demonstrated the level of interdependence between data management and decision automation. The amount of decision automation is limited at some point by the level of data management. A complex decision algorithm is of limited value if results can't be quickly and accurately evaluated. A planning solution that requires too much time to collect and prepare data will prevent planners from having time to generate a good plan and increase the risk of data issues impacting the timing and/or accuracy of that solution.

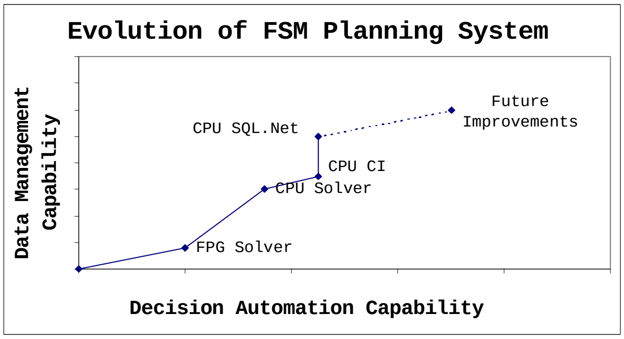


Figure 3: Data management and decision automation evolution of FSM systems

ASSEMBLY/TEST MANUFACTURING

ATM planning has the responsibility to match wafer supply from Fab to demand in the market by routing the right wafers to the right A/T factories and planning the right starts into the A/T factories. The ATM planning process had become very complex not only due to the increase in the number of products, but also the number of A/T factories and Fabs. The planning process for the ATM factories had become so complex for the manual process that it was not only divided by products but also decomposed into a request response process. The product division planners would do a detailed analysis on demand including a high-level analysis on supply resulting in a preliminary plan. They would then submit their preliminary plan and supply request to the ATM factory planners, who would respond based on a detailed analysis of the supply including splitting the builds to various A/T factories. A lot of time and energy was spent in this process including regular iteration of this request and response cycle.

The optimization techniques for planning were extremely well received as a tool that would prevent planning failures by automating the planning process, creating an effective work-life balance, and producing a better plan. The approach taken was to incrementally solve the problem by first automating the response process and then step by step integrating the whole request-response process.

A major challenge for the project was getting clean and consistent data. A lot of the Plan of Record (POR) data was not clean, and the clean-up effort had been undertaken as a multi-year project. To provide good quality data, it was decided to get some data from the POR systems and some from the manual system used by the planners, which included formalization and maintenance of data that were stored and maintained in planners' heads. A plan to move from the manual data sources to the POR data sources was also formalized to intercept the data clean-up efforts.

The team had extensive user involvement and commitment that helped in finding the requirements faster. The technical team was very flexible with the requirements and created prototypes to test and confirm the business rules. With exceptional teamwork and a dedicated effort the team was able to put the solver tool in production for the response process for one product in four months in Q4'03. After that the response solver tool was proliferated to five products including the Intel® Pentium® 4 processor and products built on Intel® Centrino™ mobile technology, and will continue for new products.

After the successful implementation of the solvers for the response process, the team took the next step of integrating the request-response process. The solution needed to comprehend the different time zones of different planners, so that it would be robust for multiple users, and have traceability and scalability. The platform chosen for developing the solution was .Net with an SQL database. The team incorporated POR data sources for demand, but still used the manually maintained mappings of Figure 2. With dedicated team effort and dedication of the users, the tool was successfully implemented for one product in Q1'05. With the help of automation and optimization the business process has been completely reengineered to integrate the demand and supply planning process. This tool is now the POR tool for one of the Pentium products. It will be proliferated to all products for which we have the response tool, and it will be used to plan any future products including the multi-core products.

ATM planning is an example of how operations research has been used at Intel to not only build better plans but also re-engineer the business process. Integration of the request-response process would not have been possible without the optimization and automation tools. This implementation and integration of the planning processes

® Intel and Pentium are registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

™ Centrino is a trademark of Intel Corporation or its subsidiaries in the United States and other countries.

serves as an example for future enhancements and integration of the supply-chain planning for Intel.

These tools have not only provided strategic value by paving the way for future planning processes and systems, but have also provided tactical benefits. The response tools have provided an ROI of tens of millions of dollars per year due to better plans and have improved the work-life balance for the planners. The need to work on the weekends has been completely eliminated, working late has been reduced considerably, and work is more evenly distributed over the week. The integrated solver has reduced the time required for reconciliation between the central and A/T factory planning by 75%, and has improved the quality of the plan with better utilization of capacity and better demand support by 5% to 10%.

MATERIALS PROCUREMENT

Supplying Intel's A/T factories that build microprocessors and chipsets with the appropriate substrates is a complex task from a number of different perspectives. The combinatorial complexity of this planning problem stems from the large number of suppliers, package families and individual substrates, and A/T factories that are spread across many geographies, as shown in Figure 4. Note that each supplier and each factory use many substrates, and each substrate is made and used in many places to hedge uncertainty. Financial complexity is due partly to the very large expenditures made in substrate manufacturing and consumption, and partly to the complex contractual arrangements that surround substrates. These contracts protect all parties against short-term demand shifts between products and against uncertain future markets as new substrates are developed. Moreover, there is evolutionary complexity as the number of suppliers, packages, and factories grows over time, as the lifecycles of products decrease, and as the ramp rates for supplier and factory qualification increase.

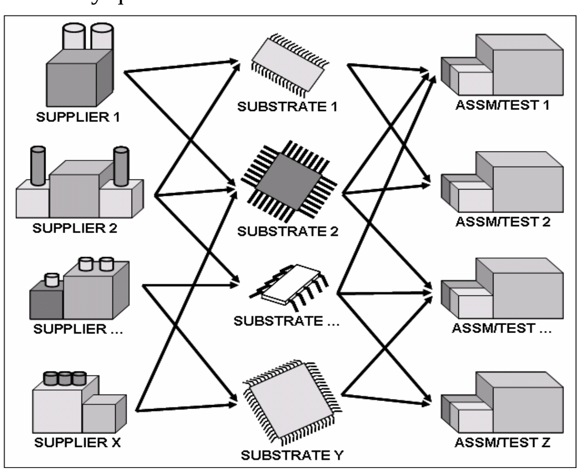


Figure 4: Supplier, substrate, buyer complexity

In past years, before this problem became so complex, planning individual substrates by a trial and error methodology was common using multiple custom spreadsheets and extensive manual data management. Business rules, some embedded in contracts, some passed down the management chain, were contained in planners' heads and applied during spreadsheet runs. Understanding of suppliers' capacity was based on experience. Planning was undertaken in parallel for many substrates with the risk of simultaneous calls to the same supplier requesting conflicting actions. This process could include multiple iterations between suppliers and Intel planners and still produce a suboptimal plan for everyone involved.

This materials procurement process has been dramatically improved by the development of a centralized data base and a Web-based optimization tool. The data base provides a single repository for data about all suppliers, package families and individual substrates, and A/T factories. More importantly, it contains capacity models co-developed by Intel and its suppliers that accurately represent flexibility between substrates. The optimization tool is based on an LP formulation of the planning problem and addresses all suppliers, substrates, and factory needs simultaneously. Very significantly, it encodes the wide variety of business rules that must be considered.

This new business process is now executed in less than 10% of the time required by the old process. The resulting plan is capacity feasible without involving extensive iterations with suppliers and has saved Intel tens of millions of dollars on its expenditures since the system was implemented in Q3'03. Given solver run times that are less than five minutes, and a Web-based interface that provides overall system and data transparency, further savings are expected as the system is employed beyond simple planning. For example, the tool can support what-ifs around the capacity needed for substrates in development for future products as well as for future demand scenarios. In addition, it can be used to explore various pricing and contracting combinations before and during business negotiations. This successful approach to substrate planning is being proliferated to other materials planning problems in business groups across Intel.

BENEFITS

These efforts have provided qualitative and quantitive benefits to the organization, from improvements in data management to decision automation. The benefits of these efforts were realized in areas of productivity, data quality, solution quality, and continuous improvement.

Productivity

Copy and paste data collection tasks and subsequent checking were replaced by direct database links. Faster plan generation was provided through automation of decision algorithms with mathematical programming solvers. Faster plan analysis was realized through increased visibility into plan details and simplifies consideration of alternate scenarios. Dynamic reporting allowed for comparison of scenarios (session to session and run to run) to more quickly identify dataset differences and assess the quality of each solution. And of course, documentation and automation of business rules make it more efficient to train new planners.

Data Quality

Copy and paste data collection errors were eliminated by the direct linking to external data sources. The auditing of collected data for predetermined data quality issues was automated and formalizing the traceability of required manual adjustments resulted in more rapid issue resolution. There was improved identification of root causes of data issues with data quality checks and reporting visibility.

Solution Quality

There was a much more uniform understanding of business rules across each business through documentation and subsequent translation into mathematical programming constraints and objectives. Encoding the rules into the LP tools guaranteed a consistent application of business rules from plan to plan and from planner to planner that was never previously achieved. The optimization solver generates the best solution given the supplied data, as opposed to manual heuristics that previously risked stopping at a feasible solution. Furthermore, better data management support and faster solvers speed up the scenario setup–plan generation–solution analysis cycle to allow more complete consideration of alternate solution options.

Continuous Improvement

This transparent system has improved collaboration and coordination between various planning arms. Now different planning groups have better visibility and understanding of each others' capabilities and limitations. A standard tool with comparable rules across divisions has made it easier to recognize differences and similarities for continuous improvement. The business rules can be changed more quickly and more reliably to reflect changes in Intel's strategies, changes in Intel's markets, and improvement ideas from the planners using the tools on a regular basis. This has improved collaboration to improve the planning process. From a broader perspective, these new transparent planning tools have started to increase collaboration between different functions in the company. For example, Sales and Marketing now knows the business rules employed in planning and can use this information to better meet the needs of customers.

FUTURE IMPROVEMENTS

Both the old and the new planning processes include the assumption that the forecast parameter values supplied to the planning tool (e.g., demand, capacity, yield, TPT) are all suitable to be used in planning. In fact, all of the parameters are measurements made on stochastic processes. The forecasts are constructed by looking at historical data and current improvement projects to estimate what will happen in the future. Some of the stochastic processes on the supply side are well understood. For example, it is possible to reproduce the distribution of TPTs of items passing through a factory by building a discrete event simulation that includes random machine breakdowns, the unavailability of equipment technicians during breaks, and so on. For the demand side, the situation is much different since the underlying stochastic processes are much more difficult to characterize. Our investigations of forecast error (forecasts vs. actuals) indicates that planning systems should consider the limits of our ability to make accurate forecasts of the expected supply that Intel can produce and the expected demand that customers will purchase. Future efforts to improve planning system design and operation given this variability will make Intel's planning systems more valuable in determining how to utilize capacity resources to effectively satisfy market demand.

CONCLUSION

Development of automated data systems and optimization-based tools has revolutionized Intel's supply-chain planning processes. Benefits have been realized in data and solution quality as well as planner productivity across all product divisions and all manufacturing organizations. Planning time has decreased dramatically, supply costs have been reduced, and demand satisfaction has improved. But perhaps the most important contribution of the efforts described here is the facilitation of continuous improvement. With better data, documented business rules, and fast planning and analysis tools, Intel planners have the time and facilities to define world-class performance.

ACKNOWLEDGMENTS

We acknowledge our manager, Karl Kempf, who has always encouraged and supported us to challenge the status quo and push for continuous improvement. We also acknowledge the numerous planners whose willingness to acknowledge the numerous planners whose willingness to share the "secrets" of the current process and accept improvements allowed us to realize these successes.

REFERENCES

[1] W. J. Hopp and M. L. Spearman, *Factory Physics: Foundations of Manufacturing Management*, New York: McGraw Hill, 1996, Ch. 16 (Aggregate and Workforce Planning).

[2] S. Chopra and P. Meindl, *Supply Chain Management: Strategy, Planning, and Operation*, Part 2 (Planning Demand and Supply in a supply-demand network), Prentice-Hall, Upper Saddle River, NJ, 2001.

