85

An Application of Lean Principles within a
Semiconductor Manufacturing Environment

By

Roy C. Wildeman, Jr.

Bachelors of Arts in the Engineering Sciences, Dartmouth College (1999)
Bachelors of Engineering, Thayer School of Engineering at Dartmouth College (1999)

Submitted to the Department of Mechanical Engineering and the
Sloan School of Management
in Partial Fulfillment of the Requirements for the Degrees of

Master of Business Administration
Master of Science in Mechanical Engineering

In Conjunction with the Leaders for Manufacturing Program at the
Massachusetts Institute of Technology
February 2005

Signature of Author_____
January 14, 2005
MIT Sloan School of Management
Department of Mechanical Engineering

Certified by_____
Stephen Graves, Thesis Advisor
Abraham J. Siegel Professor of Management

Certified by_____
Daniel Whitney, Thesis Advisor
Senior Lecturer, MIT Engineering Systems Division

Accepted by_____
Margaret Andrews
Executive Director of Masters Program
Sloan School of Management

Accepted by_____
Lallit Anand
Chairman, Committee on Graduate Studies
Department of Mechanical Engineering

BARKER

1
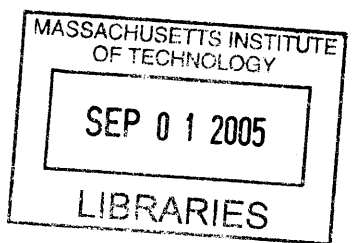
This page is intentionally left blank

An Application of Lean Principles within a Semiconductor Manufacturing Environment

By

Roy C. Wildeman, Jr.

Submitted to the Sloan School of Management and the Department of Mechanical Engineering on January 14, 2005 in partial fulfillment of the Requirements for the Degrees of Master of Business Administration and Master of Science in Mechanical Engineering

Abstract

Intel Corporation's Fab 23 is committed to implementing lean manufacturing to reduce their production cycle times and cost. This thesis is focused around the development of the principles of lean that are most relevant to Intel's complex manufacturing flow and then the application of these principles to improve the operations in a focused area, the Sorting floor. Direct examination of the work in Sort raises the awareness of inefficiencies from overproduction and inventory; viewing this work as a series of structured activities, customer-supplier connections, and simplified flows further crystallizes the need for a structured approach towards WIP management. A pilot implementation of a CONWIP control of inventory demonstrates reductions in cycle time variability and provides a foundation for further improvements. In conclusion, the challenges experienced with changing the manufacturing systems in Sort were largely organizational and likely to be seen in many other operational areas at Fab 23.

Thesis Supervisor: Stephen Graves
Title: Abraham J. Siegel Professor of Management

Thesis Supervisor: Daniel Whitney
Title: Senior Lecturer, MIT Engineering Systems Division

This page is intentionally left blank

Acknowledgements

This page is intentionally left blank

Table of Contents

1. INTRODUCTION

Increasingly, American companies are off-shoring their manufacturing operations – a trend which creates an uncertain economic climate domestically alongside lots of political controversy. Amid the shakeup of globalization, one thing seems certain: operations that are plagued with inefficiencies can and will continue to be replaced by those which are more nimble and efficient across an international playing field.

To remain competitive in the years ahead, many manufacturers must recognize their competitive reality and continually seek to remove the waste that has accumulated in their processes. In many ways, this thesis research is a story about one such entity – Intel's Fab 23[1] in Colorado Springs, Colorado. In benchmarking other world-class manufacturing systems, Fab 23 is attempting to leverage principles of lean manufacturing towards a greater competitive advantage that will sustain their long-term survival in an increasingly competitive world.

Many pundits question the applicability of lean to the kind of complexity found within a semiconductor production line. In response, this thesis is centered on Fab 23's interpretation of the *principles* of lean – that is, fundamental laws that govern process excellence – and the general applicability of these principles through a targeted study of the facility's sorting operations. By experimenting with lean thinking in Sort, this research provides a deeper understanding of the types of continuous learning, benefits, and barriers that are unique to a lean initiative within a semiconductor environment.