AUTHORS' BIOGRAPHIES

John W. Bean is a senior supply-chain engineer in the Decision Technologies department in Intel, Chandler, Arizona. He graduated in 1994 with a B.S. degree in Manufacturing Engineering from Brigham Young University, and in 1997 he graduated from the Leaders for Manufacturing (LFM) program at the Massachusetts Institute of Technology (MIT) with M.S. degrees in Material Science and Management. He has worked on WIP management, information systems development, and optimization solutions in FSM Industrial Engineering and Planning. His e-mail is john.w.bean at intel.com.

Amit Devpura is a senior supply-chain engineer in the Decision Technologies department in Chandler, Arizona. He graduated with a Bachelors degree from IIT Delhi, India in 1995. He received an M.S. degree in Mechanical Engineering and a Ph.D. degree in Industrial Engineering from Arizona State University (ASU) in 2000 and 2003, respectively. He started working at Intel in the Q&R organization in 2000 and has been working in Decision Technology since 2003, where his focus has been on implementing decision science tools in ATM planning. His e-mail is amit.devpura at intel.com.

Michael C. O'Brien is a senior supply-chain engineer in the Decision Technologies department in Chandler, Arizona. He graduated in 1991 with a B.S. degree in Systems Engineering and in 1996 with an M.S. degree in Industrial Engineering from the University of Arizona. He has worked on shop-floor scheduling projects, tactical and strategic planning systems, and has operated as an internal consultant providing simulation, optimization, and statistical support in a wide variety of business environments. His e-mail is michael.obrien at intel.com.

Shamin Shirodkar is a senior supply-chain engineer in the Decision Technologies department in Chandler, Arizona. He has a B.S. and M.S. degree in Industrial Engineering and is currently pursuing his Ph.D. in Applied Mathematics at Arizona State University. He has worked on applying the Theory of Constraints to Ramping Semiconductor Fabrication facilities and on implementing a variety of decision science tools in the supply-chain arena both internal to Intel as well as collaboratively between Intel, its material suppliers, and manufacturing subcontractors. His e-mail is shamin.a.shirodkar at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

Inventory Modeling

Kurt L. Johnson, Information Services and Technology Group, Intel Corporation

Index words: inventory modeling, Response Surface Model (RSM), distribution network, stochastic models

ABSTRACT

Original Equipment Manufacturer (OEM) customers and distributors across the electronics industry are pursuing all available means to reduce their materials costs while maintaining high service levels to their end customers. The OEMs have asked their suppliers to implement various forms of Consigned Inventory and Vendor Managed Inventory (VMI) programs. Some distributors stock high levels of Finished Goods (FG) inventory in order to maximize their service levels to the end customers and they use return policies to minimize their risk by returning excess product. In 2002, inventory across Intel was $2.3B or 8.5% of sales, including $0.7B of FG inventory. If the approach of these OEMs and distributors were to be adopted, we could expect the FG inventory to increase by >20%, resulting in higher levels of inventory risk to Intel.

This paper describes a new way of optimizing Intel Corporation's supply chain, from factories to customers. The methodology that will be discussed uses statistical methods to characterize the order distributions of customers and the distribution of times to ship products from different points in the supply chain (factories to customers). These results are then used to build a stochastic simulation model that can be experimented on to gather data that contain information on interaction effects and inventory pooling effects. Response Surface Modeling (RSM) methods are used to set up the experimental design and to analyze the results. This allows a statistical model to be developed that allows the user to explore the effects of varying inventory levels at different locations on customer-service levels. By using this methodology, optimal placement of inventory (minimizing the inventory while providing the desired service level) can be achieved.

INTRODUCTION

As the Intel distribution network continues to expand in order to reach new markets, the complexity and impact of the management of the network has grown. With this growing complexity the need to optimize distribution network processes that affect Order Fulfillment Quality (OFQ) is critical to avoid lines being down at customer manufacturing sites. OFQ is made up of four key elements:

1. The right product was delivered to the right customer.
2. The product arrived undamaged to the customer.
3. The correct amount of product was received by the customer.
4. The product arrived at the customer site at the agreed-upon time.

The last of these elements, the product arrived at the customer site at the agreed-upon time, is highly impacted by the availability of product and the placement or staging of product in the distribution network. A typical representation of the distribution network structure for a given geography is shown in Figure 1. It shows that product can be held in inventory at warehouses or distribution centers.

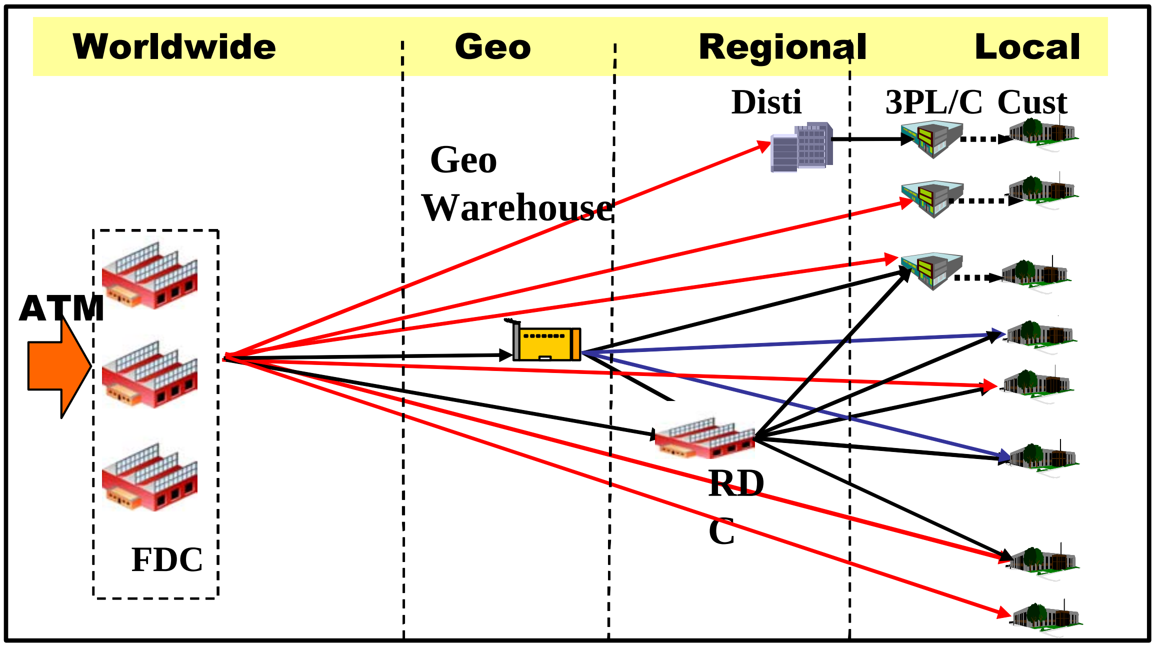


Figure 1: Typical distribution network

In this paper we discuss an alternative methodology that was successfully used to optimize inventory placement within the distribution network. The methodology includes developing a stochastic simulation model and then running experiments on that simulation model to develop a Response Surface Model (RSM) to characterize the effects of different allocations of inventory across the distribution network. This approach to experimenting on a simulation model was used because experimentation on the physical distribution network would have these undesirable consequences:

1. The customer may have lines down because of lack of product.
2. The time to get results could run into years.
3. The cost and difficulty of managing these experiments would be high.

In this paper we focus on statistical methods used to build a stochastic simulation model of the distribution network. This model takes into account the variables in the delivery process that are uncontrollable, called noise variables (i.e., shipping times from one location to another and product order patterns). It then addresses the use of RSM to explore and optimize the effects of variables that can be controlled, called controlled variables (i.e., amount of inventory to be placed at different locations and the replenishment periods) on the service level (i.e., did the product arrive when it was supposed to). Optimization in this sense is defined as meeting service-level goals with minimal costs to the distribution network.

DISTRIBUTION NETWORK MODELING

The methodology used to optimize inventory placement within the distribution network consisted of two modeling phases (stochastic simulation of the distribution network and statistical modeling using RSM) as well as a financial analysis. The stochastic simulation models were developed to create virtual models of the distribution network that could be experimented on without impacting our customers and were run in a relatively short period of time. The statistical models, RSMs, were based on data from experimental runs of the simulation models and provided information on service levels vs. amount of inventory placed at various points in the distribution network. Since it is possible to have a variety of inventory configurations that yield a particular service level to a customer, a financial analysis was done to optimize the specific configuration to minimize cost while providing the desired service level. This methodology was repeated for each major geography (Asia, Europe, and the Americas), and by each product type (tray and boxed CPUs, and motherboards).

Stochastic Simulation Modeling

The stochastic simulation model was built using a software package called e-SCOR*. To achieve the stochastic nature of the model, estimates of the variability for the noise factors had to be quantified and built into the model. These noise factors are uncontrollable, and where historical data were not available, were assumed to be random. They include order patterns, shipping or transit times, and throughput times through the warehouses and distribution centers.

For order pattern variability, historical data were available and used in the simulation model. Historical order pattern data of the same product type (i.e., CPU order pattern history for CPU products) were used in the simulation model to represent that source of variation. The use of historical data is the ideal method for representing variability of noise factors.

When historical data do not exist for noise factors, then other methods for representing those sources of variation are used. Some of these methods include using a triangular distribution, a uniform distribution, or some assumed distribution, such as a normal distribution. The parameters of those distributions can be estimated by individuals considered knowledgeable in these areas.

For transit times and throughput times no historical data were available, and a triangular distribution was used. Individuals knowledgeable in shipping and customs were asked what the minimum, typical, and maximum transit times are for given shipping lanes. These were used as the estimates of a triangular distribution. The same process was used with experts in the warehouse and distribution centers to get triangular distributions for throughput times.

Once the simulation model was developed in e-SCOR by geography and product type, this model was used to represent the physical distribution network. The model in e-SCOR also required that the controlled factors of amount of inventory in each location and rules for replenishment were to be entered into the model. In order to generalize the simulation results to multiple products, the volumes of inventory were normalized to Days of Inventory (DOI).

A single experimental run of the simulation in e-SCOR was run to investigate a particular configuration of the distribution network. Each configuration was an investigation of how the distribution network behaves under a given set of conditions. These conditions are comprised of the days of inventory to be targeted at each warehouse and distribution center along with replenishment rules. The output of each experimental run is the service level attained for the customers in a region for a given configuration of the distribution network.

The investigation of all possible configurations was not feasible as the time to run the simulation for a given configuration was approximately one hour.

Response Surface Modeling

Experiments were run on the virtual model of the distribution network built in e-SCOR. These experiments had the purpose of quantifying the effects of the amount of inventory placed at various locations within the distribution network on the service level (measured as the percentage of time that product arrived to the customer at the agreed-upon time). Since volumes of products vary for different stages in a product's life cycle, inventory levels were measured as DOI, where a DOI is the average daily volume of product that passes through a location in the distribution network.

Central Composite Designs (CCD) were the family of experimental designs that were run. In a CCD, the factors under investigation (in this case DOI) are set to have the points of a typical full-factorial design (all combinations of low and high values) along with augmented axial and center points. Multiple runs of the center points are run to get an estimate of the error in the model. Figure 2 shows a graphical representation of a CCD.

* Other brands and names are the property of their respective owners.

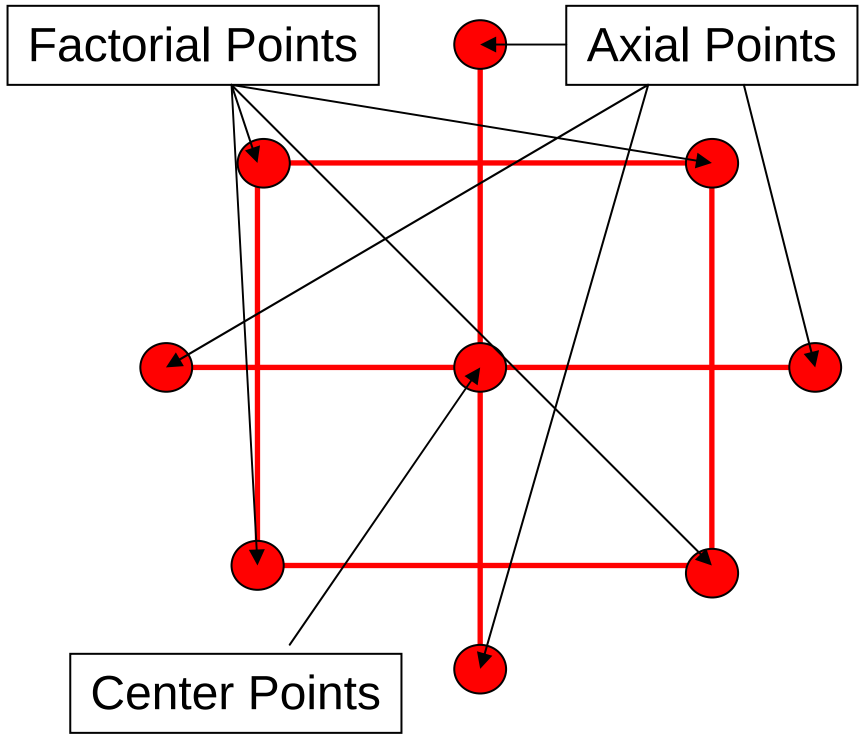


Figure 2: Central composite design

The factor settings in a CCD were chosen so a quadratic model with interactions could be estimated. The mathematical form is

$$y = b_o + b_1x_1 + b_2x_2 + \ldots + b_nx_n + b_{12}x_1x_2 + \ldots$$
$$+ b_{(n-1)n}x_{(n-1)}x_n + b_{11}x_1^2 + \ldots + b_{nn}x_n^2, \text{ where}$$

y = estimate of the service level.

Only those terms of this model that are statistically significant (p-value <= .05) were included in the model.

From this mathematical model a graphical response surface was made to visualize the effects of different levels of DOI on the service level. Figure 3 shows an example of the graphical output that was obtained.

As can be seen in Figure 3, many different possible combinations of inventory levels yield the same estimated service level. For a desired service level the optimal settings are determined by a financial analysis that finds the minimum cost of inventory in the network. A more simplistic method would be to minimize the overall DOI in the distribution network. Another consideration in choosing inventories is the sensitivity of the service level to changes or departures from the targeted DOI at the various locations.

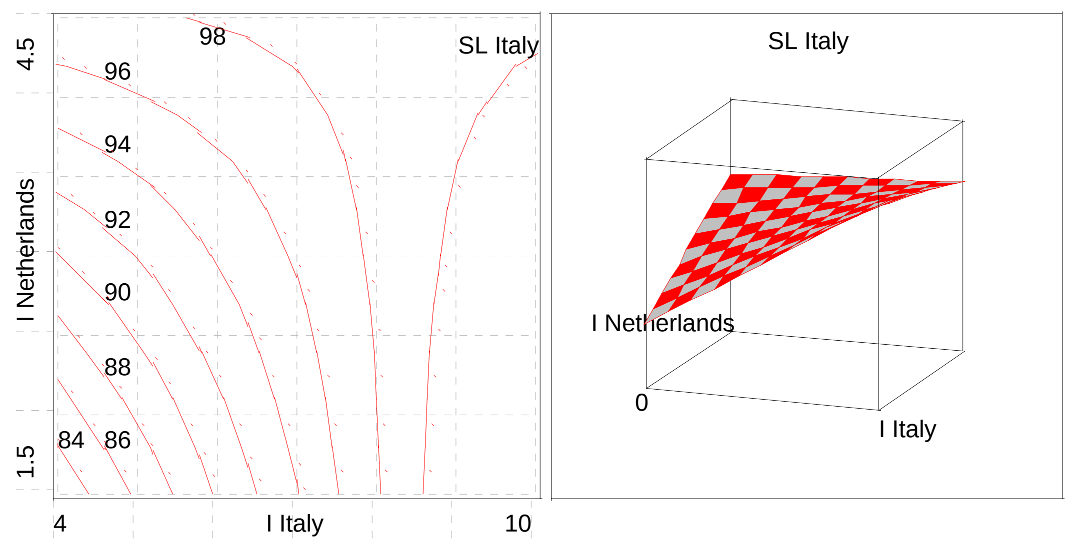


Figure 3: Response surface output

RESULTS

The results of these efforts were originally reported in a paper by Michael Waithe, et al. at the 2004 Intel Manufacturing for Excellence Conference (IMEC) titled "Use of Simulation Modeling to Reduce Product Distribution Costs and Develop Rules for Supply Chain planning." Results for our top four OEM customers include a drop in average days of inventory from a high of 11.8 in Q3'01 to 8 days by Q3'03, a $145M reduction in working capital.

These solutions have enabled Intel to successfully win or retain business at a lower cost. In response to 37 consignment requests, Intel provided consignment twice and offered a lower cost inventory hub solution in twelve instances, realizing a $7-13M Intel cost savings per customer depending on customer size and complexity. The lower cost alternative offered to one customer avoided a loss of motherboard business and enabled a $150-200M growth in annual revenue.

The modeling output data were validated by actual operational results by a second customer. This customer realized an inventory reduction of 10 days. Their inventory was reduced from an average of 11 days in 2002 to 1 day in 2003 after the Jointly Managed Inventory program implementation.

The Supply Network modeling data also demonstrated that pooling geographic supply would decrease overall network inventory and not just shift customer inventory to Intel's shelves. Scenario analysis around key variables such as demand variability, forecast error, frequency of delivery, TPT improvements, and inventory placement identified additional opportunities.

The data from the models were instrumental in the decision to expand the JMI program to the distributor channel for Boxed CPUs. The data indicated an opportunity to reduce channel inventories by ~21% and have been validated by pilots in the European and Asia Pacific geographies.

DISCUSSION

The modeling, both stochastic and statistical, yielded results that did not contradict the general results that would be seen if only the impact of variation from a strictly theoretical perspective was studied. That is to say that one would expect a benefit (reduction in inventory) to be seen by pooling inventory upstream from the end customer. The rationale is that if each of "n" customers has an average demand D_k and the sum of the average total demand of those "n" customers is D_{total}. That is:

D_k : average demand of customer k

σ_k : standard deviation of demand of customer k

For the "n" customers with independent demands D_1, D_2, … D_n which are met from a single pooled inventory location, the total average demand at the pooled location is

$$D_{total} = D_1 + D_2 + \ldots + D_n$$

and the standard deviation of total demand is

$$\sigma_{total} = \text{Square Root } (\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2)$$

This indicates that the overall variability of an upstream location is less than the sum of the variability of the end customers. This result was confirmed by the modeling. The additional question answered by the modeling was how much pooling could be utilized while still meeting the end customer service level.

In the response surface modeling a quadratic model was fit to the data. Since service level is monotonically increasing with respect to increasing inventory, this assumption does not hold true. However, in the area of investigation where service level is between 80% and 95%, the quadratic model holds up. If one was interested in the 95% to 100% region for service level, different mathematical forms would need to be investigated.

This approach to building a stochastic simulation model and then running experiments on those simulation models is a viable approach to improving processes that might otherwise be difficult to run experimentally.

CONCLUSION

Through the use of stochastic simulation models and traditional ways of running experiments (i.e., RSM), it was possible to make improvements on the strategies of placing inventory throughout the distribution network. This was accomplished first by quantifying the sources of variability to emulate the physical distribution network through a simulation model, and secondly by varying controllable factors (DOI) in a structured manner to characterize how these factors affect the service level that regional resellers and OEMs receive.

ACKNOWLEDGMENTS

We thank Michael Waithe, Spencer Merril, and Daryl Lambert.

AUTHOR'S BIOGRAPHY

Kurt Johnson currently works as a quality engineer in the Information Services and Technology Group at Intel Corporation. His professional interests include applying statistical methods to improve business processes and systems. Mr. Johnson has a B.S. degree in Mathematics from the University of Utah and an M.S. degree in Statistics from Brigham Young University. His e-mail is kurt.l.johnson at intel.com.

RosettaNet for Intel's Trading Entity Automation

John Cartwright, Information Services and Technology Group, Intel Corporation
Jay Hahn-Steichen, Information Services and Technology Group, Intel Corporation
Jackson He, Intel, Digital Enterprise Group, Corporation
Thurman Miller, Information Services and Technology Group, Intel Corporation

Index words: RosettaNet, Trading Entity Automation (TEA), e-Business, XML, Web services

ABSTRACT

As one of the founding members of the RosettaNet* consortium, Intel has aggressively pursued utilizing RosettaNet to support its supply chain. Over the past five years, Intel has implemented over 1000 Trading Entity (TE) touch points, encompassing 24 different RosettaNet Partner Interface Processes (PIPs*) enabling more than 50 unique business transactions with over 200 TEs. In 2004 alone, Intel realized nearly $40M ROI in business value.

We begin with an overview of key RosettaNet technical components. We then summarize the success Intel had over the past years in building new business processes and the e-Business infrastructure of RosettaNet. Finally, we explore the future of Business-to-Business (B2B) exchanges and the next generation of B2B architecture.

INTRODUCTION

In 1998, Intel was a leading advocate among a group of companies promoting the concept of Trading Entity Automation (TEA). This resulted in the formation of the RosettaNet Consortium. Intel devoted resources and funding to drive the definition of the essential elements of RosettaNet TEA: the RosettaNet Implementation Framework (RNIF), RosettaNet Dictionaries, and Partner Interface Processes (PIPs).

At the same time as leading RosettaNet standards development, Intel also actively engaged in using these protocols to build TEA solutions for improved business agility and productivity. In early 2000, Intel was one of the first two TE companies to implement RosettaNet. Over the past five years, Intel has made RosettaNet part of Intel's overall e-Business Business-to-Business (B2B) infrastructure and supply-chain automation processes. Intel has enabled over 1000 TE touch points, encompassing 24 different RosettaNet PIPs, enabling more than 50 unique business transactions with over 200 TEs. In 2004 alone, Intel realized nearly $40M ROI in business value from using RosettaNet.

While much of the savings is attributed to automating previously manual processes (such as FAX, Web applications, e-mail), the largest ROI has come from new business models that were unattainable without automation. One of these models is the Outsource eSolutions/3PL model (OeS/3PL). This complex business model involves third-party logistics companies, subcontract manufacturers, suppliers, OEMs, and other customers. Thirteen RosettaNet transactions were utilized across the supply chain to effectively communicate both raw and finished goods material movements.

One of the main challenges in creating a B2B platform is the need to insulate TEs from internal changes. Our internal enterprise systems are very complex, running a variety of Enterprise Resource Planning (ERP) systems for different parts of the supply chain. Each of these internal systems uses a different technology for ERP, OS, database, and interfaces. In this paper, we show how Intel provides a standard B2B platform to allow these disparate systems to effectively and consistently connect to our RosettaNet gateway (public process). We also detail how our architecture insulates internal applications (private processes) from external TEs–allowing Intel to continually evolve its application and technology infrastructure with little or no effect on the TE side of a transaction.

As Intel built its robust public and private process infrastructure, we also expended a great deal of time to optimize the entire B2B supply chain. This ensured that B2B solutions always have a RosettaNet-compliant interface enabling the following:

- Non-proprietary communication with TEs.

* Other brands and names are the property of their respective owners.

- Ease of integration with other RosettaNet-capable TE organizations.

This has been an effective strategy to encourage our TEs to adopt RosettaNet.

ROSETTANET STANDARDS AND TRADING ENTITY AUTOMATION

Intel was a founder of the RosettaNet Consortium (www.RosettaNet.org*) when it first started in 1998. The RosettaNet Consortium now has more than 500 member companies and has become a leading organization in the creation, implementation, and promotion of open e-Business standards and services.