1.1 Thesis Structure

A brief overview of the thesis structure is as follows:

Chapter 2 describes the history of Fab 23, the current business climate for flash memory, and summarizes the strategic operational goals the fab needs to achieve in order to remain competitive in the future.

Chapter 3 provides an overview of the lean initiative at the fab, called Manufacturing Excellence (or mX). The principles of mX will be described in detail along with a summary rationale for focusing this research on the facility's sorting operations.

Chapter 4 is an overview of the Sort 23 operations, including the high-level process flows, management structure, and economic drivers of the group.

Chapter 5 is a deeper study into the operating characteristics and latent wastes within the sort production process. Both direct observation and statistical analyses are used to gauge the level and severity of these different wastes while creating tension for change.

Chapter 6 presents several specific hypotheses for change towards Sort's ideal state. Each opportunity is viewed through the lens of the mX rules i.e. seeing work as a series of

---

[1] Fab is a term for a fabrication facility, or factory, which manufactures semiconductor products.

structured activities, customer-supplier relationships, and simplified flows that can be continually improved upon towards the ideal state.

Chapter 7 describes the methodology and results from an implementation of one of these aforementioned opportunities, namely a CONWIP control of inventory to promote more consistent and streamlined cycle times across Sort's product flows.

Chapter 8 highlights some of the key challenges experienced with changing the manufacturing systems in Sort. The research concludes that many of the barriers to mX remain largely organizational and are likely to be seen in many other operational areas at Fab 23.

2. BUSINESS BACKGROUND

"I think it's safe to say we're emerging from the worst IT/telecom recession in our lifetime. Ahead of us are a set of opportunities unparalleled in our past.'"
- Intel Capital President John Miner (Fackler, May 2004)

2.1 Fab 23 History

Fab 23 in Colorado Springs is Intel's sole manufacturing source for a mature flash memory technology. Flash memory chips retain memory even when a device is shut off, and they are commonly used in cell phones, digital cameras, and PDAs. Originally, Intel purchased the factory from the Rockwell International Corporation in February of 2000 to rapidly add more manufacturing capacity in support of the booming demand for Intel products (Intel, June 2004). Representing an investment of $1.5 billion to manufacture not only flash memory but also logic devices, Fab 23 quickly set a new Intel record for the fastest start up for a 200mm manufacturing site. In support of this rapid ramp, the Fab 23 employee base was expected to grow over 1,000 as the factory became fully loaded.

With the unexpected technology recession in late 2000, Fab 23 was suddenly facing a radical change in external environment. In response to the market downturn, the fab embarked on a cost-cutting campaign, dubbed PnL (for Profit and Loss), to coordinate and align functional departments towards operating the factory as an autonomous business entity within Intel's Fab and Sort Manufacturing (FSM) division. By internally focusing on their cost accounting practices, management at Fab 23 was able to drive more business-centric thinking across its organization, focus efforts on controllable cash flow as a ballast against their fixed depreciation stream, and ultimately drive their cash cost per 3K wafer start [2] down 40% over a period of 14 months. The PnL initiative was a major factor in enabling the fab to remain viable as a single-technology supplier during the economic recession. Significantly, the period was also noted for many lay-offs, infrastructure shut-downs, and other painful spending cuts felt across both Fab 23 and other Intel facilities.

2.2 Business Environment

The semiconductor industry has been steadily recovering over the last two years; the three-month average of worldwide chip sales rose to $16.94 billion in April, 2004 - the highest monthly level since January 2000 - and a growth of 36.6% from the previous year (Andrey and Greenagel, June 2004). The demand for flash memory has been strong as well, with 2003 sales volumes up almost 27% from the previous year (Kanellos, November 2003).

The overall flash market is divided into two main technologies: NAND (for "NOT-AND" - a reference to how data is retrieved) and NOR ("NOT- OR"). Intel, Advanced

---

[2] A cost accounting metric used across Intel's Fab Sort Manufacturing division. 3K Wafer Start is a standard measure of fab volume, indicating 3,000 wafers start processing over the course of a week.