RosettaNet defines a set of XML-based protocols to facilitate secure electronic exchange of standardized business documents between TEs over the Internet. RosettaNet has three key standard specifications:

- RosettaNet Implementation Framework (RNIF)
- RosettaNet Business and Technical Dictionaries
- Partner Interface Processes (PIPs)

(For more detailed information on these specifications, please visit the RosettaNet standards website.) [1]. The RNIF specifies the protocols for XML-based message packaging, secure and reliable routing, and basic TEA constructs. The RosettaNet Technical and Business Dictionaries describe valid data formats for business transactions. PIPs define sequences of business transactions and interchanges, as well as expected responses.

As illustrated in Figure 1, the RosettaNet RNIF, Dictionaries, and PIPs form the foundation of the RosettaNet infrastructure. The first step of a RosettaNet implementation is to build/enable a RosettaNet infrastructure. By design, RosettaNet is based on XML and platform independence. TEs can implement RosettaNet on different platforms provided these follow the RosettaNet specification.

After a RosettaNet infrastructure is built and tested, TEs need to agree on the types of business transactions they will conduct over RosettaNet. There are many types of business transactions that are defined in RosettaNet PIPs; for example, product catalog and purchase order. TEs must choose what PIPs they will support.

RosettaNet defines standard protocols for public processes to be used between TEs–the processes that all TEs will follow to accomplish business transactions. However, to accomplish trading entity automation, each trading company must integrate its public processes with its own private processes; a linkage between the communications gateway and back-end ERP systems. Private processes are TE-dependent and are outside the scope of RosettaNet standards.

Once a RosettaNet infrastructure is successfully deployed, and business transactions (PIPs) are selected, a typical RosettaNet transaction proceeds as follows:

- A business request (e.g., purchase order) is generated in a back-end ERP system (private process). The purchase order is packaged following the formats defined in the RosettaNet Dictionaries and organized in the message sequence described in the PIPs.
- A validated PIP package is sent through the RosettaNet infrastructure over the Internet to the targeted TE.
- On receipt of the PIP package, an acknowledgement is issued to the corresponding TE. The package is unpacked and interpreted, then the business request is sent to the back-end ERP (private processes).
- After the back-end generates responses to the business request (e.g., successfully fulfills a purchase order, or rejects an order), a response PIP package is formed and sent back to the corresponding TE.
- Upon TE receipt of the response to the original business request, the message is unpacked, interpreted, and appropriate actions followed.
- Throughout each business transaction, the RosettaNet infrastructure will conduct data transfer over secured channels and keep track of each transaction step in a non-repudiation database.

Developing RosettaNet solutions requires reengineering/integration of internal business processes and collaboration of external processes among TEs. It involves IT professionals, business process experts, and legal teams.

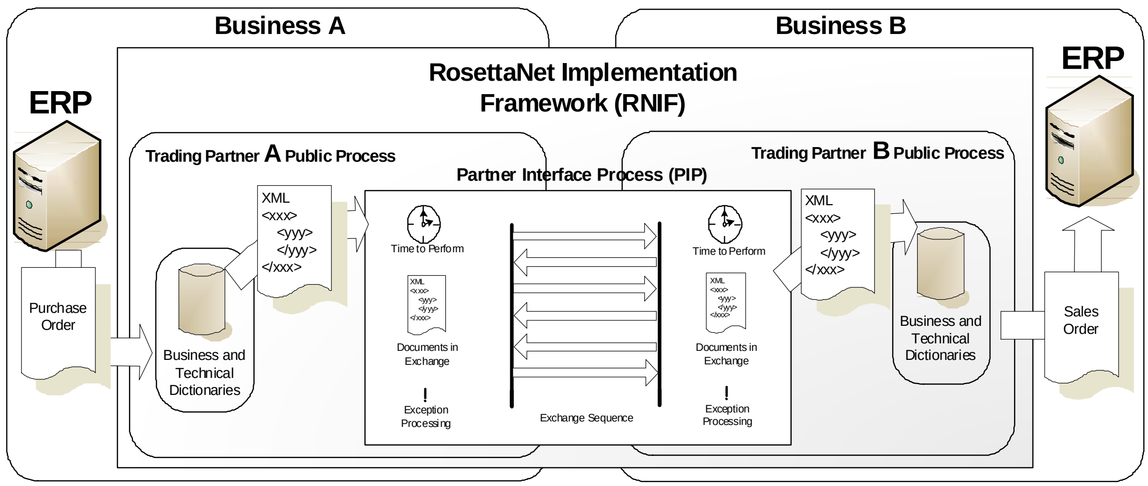


Figure 1: RosettaNet Trading Entity Automation (TEA)

ROSETTANET'S VALUE PROPOSITION

RosettaNet provides value to companies in a variety of ways. First, by having a commonly understood public interface process, Intel and our TEs will spend less time negotiating how we are going to perform our business interactions. This also solves the problem of having to develop and support niche B2B solutions.

Second, under the RosettaNet umbrella, companies are able to share their best known practices for supply-chain integration and strategize on how to become more efficient without the fear of anti-trust legislation.

Third, companies are able to leverage a common B2B infrastructure that is capable of connecting with all tiers and verticals of TEs. The ROI of the B2B infrastructure is dependent on the utilization. RosettaNet provides a partner interface standard that can now be used with all the tiers and verticals of TEs that we need to engage with.

Intel's B2B Platform

From the outset of Intel's B2B implementation in 2000, we planned for a future in which we would be connecting hundreds and eventually thousands of trading entities to our back-end ERP systems. One of the realities Intel has to deal with while achieving this goal is that our TE population is at significantly varied stages of technology deployment and sophistication.

Many of these TEs were already doing electronic exchanges using Electronic Data Interchange (EDI) and would be transitioning their internal systems to communicate over RosettaNet. Some small and mid-sized enterprises (SMEs) did not have the IT staff or budget for a large-scale RosettaNet implementation. To complicate matters, Intel's internal enterprise systems are very complex, running a variety of ERP systems for different parts of the supply chain. Each of these internal systems has different ERP application, OS, database, and interface standards. However, all entities with whom we trade are able to do some form of electronic transmission. Because it is undesirable for these trading entities to have limited access merely because of the cost of enabling an emerging industry standard. Intel has adopted a policy of promoting selected standards with TEs, while enabling technology-specific gateways as warranted by business needs.

The challenge then became: how can we connect such a diverse TE population with the equally diverse Intel back-end ERP systems–and insulate all of these constituents as systems evolve and change?

We developed an approach to public gateway design for connecting all Intel trading entities:

1. Enable RosettaNet.
2. Allow existing EDI transactions to continue.
3. Create a Web-suite for use by less sophisticated trading entities limited to four foundational transactions (PO, Forecast, Invoice, and ASN).
4. Create a File Transfer solution for non-RosettaNet and non-EDI business transactions.

Early on, Intel made the conscious decision to remove any RosettaNet knowledge and routing information from back-end systems. Instead Intel created a re-usable Middleware Services Platform (MwS) that provides the "glue" between Intel back-end systems and the public gateways used for transactions with our TE population.

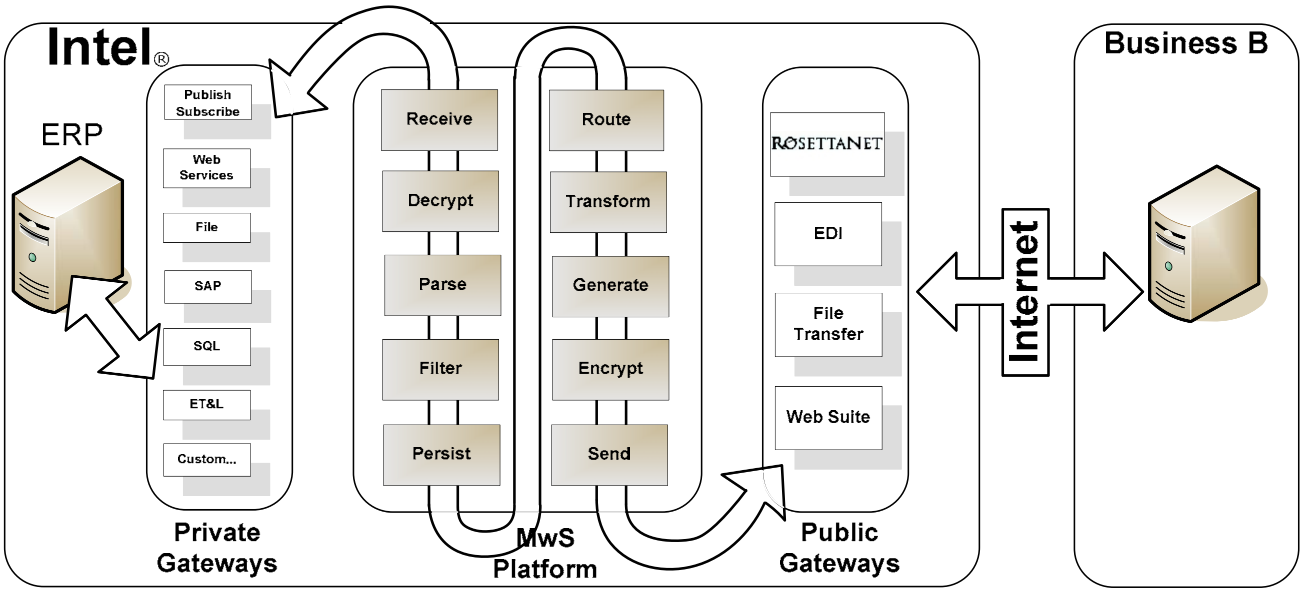


Figure 2: Intel implementation of RosettaNet B2B infrastructure

There are three primary benefits from this architecture:

1. Back-end systems do not care how the data are presented to the TE.
2. TEs are able to migrate from one delivery mechanism to another with only data changes on the Intel side.
3. Intel back-end systems can evolve and change without affecting the existing and growing TE population.

A secondary benefit is that within the MwS, Intel is able to add or change existing gateways seamlessly. An example of a new gateway we are preparing is the RosettaNet-compliant Multiple Messaging Services (MMS) gateway. Once this is added, it will allow RosettaNet payloads over Web services.

The MwS platform is a highly secure, fault-tolerant set of systems capable of scaling out to meet ever-growing transaction volume. Intel chose a business process orchestration engine from a major vendor as the foundational software for our MwS platform. This software contains the necessary tools to enable most of our functionality, and it is also highly adaptable–allowing Intel to add new features and functionality very quickly.

The creation of the MwS platform has created a core system development competency at Intel, that of moving data between systems and companies. The group with this focus develops and supports a wide variety of data exchanges. This allows the various core ERP systems developers to focus on their systems and not be encumbered with the complexities of standards and protocols outside their core competency.

To date, we have developed almost sixty separate business solutions connecting over 200 TEs. We process in excess of 200,000 transactions daily–and have never lost data.

To reduce the need to manage all forms of transmissions and choreographies, Intel has enlisted the help of hub providers for some specialized transmission types. These hubs are in effect an extension of Intel and have provided a cost-effective solution for the following:

- Converting TEs running on EDI to RosettaNet. This will allow us to end of life (EOL) our EDI infrastructure without requiring all TEs to convert to RosettaNet.
- Offering File Transfer Protocols (FTP) at a fraction of the cost of deploying and managing our own servers for these protocols. This has allowed us to standardize on a single technology (Web Services), greatly reducing complexity and support costs.

Savings incurred by use of this B2B platform approach are as follows:

- Application development/re-use.
- Shared infrastructure.
- Support/Maintenance.