Micro Devices, and Fujitsu produce primarily NOR flash devices, which have lower density, lower writing speed, and higher reading speeds. NOR devices are also less prone to data corruption and are typically used to store information on cell phones, PDA's, and set-top boxes (Kanellos, November 2003). Reciprocally, Samsung and Toshiba together control about 90 percent of the $4.7 billion market for NAND devices, which can hold more data than NOR, are cheaper at comparable densities, have a higher writing speeds but slower reading speed, and are commonly used in digital camera flash cards and MP3 players (Yoon and Sorid, May 2004).

NAND is the faster growing technology, reaching 40 percent of the overall flash market from its level of 10% two years ago ("Memory", March 2004). The shift corresponds to a stronger relative demand for digital cameras and MP3 players and has resulted in a turnover in market leadership; Intel had previously led the flash memory market with its focus on NOR technology for over a decade but, in the third quarter of 2003, Samsung, Toshiba, and Spansion (a recent joint venture between ADM and Fujitsu) all overtook the former number one as Intel's market share dropped from 26.1% in 2002 to 13.5% market share at the end of 2003 (Kanellos, November 2003).

Historically, scarce fab capacity worldwide has fueled high margins for the major players in both groups of technology. However, in the spring of 2004 Samsung & Toshiba began slashing their prices to undermine the threat of emerging rivals (Yoon and Sorid, May 2004). Additionally, benchmarking studies revealed Intel flash memory manufacturing lagged all the high-volume manufacturers across both NAND and NOR technologies in terms of cycle time, measured by manufacturing days per mask layer. The relative growth of NAND technology, the aggressive build-up of production capacity by established players, and the potential entry of new players into the NOR technology space all prove to be threats to Intel's margins and share of the flash memory market space.

2.3 Key Takeaways

As a single-technology fab currently producing only a mature NOR technology, Fab 23 must focus on improvements in both cost and cycle time to strategically bolster its competitiveness among both incumbent and new entrant flash memory manufacturers. Additionally, the fab has a production commitment from Intel to manufacture an added communications technology in the near future. In order to compete in this new market, cycle time reductions are extremely critical, as customers in the communications market are likely to demand radically lower lead times than Intel has traditionally been asked to deliver.

Lean manufacturing is a well-established improvement methodology which emphasizes both waste elimination and speed – a good match with Fab 23's strategic objectives of lower cost and faster cycle times. Next, we will cover the nature of lean at Fab 23, characterized by the Manufacturing Excellence (mX) initiative.

3. OVERVIEW OF MANUFACTURING EXCELLENCE (mX)

"Compete Externally & Lead Internally"
- Recurring communication at Fab 23

In early 2003, Fab 23 collaborated with an internal benchmarking group to explore lean manufacturing as a method to not only capture the common thinking and language that made their recent PnL initiative a success but also drive new improvements in cost and manufacturing cycle time. By emphasizing openness to experimentation and *internal leadership* towards operations excellence, Fab 23 strives to ultimately enable Intel's Fab and Sort Manufacturing (FSM) division to *compete externally* in the flash memory market. The evolving lean initiative is dubbed Manufacturing Excellence, or mX, and promotes change along both technical and organizational fronts.

3.1 Technical Goal

The primary technical goal of mX is to understand how 'pull' methodologies can be used to improve the cycle times of their complex operating flow. Most generally, a pull system prevents inventory from growing beyond a specified limit by limiting the release of work into particular segments of a process (Hopp and Spearman, Spring 2004, p. 147). The benefits of such disciplined flow are well documented and typically include:

1.) A means to reduce WIP and corresponding cycle time
2.) Smoother production flows for more predicable output
3.) Improved product quality with faster line velocity and a built-in intolerance for scrap/rework
4.) Continuous reduction in manufacturing costs through the systematic exposure of latent wastes, or "rocks" (Hopp and Spearman, Spring 2004, p. 137)

Pull methodologies have not been widely adopted in the semiconductor industry for a variety of reasons. First, the re-entry nature of semiconductor processing flow creates a highly-complex environment, where multiple process steps can go through a particular toolset and even multiple toolsets that can process a single step. Secondly, the characteristic variability in tool availability becomes a key obstacle, as unpredictable downtimes become a significant challenge when attempting to structure the strict inventory levels and deterministic run times intrinsic to traditional pull systems.

3.2 Cultural Goal

The primary cultural goal of mX is to engage Manufacturing Technicians (MTs) towards systematically eradicating the waste that continually burdens their factory jobs. A cornerstone of successful lean manufacturing systems is allowing workers to display their capabilities in-full through active participation in running and improving their own workshops (*Toyotaproductionsystem.net*, July 2004). In order to realize this kind of change on the floor, work standardization and structured experiments must be accepted norms that enforce coordinated improvements across different shifts.

In the ideal state of mX, every employee at Fab 23 views their daily problems as opportunities for learning and systematic waste reduction. In this vision, everyone also uses root-cause analysis as a normal business practice, and systems-thinking is a proliferating skill across all levels of the organization.

3.3 Implementation Strategy

3.3.1 Lean Rules & Principles (not Tools)

At the start of the initiative, the mX team experienced a few set backs. The creation of an Improvement Suggestion Portal was intended to help engage MT's in promoting waste elimination. However, the tool inadvertently revealed a frustrating scarcity of resources to support the surge in improvement suggestions. Some managers also raised concerns about the addition of a redundant improvement channel which paralleled systems already being used. Additionally, extensive time was invested in a detailed value-stream mapping effort which ultimately led to common-sense conclusions. Frustrated by a lack of progress, the mX team embraced a different implementation strategy which de-emphasizes lean tools and focuses on a shared way of thinking and learning the key principles inherent in lean manufacturing systems.

Prior research on the Toyota Production System has concluded that the recognition and scientific improvement of work activities, connections, and flows throughout an organization's processes is an intrinsic strength behind Toyota's operational excellence (Spears and Bowen, 1999, p. 98). With the counsel of the Lean Learning Center in Novi, Michigan, Fab 23 has formalized this research into 4 rules that govern and direct new ways of thinking about opportunities for improvement (Flinchbaugh, August 2004, p. 2):

1. Structure every activity
2. Clearly connect every customer-supplier relationship
3. Specify and simplify every flow
4. Improve through experimentation at the point of activity toward the ideal state

Here, the ideal state is defined by delivering exactly what our customer wants - exactly when they want it, at the price they want - with zero waste and everyone safe.

Additionally, the mX team has communicated 5 principles which help provide a common vocabulary, lens, and methodology towards the ongoing coaching and support of all mX opportunities:

1. Directly and deeply observe the work in question to understand the current reality
2. Systematically recognize and eliminate waste
3. Systematically solve problems through the scientific method
4. Establish high agreement on the what and how of the proposed change
5. Promote a learning organization that continuously learns, applies, and reflects on its change

These rules and principles of mX are summarized into a "House of mX" schematic, shown in Figure 3.3.

Figure 3.3: The "House" of Manufacturing Excellence (mX) principles.
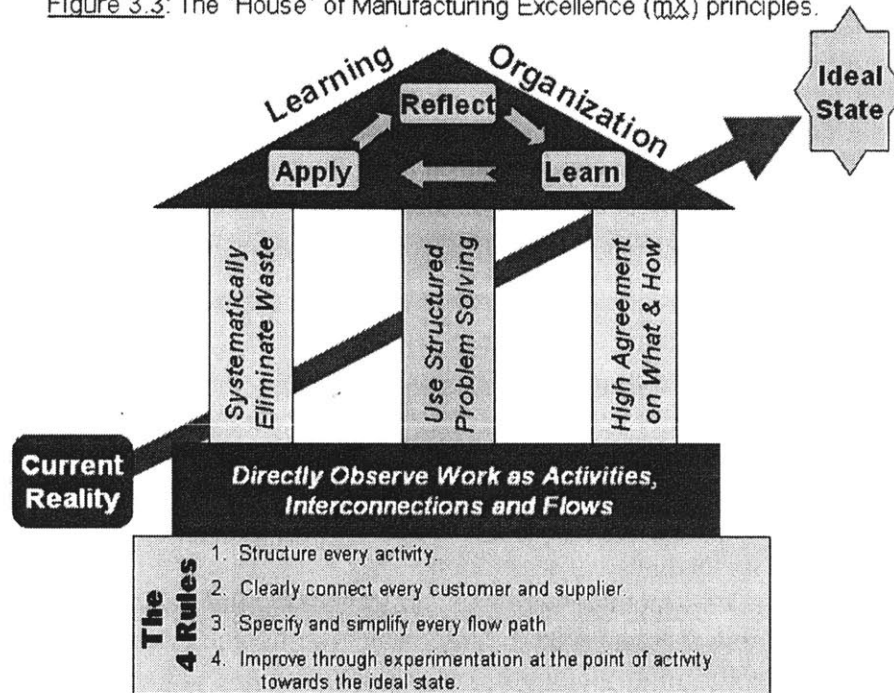
3.3.2 Pulling in Change on two fronts: "Mile-Wide, Inch-Deep" and "Inch-Wide, Mile-Deep"