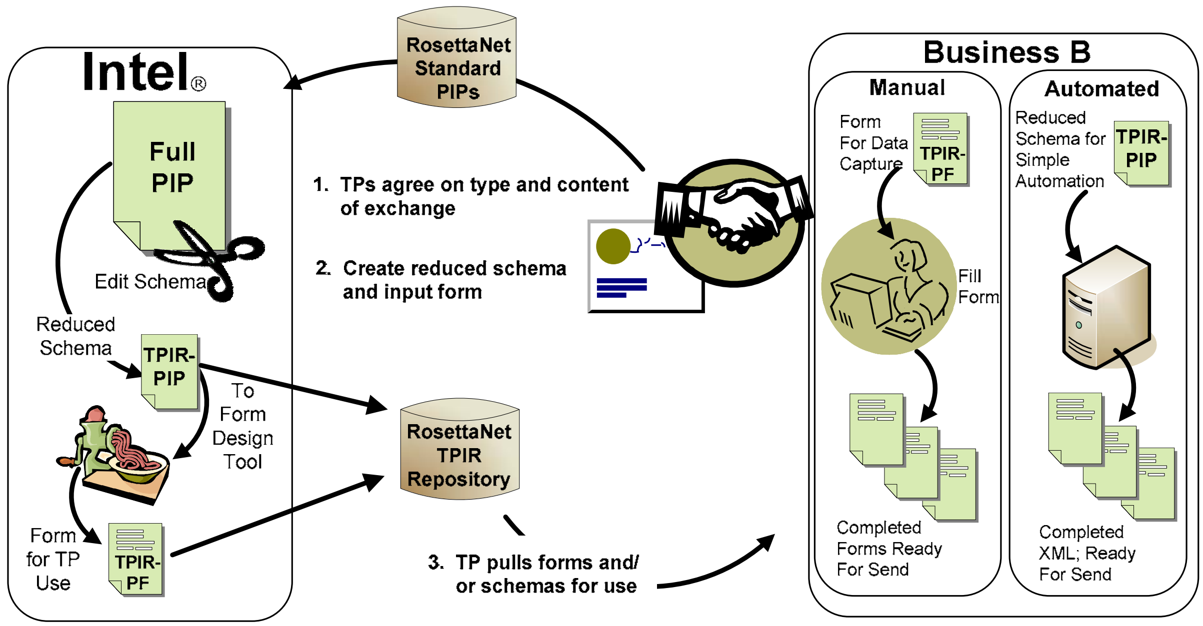


Figure 3: RAE process model (High Level)

Next Generation B2B

Intel has always been in the forefront of promoting and developing new technologies including Web Services. We are actively engaged in developing a Web Services infrastructure in our IT environment to promote Web Services protocols as the building blocks for future RosettaNet standards.

Large multi-national corporations in the high-tech industry have broadly and successfully adopted many of the RosettaNet standards. However, SMEs have been very slow to take advantage of the RosettaNet standards. These are their common reasons:

- It costs too much to develop RosettaNet infrastructure.
- They have no expertise to integrate RosettaNet PIPs into their private processes.
- The time required to implement RosettaNet doesn't meet business needs.

To address these issues Intel formed the RosettaNet Automated Enablement (RAE) program in RosettaNet. The RAE program identifies the various types of TE groups where RosettaNet adoption is low, defines the business constraints that have hindered adoption of RosettaNet, and goes on to define the solution to those SME constraints.

The technical intent of the RAE program is to facilitate RosettaNet usage among a broader cross-section of the supply and demand base, without requiring a substantial time or financial investment. The program will accomplish this by augmenting current RosettaNet technology with new methods, processes, and PIPs that have been converted from Document Type Definition (DTD) to XML Schema format. RAE solutions can then be used as one of a portfolio of technologies to provide lightweight B2B connectivity.

The high-level RAE process flow is illustrated in Figure 3. This process model illustrates how the various RAE components interact with each other to provide a comprehensive SME solution.

A Multi-National Corporation (MNC) can create a Trading Partner Implementation Requirement-Partner Interface Process (TPIR-PIP) by constraining the RosettaNet PIP standard. An XML-based schema TPIR-PIP is created by an MNC by using one of the new XML-based schema RosettaNet PIP standards as a starting point. The RosettaNet PIP standards can be refined by a MNC to constrain or limit PIP content to remove ambiguity and provide clarity for how, specifically, an MNC wants to conduct e-Business with its TEs. The TPIR-PIP is a machine readable document that can then be used to auto-configure the TE's gateway.

Any given PIP may be used between companies for a number of business functions. For example, Advanced Shipment Notification (ASN) may be used as notification of inbound raw materials shipments, outbound shipments to customers, shipments to contract manufacturers, and shipments from contract manufacturers to third-party logistics providers. As each of these functions serve a different business requirement, each has different messaging requirements. The TPIR-PIP allows a company to constrain the community PIP definition and also allows them to assign a unique identification to each of the TPIR-PIPs they create. At runtime, the TPIR-PIP identifier is embedded in the message header that enables the TE to understand the unique role a particular PIP is being used for.

For TEs that do not have the expertise or business ROI to integrate RosettaNet gateway processing with their back-end systems, or in cases where a TE does not have an automated ERP system, RAE provides a new RosettaNet capability that allows the TE to browse the XML message via a form. This is enabled through the Trading Partner Implementation Requirement-Presentation Format (TPIR-PF) specification. The TPIR-PF provides the presentation metadata needed to render the message in the absence of integration with an application. TPIR-PFs allow an SME to view PIP information and/or manually enter data (as an alternative to integrating the TPIR-PIP with their back-end system).

RosettaNet did not develop a whole new standard to describe the presentation metadata. Instead they selected PDF/XPF, an open standard adopted by major vendors specifically because of its widespread use, stability, and broad capabilities. RosettaNet does not prescribe which tools can be used to implement the TPIR-PF, so many will use widely available and freely distributed, PDF/XPF-compliant tools from well-known vendors to display and respond to the messages.

The first step in developing a TPIR-PF is to define the TPIR-PIP. The TPIR-PF form design session binds the TPIR-PIP schema with the TPIR-PF form that is being designed. This is an important point because the form takes on the same constraints defined in the schema. If the TE tries to enter a value in the form that does not conform to the schema, an appropriate error is returned and the SME is prevented from sending the message.

The RAE specification also defines a Registry interface specification. The Registry provides a TE repository that stores both TPIR-PIPs and TPIR-PFs and can be accessed by any trading partner that wishes to conduct e-Business with an MNC. The Registry allows TPIR-PIPs and TPIR-PFs to be posted, stored, and retrieved. All TPIR-PIPs and TPIR-PFs are under version control within the Registry. The Registry provides for the automated provisioning of the TPIRs to the entitled subscribers of the TPIRs.

SUMMARY/KEY LESSONS

Intel's aggressive adoption and continued influence into the RosettaNet consortium has reaped a tremendous ROI by making Intel's business with external TEs highly automated and very agile.

One of the key things we learned from using RosettaNet is that RosettaNet has enabled new business processes that could not have been done without using a standard process and supporting message. Other items that affected the value of using RosettaNet include the following:

- Leveraging collateral that others have developed.
- ROI is based on utilization of capability.
- Consolidate and eliminate redundant B2Bi capability.
- Plan for how to build out (enable) new TEs.
- Senior business stakeholder support essential for initial RosettaNet development and deployment to be successful.

Our MwS platform not only insulates Intel internal systems from TEs but it can also quickly adapt to new standards and technology.

Finally, and most importantly, Intel continues to influence the industry by driving and participating in key milestone programs such as RAE and MMS.

We welcome all industries to actively take part in shaping and implementing the vision of a globally connected supply chain.

ACKNOWLEDGMENTS

We thank our reviewers Tom Sanger, Alan Court, and John Ficca.

REFERENCES

[1] RosettaNet Consortium, *RosettaNet Implementation Framework Core Specification, Version 2.0*, March 06, 2002.

[2] eBG OeS2/3PL paper.

[3] D. Mertz, *Understanding ebXML*, June 1, 2002, IBM Developer Library: http://www-106.ibm.com/developerworks/xml/library/x-ebxml/*

AUTHORS' BIOGRAPHIES

John Cartwright is the forward engagement manager for Intel's Buy-Side B2Bi organization. He has 20+ years of IT development and Supply-Chain Integration experience. He has been on loan to RosettaNet for five years. During his tenure with RosettaNet, John has been the product manger of Cluster 2, Product Information and he has developed Cluster 7, Manufacturing. John has also been the program director for the Direct Ship/3PL, Shipment Booking & Status and the RosettaNet Automated Enablement Programs. He holds a B.S. degree in Business Administration and Computer Applications from CSUF and an M.S. degree from USC in Systems Management. His e-mail is john.m.cartwright at intel.com.

Jay Hahn-Steichen is the forward engineer with Intel's Integration Platform Services group. His primary current responsibility is enterprise manageability strategy and design. Previously, he was part of the B2B infrastructure

design team. He has over 20 years of IT experience. His e-mail is jay.hahn-steichen at intel.com.

Jackson He is the lead architect of Intel's Digital Enterprise Group, responsible for enterprise solution evangelizations and standards developments (OASIS, WS-I, DMTF, etc.). He was the lead architect on loan to RosettaNet from 2001-2002. He has over 20 years of IT and computer technology development experience. He holds Ph.D. and MBA degrees from the University of Hawaii. His e-mail is jackson.he at intel.com.

Thurman Miller is the middleware services organization manager. He is responsible for managing a global organization that supports the engineering, development, and architecture roadmap for Middleware Technologies including B2B and EAI. He has a B.S. degree in Computer Science from the University of Kansas. His e-mail is thurman.b.miller at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

RFID: The Real and Integrated Story

Craig Dighero, Technology and Manufacturing Group, Intel Corporation
James Kellso, Technology and Manufacturing Group, Intel Corporation
Debbie Merizon, Technology and Manufacturing Group, Intel Corporation
Mary Murphy-Hoye, Corporate Technology Group, Intel corporation
Richard Tyo, Technology and Manufacturing Group, Intel Corporation

Index words: RFID, Radio Frequency Identity, smart objects.

ABSTRACT

There is considerable hype and misinformation regarding the use and benefits of Radio Frequency Identity (RFID) technology in supply-chain operations. The reason for this is that while many people talk about RFID technology, very few companies have successfully used it, and fewer have been able to define the actual benefits of using it. Intel's supply network organization formed a unique collaboration with one of our major Original Equipment Manufacturing (OEM) customers to run a proof-of-concept experiment utilizing RFID tags in the combined supply chain of Intel's Malaysian assembly/test facility and our customer's Malaysian laptop assembly plant. They set out to determine the feasibility and operational benefits of this emerging capability.

One of the key goals of this project was to learn about RFID and what implications it will have on how we interact with the supply chain in the future. Another key goal was to partner with Stanford University to identify the necessary elements, as well as construct an industry-first Return On Investment (ROI) model. The project included working with a key OEM customer, Stanford University, MIT, and numerous suppliers of RFID equipment and readers to build a working system in which we could move and track microprocessors through the supply chains of both companies from the back end of the Intel PG8 Test facility onto our customer's consumption point in their manufacturing facility. One of the major benefits was that we were able to forge a new kind of relationship with one of our key customers.

We discuss the following key points in this paper:

- The ecosystem as it exists today.
- Our strategy and approach to understand the technology.
- What we really did with our OEM customer.
- What the value proposition looks like.
- What we learned.

We chronicle the early path-finding project from inception to completion of the shipping of over 70,000 CPUs to our customer in a four-week period. We further attempt to dispel the myths and articulate the realities of what this technology can really offer as it applies to supply network design and optimization in the future.

In the future, we hope to include a proof-of-concept project with a key transportation provider and outline our efforts to align the various entities within Intel that are engaged in RFID experiments. We will also be working on an inventory visibility project with a key boxed product distributor.

INTRODUCTION

As a global, world-class, semiconductor manufacturer, our approach to finding value from Radio Frequency Identity (RFID) technology was quite different at the beginning of 2004 than most other early adopters. While others were being forced to implement RFID as a result of major retailer and government mandates to meet "slap and ship" solutions, we determined that we wanted to explore the idea of how (or if) this technology could help our company and our supply network travelers transform our business practices, and potentially realize improvements across the entire spectrum of our supply networks.