In early 2004, the mX initiative was being directed by a core of four managers who devoted a significant portion of their time to coaching the rules and principles of lean. As a means to further proliferate this new way of thinking, the core team organized and trained a set of mX Champions to identify key mX opportunities and drive tactical improvements through the coaching, experimentation, and reflection of mX rules and principles in their home functional areas. By targeting troublesome, ongoing problems with mX thinking, Champions could start to solve thorny issues, credit their lean thinking with the success, and hence generate more interest and demand for further mX thinking among their constituents (Klein, October 2004). This approach encompasses the mX core team's "Mile-wide, Inch Deep" campaign.

In parallel, an "Inch-Wide, Mile Deep" initiative was pursued. In this strategy, a subset of operations within the entire flash memory manufacturing line can be selected as an isolated area for open experimentation and deeper learning about lean. In surveying possible candidates, the facility's Sorting operations were selected as Fab 23's first *mX Learning Lab* for several reasons:

- Sort 23 has been consistently rated the best Sort operation within Intel in terms of highest equipment utilization and lowest cost per wafer-sorted in recent years.

This history and confidence for success led to a high willingness to experiment with lean among the Sort 23 personnel.

- Flash Memory Sorting is a relatively simple process flow with only a few re-entry loops. Equipment availability is also very high, and the processing times across operations are consistent. These conditions lend to a favorable, stable environment in which experiments with WIP management and pull methodologies could be explored.
- The flexibility in micro-contamination regulations and clean-room standards presents significantly lower barriers towards implementing any layout changes, visual controls, or visual flow solutions that may be generated from lean thinking.

3.4 Thesis Focus

At the tactical level, the LFM internship was an opportunity to participate in the Sort Learning Lab to help identify and resolve ongoing problems through the use of mX rules and principles. The methodology for this research work was as follows:

1. The work in Sort was directly observed in order to gain an appreciation and sense for the type and quantity of waste in the area.
2. The mX rules of structured activities, customer-supplier connections, and simplified flows were applied to target latent inefficiencies and generate new opportunities for improving both cost and cycle time.
3. A tactical experiment targeting one of the key opportunities, WIP control, was then structured and implemented.
4. The results of the experiment were interpreted for further learning and refinement of the change.

Ideally, this scientific approach towards cost & cycle time improvements could be learned from and replicated to other areas in the Fab/Sort production line at Fab 23 for further mX proliferation.

Before directly embarking on the direct observation of the wastes in Sort, however, a brief overview of both the process complexity of flash memory sorting and the formal organization of the Sort 23 group is important.

4. FLASH MEMORY SORTING OPERATIONS

"Folks usually feel Sort is somewhat simple compared to the fab operations. The reality is very different...and the complexity is always a challenge for us to manage."
 - Sort 23 Operating Manager

In order to provide a good context for the research undertaken for the thesis, an overview of the processes, organization, and cost structure of Sort 23 is provided. Hopefully, the reader may also gain some appreciation for the daily complexity that confronts this Operating Manager!

4.1 Sort Processing

The purpose of a "sort" operation is to separate good and bad wafers and individual die thereon, taking wafers in from upstream fab operations and forwarding them to final manufacturing steps. Sort 23 receives its wafers[3] from two upstream fab operations: Fab 23 in Colorado Springs and Fab 11 in Rio Rancho, New Mexico. After processing the wafers, Sort then delivers the fully-sorted lots to a series of downstream Assembled Die Inventory (ADI) warehouses (which supply Intel's Assembly and Test operations with incoming inventory). Figure 4.1 shows how Sort 23's inputs and outputs fit within a greater context of Intel's flash memory supply chain.