We wanted to test our belief that smart object technology (of which RFID is a part) would offer both operational benefits and the opportunity to make major changes in the supply network ecosystem as a result of additional information. We therefore established a strategy of looking end-to-end across our supply network to determine potential areas of benefit. We then developed short proof-of-concept studies to test potentially high-value areas for the technology, one at a time, using a building-block approach. Over time we will connect the various building blocks into a larger, integrated project that will test the technology throughout the supply chain.

For this specific proof of concept, we chose the segment of our supply chain that included only our final point of manufacturing through distribution out to our customer. We desired to understand several items in this experimental segment. The mundane functions of tag readability, reader and printer reliability, tag, box, and box on pallet placement were the basics, of course; we wanted to know if the technology worked. However, beyond those fundamentals, we were interested in determining if the establishment of a new level of relationship with our customer and the actual integration of information between two companies would provide opportunities for improvements in our planning and operational systems. We did not know specifically where these might be but hoped that they would exist. Of specific interest was the topic of increased visibility across the company boundaries and the potential for benefits as a result.

This "learn as you go" approach was especially valuable with the specific problems encountered in RFID technology itself. In this paper, we chronicle not only our approach but our findings in the areas of technology, operations, value, and business improvement.

THE ECOSYSTEM: WHAT RFID IS

An RFID capability is made up of several technology components that can be embedded into a business environment to improve and transform key supply-chain processes. In this section we describe the fundamental RFID technologies as well as provide insight into the implications and value of integrated RFID system design.

Figure 1: RFID tags

The most basic components of RFID are wireless radio frequency "tags" (Figure 1); these are small devices with a transponder and an antenna that emit data signals when queried/powered by an RFID reader tuned to the tag frequencies.

Multiple frequencies have been identified for RFID that have different purposes, distance capabilities, and costs. Figure 2 provides usage information about Low Frequency (LF), High Frequency (HF), Ultra-High Frequency (UHF), and microwave technologies. HF (13.56 MHz) has been used for some time in various applications. UHF (915 MHz in the US) has been focused on for the recent retail supply-chain mandates and investment due to its distance and cost attributes.

	Frequency	**Distance & Cost**	**Example Application**
LF	**125 kHz**	**Few cm, ¢**	**Auto-Immobilizer**
HF	**13.56 MHz**	**1m, 50¢**	**Building Access**
UHF	**915 MHz**	**7m, 50¢**	**Supply Chain / retail / CPG**
µwave	**2.4 GHz**	**10m, $'s**	**Traffic Toll**

Figure 2: RFID frequency and use spectrum

RFID tags have superior capabilities and benefits over barcode technology in the following ways:

- A power source is not required for passive RFID tags, which is a key defining benefit.
- Non "Line of Sight," high-speed, and multiple reads are possible, changing the nature of how this technology can be applied.
- The RFID Electronic Product Code (EPC) standard extends the UPC standard, by providing for an individual unit to have a specific and unique identity (ID).
- RFID tags can have read *and* write capability for both ID and other data.
- An RFID tag and a battery create an active RFID capability with the attributes of a wireless sensing and communication device (e.g., sensor/mote).

RFID Readers/Printers

RFID readers are placed at designated points along the supply chain (such as when arriving at or leaving a distribution center). Readers send an RF signal to power and activate the tags, process the signals, and receive data. Reader-collected data can then be filtered and proliferated to product information databases and business services. RFID readers can also write to a read/write-capable tag, changing the data on the tag during their lifetime. RFID printers are also available to print information on tag-embedded paper labels as well as write to the RFID tag silicon.

It needs to be pointed out that while RFID reader/printer capabilities are evolving rapidly, the UHF RFID industry and technologies are still immature. The onus rests with the RFID user community to define well-architected, cost-effective RFID integrated systems, and to drive more robust and capable technology components.

RFID Standards

A cross-industry-standards consortium, EPCGlobal, created out of the research work done at the MIT Auto-ID lab, is defining standards for data exchange and architecture. EPC-formatted data on an RFID tag provide an industry-standard way to identify and exchange information about an item. The EPC standard includes a product serial number and can provide links to information such as country of origin or use-by dates.

Integrated RFID Design

RFID is not simply putting readers and printers into a distribution center and tagging boxes and pallets. There are many other ecosystem implications that determine the design (capabilities and constraints) of an integrated RFID system that are more than the sum of their parts.

Ecosystem Implications

The ecosystem implications are as follows:

- **Environment:** RFID readers use radio waves to power nearby tags. Radio waves are subject to interference and can be impacted by devices tuned to a similar frequency (e.g., barcode readers, WLANs, etc.) by the material content of the tagged objects (metal, liquid) and by the form factor of the tagged parts, boxes, and pallets**.**

- **RFID and IT infrastructure integration:** To create a working RFID system, an integrated design is necessary of the specific reader, tag, and printer as well as a reader form-factor design (portal, etc.). The combined interactions of the RFID components in a "setting-based" design will dictate how well the RFID system will perform. In addition, the RFID reader infrastructure requires self-manageability characteristics (e.g., RF profile characterization and debugging, etc.), cross-reader interaction capabilities, and a well-architected alignment to existing computing infrastructure.

- **Application architecture:** There are two aspects of application architecture that must be considered with RFID:
  - The real-time interaction architecture between users, reader, tags, and tagged objects in the specific business environment (e.g., a distribution center) to capture the physical workflow activities.
  - The "middleware" architecture that creates the bridge between the physical workflow and the higher-level enterprise applications, such as the following:
    - Managing the real-time RFID-generated data flow.
    - Filtering and directing information back into the RFID infrastructure.
    - Performing aggregation and communication to/from the enterprise systems.

  Critical components of RFID application architecture are the two-way filters that maximize real-time local decision making as well as enable global strategy and business rule setting.

- **Information management:** RFID necessitates a new approach to information architecture and management due to the creation of a unique identity associated with a specific physical object. Identity and other information about an object can be embedded in the RFID tag and can persist throughout the lifecycle of the object. Perfect visibility into the physical movement of objects through existing business processes enables new associations between the physical workflow and the logical systems. A few of the implications of this include the following:
  - Persistent object-contained information and identity.
  - Highly distributed information structure and physical data storage.
  - Object nesting (e.g., units to boxes to pallets, etc.) and resulting information hierarchies.
  - New information, associations, and aggregation created by physical object proximity.

  This opportunity creates potential havoc with existing information systems. We found that our current systems and processes were not structured to take advantage of this persistent level of information, nor were the data structures even in place to receive this level of information about our products. Although in some cases we desire unit- and box-level traceability, we are not structured in our systems to accommodate this level of data or the quantity of data provided.

  Furthermore, we found that our ability to communicate with our supply-chain travelers, in this case our customers, was lacking in process and in common language.

- **Regulation and compliance**: RFID requires more than procuring the appropriate hardware and software and installing it into a warehouse. Issues not normally considered during an enterprise-based technology implementation must be taken into account.

As RFID UHF takes hold in the United States due to mandates by the DOD and key retailers (Wal-Mart, etc.), additional regulation and compliance work must be done by companies operating global supply chains to ensure seamless RFID-based interaction. The 915 MHz-925 MHz UHF spectrum has been allocated for use in the United States and is driving alignment across the US-based RFID technology industry (reader and tag vendors, etc.).

However, the European Union is aligning around the 868 MHz spectrum and is facing more stringent constraints on the use of the RFID communication protocol (e.g., "Listen, Then Talk"). Many of the key Far East countries (e.g., China, Taiwan, Malaysia, etc.) have not yet aligned to a standard UHF frequency, and in many cases have already dedicated that frequency for other uses. To operate in these countries requires special permits/licenses issued on a site-by-site basis by the local government.

STRATEGY AND APPROACH

Our overall research vision defines ubiquitous computing technologies in the context of our business, and it identifies the types of technology that would make that vision real and valued in our manufacturing and supply-chain processes. RFID (providing easily accessible and unique identity) was a key and early foundational capability necessary to realize our vision. With this in mind, we began the first of a series of proof of concepts that would provide that foundation for the future.

End-to-End

While the value of an end-to-end vision shaped our RFID approach and philosophy, we realized we had to be innovative to make these ideas a reality.

First, we set strategies to shape our first year of investigation.

Second, we created a loose network of RFID investigators across the company and defined a criteria framework for their projects/trials and for aligning them with our end-to-end vision. We hired ethnographers (corporate anthropologists) to spend time in our factories and warehouses documenting the world from the product perspective.

Research seed funding was used to create a shared RFID lab in an Intel distribution center. This was a place that any project team could experiment with a broad suite of technology components and RFID artifacts in a real-world environment. However, we knew a few RFID tags in a lab wouldn't provide us with the insight into improvements, opportunities, and challenges that scale experiments would.

We needed organic experimentation in a high-volume manufacturing environment, grounded in real business problems, with quantifiable ROI and a good fit for RFID capabilities. This would provide a base of trials for us to learn from collectively. We focused on areas with the highest potential business value and began our "building-block approach" with the Intel Malaysia Assembly/Test to OEM manufacturing proof of concept.

Our approach in designing the proof of concept included a number of elements to ensure that the information captured would be reliable and usable within our existing manufacturing operations. These elements are as follows:

Deploy the proof of concept in a high-volume real-world production environment. We wanted to discover RFID's impact on our actual operations. The proof of concept was designed for, and deployed within, the production facilities at Intel Malaysia, the company's largest semiconductor assembly and test facility. In this high-volume manufacturing setting, we could extensively test the effects on manufacturing processes, material flows, information flows, business processes, regulatory environments, and resource utilization. Of particular interest was how it would interact with the wireless environment, material and informational flows, facility layout, and processing steps of producing and shipping large volumes of product.

Collaborate with a major customer. We believed that RFID, as a paradigm-shifting and possibly disruptive technology, could have significant impact on the supply chain beyond Intel's walls. Therefore, we wanted the proof of concept to extend if possible into our customers' operations. After considering a number of potential partners, we decided to work with one customer in order to limit complexity. A major PC OEM with notebook PC manufacturing facilities in Malaysia became that partner.

Focus on interactions. We knew that data would be collected in new ways and expected that the data collected would be richer and more complete than those collected by current methods. Since we anticipated that both the new methods and the enhanced data would present new challenges to existing processes and capabilities, one of our objectives was to explore how RFID changed the interactions between people, product, infrastructure, data, and supply-chain partners.

Measure and document key knowledge. We knew that success in this proof of concept was only a beginning, and that documenting and sharing the lessons learned would be the real benefit as we moved into other areas and other proofs of concept. There were two categories of learning we wanted to document: the mundane functional aspects of readability, writeability, distance of reads etc. and questions regarding frequencies, reliability, and function. We also wanted to understand the integration effects on our systems and processes that additional data would drive and offer, including the opportunity to drive inventory savings as a result of visibility. Finally, we were interested in the less quantitative opportunities offered by the technology regarding our relationship with our customers.

We utilized our engagement with Stanford and MIT professors and students to ensure that we were measuring and documenting our results and knowledge. This has allowed us to share this knowledge with our key partners and has provided the baseline for more advanced analytics to evaluate applications of the technology for the greatest business value in the future.

THE RFID LOGISTICS PROOF OF CONCEPT

The Existing Manufacturing Environment

The proof of concept took place in the facilities of Intel Manufacturing in Penang, Malaysia ("the factory"); in Intel's adjacent Malaysian Integrated Warehouse ("the warehouse"); and in the PC OEM's manufacturing facility ("the customer") nearby. The proof of concept tracked the movement of Intel® Centrino™ mobile technology-based microprocessors from the end of the manufacturing line (where individual processors are inserted into carrier trays), through Intel's warehouse, and finally to the point in the customer's manufacturing line where individual processors are delivered for insertion into notebook subassemblies.

Data

At a basic, functional level, the goal of the project was to provide real-time, location-level data via RFID. Further goals related to the potential value of that data stream were as follows:

- Gain understanding of possible data architectures and data management techniques.
- Identify how RFID can structurally affect existing data flows and existing applications.
- Determine the type of data visibility our customers will desire.
- Determine the type of data Intel would like from its customers.
- Determine the best methods of data retention for later data mining.

Process

One of the key decisions was the selection of the appropriate frequency for the system to operate in. Since the selection of frequency defines the tags and the equipment to be used, and enables the level of readability distances that are possible, the appropriate selection drives the subsequent capabilities. After several experiments with proximity and distance of read capability, we settled on the 910-920 MHz range as the appropriate frequency to meet our distance and read requirements. The specific frequency depends upon the country and the firmware used. The Tyco equipment that we used has two primary ranges. One is an EU version that is best suited for usage in the 865-868 MHz range and the other is equipment for US and Asia with a range of 910-920 MHz. We used the latter.

After we selected the generic frequency, we then worked with the Malaysian government to obtain a special use permit to operate UHF readers in their country. The government identified that we would need to operate in a narrow range of 917.5-922.5 MHz due to the GSM cell phone usage that was just below and above this limited frequency range. After getting the basic equipment, we worked with the vendor to create a firmware (software upgrade for the reader) that ensured the reader was only operating in the frequency range that we were permitted to use.

After agreeing to and enabling ourselves to operate in this frequency range, we went to the various sites in Asia and conducted a spectral analysis of the various facilities where we were going to install the equipment. This was done to ensure that we could safely operate the readers without impacting any of the other wireless devices that were operating around the 915 MHz range. We did find that we needed shielding and antenna tuning to eliminate significant interference with our hand-held barcode systems that used RF to transmit signals in our warehouses.

We had two primary hardware configurations. The first was designed to write/create the tags and the other was designed to read the tags. For the write application we would use one reader, antenna, and PC. We used additional Mylar shielding to help pinpoint and limit the range of the antenna so that we would only write to a tag that was very close to the antenna. These reading stations (portals) were designed to read tagged boxes that were either inside of a metal cage or sitting on a pallet.

Tag selection was guided by the UHF frequency decision and by available space on the intermediate box. The tag had to fit in a 3.5" x 4" space on the end of the box. After

---

® Intel is a registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.

™ Centrino is a trademark of Intel Corporation or its subsidiaries in the United States and other countries.

testing numerous 64-bit and 96-bit tags, we selected a Class 1 "butterfly antenna" tag from UPM Rafsec with a 96-bit memory capacity.

The RFID equipment (antennas, readers, and PC) at each processing station was tied into our factory network and sent data to a server for consolidation. The linkage into the server was an Intel-created middleware. The middleware managed the large amount of data generated by the RFID readers: it collected them, and parsed them, deciding which data were relevant, and it delivered appropriate data to the database. The relevancy of the data was determined by several criteria. Actual movement or change of state was key. Many-time multiple reads are accomplished but there is no movement or state change. We desired to filter that information and only capture and transmit when something of note actually happened. This becomes a laborious process of designing what is desired, testing it, looking at the results, and then redesigning the filters based on actuals.

The middleware also managed the RFID readers and antennas by telling them when to turn on and how long to run. We wanted to push as much intelligence and data management to the edge of the network and use the readers and the middleware to decipher the data that was important and needed to be sent onto the database. This was very critical in that it allowed us to not overload the database with irrelevant or redundant data.

Our customer had a similar network running at their facility. We shared data files via e-mail that were then uploaded to the receiving sites network. In future POCs we expect to share these data directly with our customers as well as integrate them into our ERP tools. Also of note here is that since we were not feeding data directly into our ERP tools, we were simultaneously running two processes throughout the proof of concept. The first process was our standard process used to build and ship our material. The second process enabled all of our RFID-based transactions. With greater integration in the future, we would be able to operate with one process that would update both RFID and other information at the same time.

The products we were tagging were Intel Centrino mobile technology-based processors boxed at 250 units per intermediate box. We had an RFID tag on the intermediate box, as well as on the overpack box. In addition to the boxes, we also tagged the transportation media (metal trolley and shipping pallet) to establish parent-child relationships.

At the physical level, data tags were loaded with a unique identifier. This unique ID was linked in the middleware and database to the specific data characteristics that we tracked for each intermediate box. This data included the following:

Manufacturing information:

- Product code (SKU) number and quantity in a group pack.
- Lot numbers, country of origin, overpack ID, and intermediate box tags.
- Transaction times and locations.

Supply-chain and order information:

- Cart name and/or pallet ID.
- Delivery Note (DN) number/House Airway Bill (HAWB) number.
- Customer Part number/Customer PO number.

The data that were in our database mimicked those which are tracked in our order management processing system. We ran the proof of concept, gathered data, summarized our findings, and began the analysis stage. In this stage we collected and summarized our key lessons learned.

RETURN ON INVESTMENT

As a part of this study, we endeavored to determine if there was a value to increased visibility due to the existence of more real-time data. Dr. Hau Lee of Stanford University had a special interest in the value proposition regarding RFID. He participated in our proof of concept, with special attention to ROI.

There are three areas that Dr. Lee investigated for possible issues regarding ROI. They are illustrated in Figure 3.

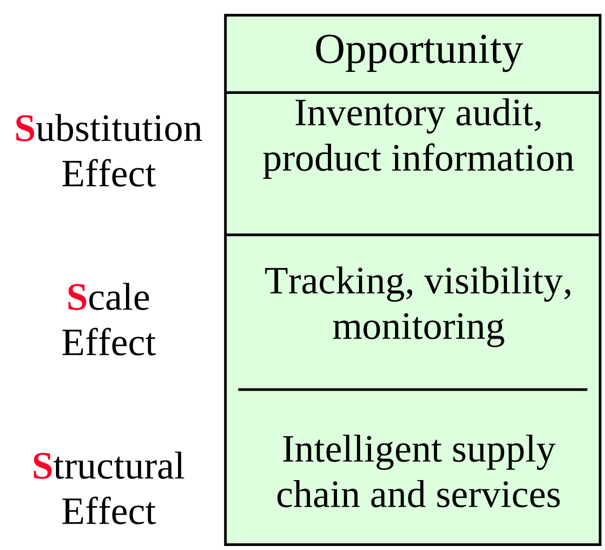


Figure 3: Areas investigated by Dr. Lee

Substitution would imply simply replacing current barcode or other data-management practices with RFID and due to the ease of data capture, this would provide payback. The work done on this POC and work that Intel Solutions Services did with Tyson foods and other clients clearly indicates with data, that substitution alone does not provide a good ROI for RFID. There is about 12 to 15 cents per read that can be saved and regardless of tag cost or level of infrastructure, this alone is insufficient to pay for any level of RFID implementation.

Scale is a much more likely opportunity for savings. This implies that one would install RFID across a wide range of products or processes. However, it is still not apparent in our research that scale will provide adequate payback to offset the investment. This is really substitution across a larger base.

It is believed that structure is the most likely area for a good ROI. Structure would imply that fundamental changes in the manner and process of the business can be enacted via the existence and use of more timely and abundant data. An end-to-end expansion should be able to create a much greater value and benefit as follows:

BY NETWORK:

> when RFID technologies are deployed throughout the supply network, so that smart objects can be traced throughout the network.

BY TIME:

> when RFID technologies deployed on a product can manage that product throughout its product lifecycle, from product generation all the way to product return and disposal.

One specific example of this premise was evaluated in detail. The opportunity to reduce safety stock inventory as a result of more frequent views of the actual consumption was tested and modeled. Dr. Lee has established a relationship between an increase in visibility and the decrease in safety stock required in a company. The formulas are illustrated in Figure 4.

Downstream visibility	*Safety stock at Mfg DC*
No	$k\sigma$
Yes	$\sqrt{1-f}\,k\sigma$

f = fraction of revealed demand over replenishment cycle

Figure 4: Value of visibility

Using this relationship, Dr. Lee was able to develop a potential savings due to visibility for a retail sample illustrated in Figure 5.

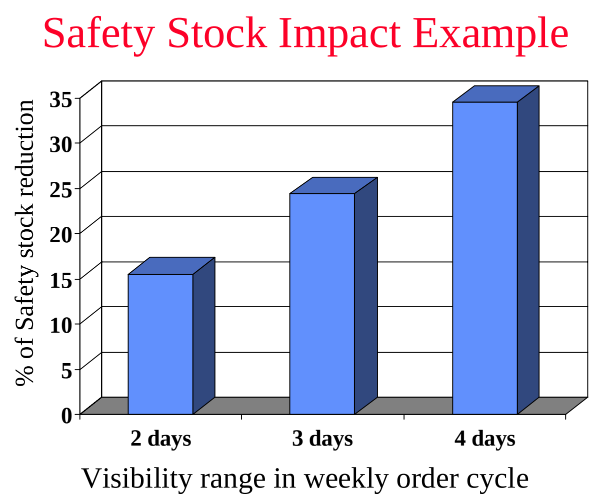


Figure 5: Safety stock impact of visibility

At this time, we are continuing the work with Dr. Lee to better quantify the opportunities for savings. There is much more work to be done, but this approach appears to provide a good opportunity for long-term savings and benefit. This does not yet address the more systemic changes in supply-chain organization that may be possible. This will be a key topic of interest in our next series of tests.

WHAT WE LEARNED

Throughout our first year of discovery, and especially with this particular RFID project, we identified several key lessons. We learned a great deal about the actual discrete functioning of the technology regarding tags, readers, frequencies, country regulations, and placement of readers, tags, and support equipment. We also discovered much about current and future possible business impacts in our supply network. Many of these lessons will help us to shape our future experiments to gain even deeper knowledge. We are confident that the technology will provide us with more timely and granular detail about our products and processes. We have yet to discover exactly how to take full advantage of this detail in our supply-chain operations. Our end-to-end thinking is helping us to formulate the next steps and experiments and take fuller advantage of the lessons we have learned so far.

Following is a discussion of some of the key lessons we learned.

Lesson 1: This New Technology Must Be Learned in the Real World

This technology must be learned by working with the actual readers, tags, and products in high-volume production environments and by working on real problems that are unsolvable by current methods. By finding real problems we could engage the key players in the supply chain and begin to understand the true value and limitations of RFID. We found significant differences in read rates, ease of use, and consistency of data in the actual production environment than we had in our lab. One issue of note was that in our lab, we made sure to gather quantifiable read rate data. In the production environment, our operators continued to jiggle the pallets around until they got full reads and the data captured did not illustrate how much time this took.

It became clear to us as we both designed our experiments and ran them, that our approach needed to be more holistic to fully understand the following: 1) the business context including problem definition, boundaries, and impact; 2) the physical environment including workflow, people, process, product, data, and layout; 3) the "logical" system and infrastructure environment including existing system interactions; and 4) the RFID component technologies *and* the specific integrated RFID solution.

Every aspect of the physical environment had to be taken into account to maximize the accuracy and reliability of the RFID technology as follows.

- On the object: placement, orientation, form factor and materials, and spacing, plus our tags on the boxes.
- Around the object: casings, carriers, transport materials, structure, and hierarchy.
- With the object: physical proximity and relationship associations.

We did learn a great deal about physical placement of tags, difficulty in reading, consistency of read, and capabilities of the equipment. Our takeaway was that each environment, box type, product, and tag type needs specific testing in the environment to refine the particulars of the circumstance, and that lab testing or carryover from previous other products is not applicable or sufficient.

Lesson 2: New Types Of Expertise Are Needed

From the beginning of the RFID effort at Intel, we had to create and apply new methods and techniques for the projects, ranging from how we capture our business requirements, to designing the data structures and applications, to building and testing robust RFID systems. In addition, we had to address new requirements and include unexpected tasks, much different from a normal enterprise application project.

Ethnography

From the outset, we were sure that simple substitution of RFID for barcode would not be a profitable endeavor. We were looking for more fundamental opportunities. We decided to attempt to understand how unique identity at the object level could or would affect our fundamental business processes. We therefore set out to capture the requirements in a new way, from the product point of view.

Intel factories are very capital intensive, and as a result the factory processes and applications are designed with equipment optimization as a priority. If we were going to use RFID technology to tag our products, we needed to understand how these objects physically moved through the supply chain, and what interactions occurred between products, people, equipment, and data (both paper and electronic). This new view of an object's workflow was captured through the use of *ethnography* (e.g., corporate anthropology), where anthropologists studied the factory or warehouse and captured information from the product's point of view. We took our corporate anthropologist into the factories in Penang to study the process in this manner. This gave us new insight into the way the factories operate and also helped us identify the best insertion points for the technology. From this study we decided that the process flow from the end of our testing operations through the boxing, warehousing, and out to the customer would offer the best opportunity to provide us and our customer with real-time availability information that we hoped would be the most useful to changing our management and delivery processes.

Collaborative Supply-Chain Conversations

When RFID project work began with the PC OEM, the level and type of conversation between us changed in very interesting ways. RFID eliminated some of the barriers in typical supplier-customer interaction, by providing a new way for us to examine and improve both the logical processes and physical movement of our objects. A view into the relationship between product planning, manufacturing, logistics, transportation, and consumption at the object level, brought us new insight into our product's lifecycle. There were several areas in which simply doing the proof of concept allowed the planners and logisticians to talk directly without the purchasing and sales organizations filtering information in both directions. Focusing on the product movement versus the commercial terms allowed both companies to derive the truly best processes for both. Several misconceptions regarding needs for labels, codes, selection criteria, etc. were cleared up and modified independent of the actual RFID technology, simply as a result of this increased and different communication.

A fairly mundane but important example of this type of conversation was illustrated by the tracking of an urgent order during the proof of concept. With data, we were able to track an urgent order pulled in from 3 p.m. to 10 a.m. at the manufacturing facility. This order was a rush to the customer. The actual tracking indicated that although it left Intel at 10 a.m., it did not arrive at the customer 30 minutes away until 2 p.m.. (The truck stopped for a two-hour lunch and prayer.) It was then placed on the shelf and not used until later the next day. The ability to track this transaction with data created a dialogue about rush orders and the need and criteria for them when our facilities are a half hour apart. This process will change even prior to use of RFID.

3D RF Characterization and Integrated Testing

As we designed our integrated RFID system, it became evident that new levels of characterization and integration would be necessary for a successful implementation. The first step was to create an RFID testing lab inside of a working distribution center. This gave us a testing environment where we could analyze a variety of RFID tag and reader technologies and frequencies to understand individual components capabilities but also determine the best pairings of tags and readers.

In-situ RF characterization was required next to tune the RFID tag-reader system by evaluating the specific environment's impact (e.g., interference, absorption, etc.) on the RF signals and the resulting system performance. The RF signal could be affected by the object materials, its containers (metal cages, desiccant bags, etc.), current infrastructure (e.g., barcode readers), and the physical layout of the factory or warehouse.

Then the physical objects could be characterized and tuned with the RFID system by generating 3D mappings of the interactions of the reader radio waves with a tagged pallet or transport container. At the same time, we had to comply with international government regulation and RF licensing restrictions. This required on-site RF assessment to verify our RFID system complied with Malaysian spectrum specifications.

Lesson 3: Expect A Long-Term Disruptive Impact On Enterprise Systems and Architectural Landscapes

In the course of our work, it became obvious that we were collecting large amounts of data on the product reads, about each transaction, not only at each leg of the supply chain but multiple times. It further became obvious that if we desired to design an end-to-end, object-centric, RFID-integrated database to cooperate with the tag information, we would need a very different architecture from our current SAP centralized environment.

With data resident with and owned by the individual unique objects, highly distributed information architectures become possible. The idea that all data must be centrally located so decisions can be made does not appear scaleable to us in an extended RFID/Smart Object environment.

In the near term, the use of stopgap middleware and appliance software to handle data filtering tasks appears to work well. We developed such a solution for our project and are evaluating various commercial products for our next efforts. Over time, we expect that the number and diversity of nodes will multiply exponentially, and the need to make localized real-time decisions will increase.

Lesson 4: Structural Change Opportunities

We approached each RFID project with the mindset of making structural changes to our processes instead of just substituting a new technology for an existing one. This made each project much larger than just replacing our existing barcode steps with RFID. Through detailed ROI analysis, we also determined that using RFID to capture *only* the same data that barcodes do today, will *not* bring significant value to a company. Improving productivity, extending visibility, improving underlying operational parameters, and making real-time decisions through the use of RFID and other Smart Object technology offer the best opportunity for payback. The work of Dr. Lee on this project leads us towards the types of benefits that may be possible but are as of yet not fully quantified. More work needs to be done in this area but we are confident that changes in the area of structural business processes will yield the most benefits.

Through this RFID project we also identified key areas where RFID technology could be used to address current business process- and system-related issues across our product's lifecycle. These areas include the following.

Manufacturing
RFID encourages/enables unit-level routing instead of our current lot-level processes. This finer level of routing visibility provides benefits that range from tighter control of material on the manufacturing floor to decreased throughput time, as individual units can be routed instead of having to wait for a full batch to be made. It also provides a simpler method of creating and tracking product routes.

Logistics
In addition to revamping and eliminating several steps in our processing flows, we believe that RFID can also decrease our workload and improve our productivity. We also identified potential benefits related to cycle counts and an ability to get material through our processing steps quicker than is currently possible.

Customers
With RFID we see several improvements in our customer interactions. The technology enables clean unit-level visibility to the material being received, and eventually consumed by our customers. This greater level of detail can be fed into our demand systems to better support our customers.

Productivity
Currently, operators spend a lot of time printing new labels, verifying data, and barcode scanning the same information numerous times into the various application interfaces. These steps could be simplified or eliminated with RFID, increasing productivity and reducing costs.

Our future work will address putting hard numbers against these areas of learning and opportunity.

Lesson 5: Getting to SCALE–Readiness Across the Supply Chain and the RFID Technology

There is a large difference between technology trials and world-wide, end-to-end deployment and support. The technology standards and RFID system infrastructure are still in their infancy so this can amplify integration and characterization issues. We are beginning to address the many tasks needed to ready the ecosystem and eliminate obstacles to widespread adoption and the value that comes from scale and critical mass.

So, with those things in mind, what does it take to “get ready”? In order to demonstrate the complexity, let us look at one facet of readiness: **standards**. What standards need to be in place to get ready for implementation?

When an RFID project is in the initial planning stages, several selections need to be made:

- Tags that operate in a particular frequency.
- Portals, readers, writers–all the devices that define the integrated RFID system design.
- Information that will be written to and acquired from the tags.
- Application architecture and interfaces (to the RFID systems and legacy enterprise applications).
- The deployment countries and associated frequency standards permissible in those countries.

A broad set of RFID standards are needed, driven from an integrated usage perspective. It is desired that the tags be standard, and thin and flexible as possible. EPCGlobal has made excellent progress in framing the standards questions and driving cross-industry solutions, and companies are aligning to these. But there is still much work to be done to finalize and publish these standards.

However, once industry standards are set, a company still has several more readiness steps to complete.

Once a company has identified its suppliers and customers that will participate in an RFID initiative, the following conditions need to be met:

- Suppliers and customers are ready *and* have implemented the use of key RFID standards in their processes and systems.
- RFID standards must include definitions for content, structure, mechanism, interaction, and interface protocol, and all these must be in place and agreed upon by both sides.
- Once external standards are defined, a company then needs to determine the set and level of standard definitions it will apply and enforce internally for equipment (including readers, tags, systems, databases, etc.), approved vendors, application and data architecture, and strategies for interaction and engagement with customers and suppliers.

This is just a small sample of definitions required for RFID standards readiness.

There are still many things to do to realize the full promise of this technology. However, that does not prohibit us from learning and getting ready across our end-to-end supply networks.

CONCLUSION

We are seeing great promise and signs that the RFID and future upcoming sensor network technologies will help to change the way we think about our manufacturing processes and the interactions with our people and our customers. With only a small experiment, working with one customer, we learned a lot and improved our processes. We know there is much yet to be learned, but now we can build on the knowledge we have gained. We will continue to conduct our proof of concepts on the factory and supply-chain "floor" with real products, real systems, and real people interacting. We will prepare ourselves to contend with and manage the expected volumes of real-time data and attempt to determine how to best use these data to realize the promise of enhancing our supply-chain visibility. We have also learned a great deal about the need for and the characteristics of those standards required to rapidly deploy this technology in a scale (world-wide) approach. We will be working to influence Intel and others to drive for internationally accepted use, data, and frequency standards in the near future.

And finally, while RFID may seem to be a fairly simple and innocuous technology on the surface, a wide range of issues and choices need to be explored and resolved for its successful, wide-scale deployment. Concerns over threats to security and privacy (both real and perceived) are driving legislative action that could raise even greater hurdles to global RFID deployment.

ACKNOWLEDGMENTS

This project could not have been done without the contributions of Dr. Hau Lee of Stanford University, Jim Rice of MIT, Chris Stroop, Wayne Bynum, Kerry Hudson, Bee Nee Thiang, Duncan Lee, and KL Choong, all of Intel Corporation. We also thank MIT MLOG students, ADC managers and technicians, and Scott Shawn, Technical Writer.

You can read additional details about this project in the *RFID Journal* and the *Supply Chain Management Review*.

REFERENCES

"Supply Chain Management with RFID," Dr. Hau Lee, Stanford University.

AUTHORS' BIOGRAPHIES

Craig Dighero is a program manager in the New Business Engagement Group and is focused on acquisition integration and RFID usage in Intel's supply chain. Craig has been with Intel for ten years and has held various supply-chain and acquisition integration roles over this time. He holds a Bachelor's degree in Business with a minor in Packaging from Cal Poly. His e-mail is craig.dighero at intel.com.

James R. Kellso has been with Intel 18 years with 16 years of industrial engineering consulting experience prior to Intel. His primary technical interests involve creating mathematical ways to represent the complexity of supply-chain operations and decisions. Jim holds a Bachelors of Industrial Engineering degree from the University of Michigan and has been a licensed professional engineer for over 30 years. His e-mail is jim.r.kellso at intel.com.

Debbie Merizon is the manager of Supply-Chain Strategy. She has been with Intel for 15 years. Her primary technical interest is enabling capabilities that drive efficiencies in supply-chain management and leading Intel's RFID Program Office. Debbie holds an undergraduate degree in Finance from California Polytechnic State University and graduated from San Jose State University's Executive MBA program. Her e-mail is deb.l.merizon at intel.com.

Mary Murphy-Hoye is a senior principal engineer in Intel's System Technology Lab currently focused on Smart Object research for emerging ubiquitous and self-aware computing applications and architectures. She currently sponsors a joint research project with Stanford and MIT, focused on exploring Smart Objects for Intelligent Supply-Chain Integration through case study work. She received her B.S. degree in Mathematics from the University of Arizona in 1979. Her e-mail is mary.c.murphy-hoye at intel.com.

Richard A. Tyo is a research integrator and TME technologist. Rick has been with Intel for seven years with 16 years of software, systems engineering, and research experience in the aerospace and automotive industries prior to Intel. His primary technical interests involve identifying promising new and emerging technologies and advancing them into mainstream operations. He has most recently been providing leadership in technical standards, government regulations and licensing, and RF engineering for the TMG RFID Program Management Office. Rick holds a Masters in Industrial Engineering and Engineering Management earned while working for NASA, and a Bachelors of Electrical and Computer Engineering from the University of Cincinnati. Rick is a licensed professional engineer. His e-mail is richard.a.tyo at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

For further information visit:

developer.intel.com/technology/itj/index.htm