Figure 4.1: Intel's networked flow for P803 & P804 process technologies. Sort 23 receives unsorted wafers from both Fab 23 and Fab 11 for testing & subsequent delivery to an Assembled Die Inventory (ADI).
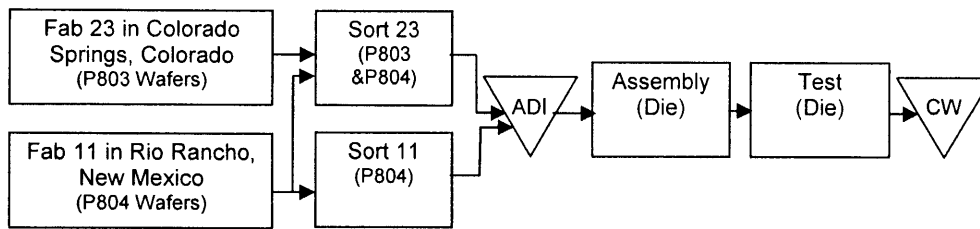
The wafers from both Fab 23 and Fab 11 arrive in lot sizes of 25. Depending on the wafers' upstream origin, Sort must then continue processing these lots through one of two standardized manufacturing processes, referred to as the P803 and P804 process technologies. In Sort, each of these process technologies is a fixed series of operations that enforce the following responsibilities:

1.) Test and remove wafers with out-of-spec electrical connection properties.
2.) Program memory functionality within each flash memory die.
3.) Stabilize the die's electrical properties and force a failure of any marginal die.
4.) Provide valuable end-of-line data for feedback to upstream operations in the fabs.

---

[3] A wafer is a silicon disc approximately 1 mm thick and 200 mm in diameter. Wafers are used to form the substrate of each flash memory die. Each wafer may produce a quantity of die devices ranging from several hundred to a few thousand, depending on the die product's type and size.

5.) Electronically 'mark' any non-functional die in preparation for the downstream
Assembly operations.

Notably, the P03 and P804 process technologies each have a different process flow to handle the unique characteristics of different flash memory designs.

4.1.1 Process Technology P803

Figure 4.1.1 shows the sorting process flow for the P803 technology products. First, all lots are taken through an E-Test process, where the interface structures created between the die during fabrication are tested for functionality. The data gathered at E-Test is compared to nominal process values; out-of-spec material is either immediately scrapped or placed on hold for further investigation depending on the reason for failure.

<u>Figure 4.1.1</u> Sort Process Flow for P803 Technology

| E-Test | → | Pre-Bake | → | Sort 1 | → | Bake | → | Sort 2 | → | Pack & Ship |

Fab 23
Delivery

Ship to
ADI
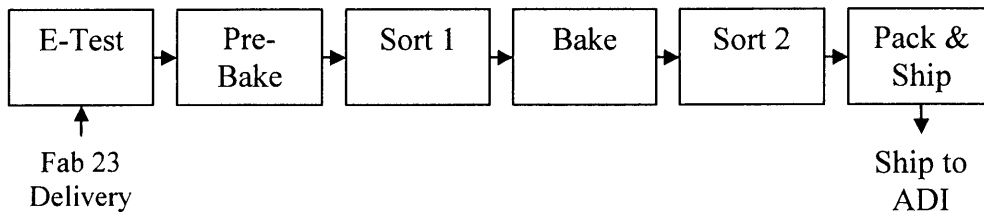
From E-test, the lots move to the first Pre-Bake operation. Bake operations stress marginal die to encourage failure as electrical charge is stabilized within each memory cell.

Thereafter, a lot is brought to the first sorting operation (Sort 1). Sort 23 operates a total of 63 sorting testers which have the flexibility to process any sorting operation (i.e. Sort 1 or Sort 2) on any process technology (P803 or P804). At each of these sort operations, the testing program, or Test Tape, checks that each individual die is functional. Die that fail are classified into bins based on the test which caused them to fail. The bin distribution of die on a wafer is another indicator which can cause material to be put on hold for further investigation by engineering. All sort test results are then loaded into a database system, and non-functional die are mapped in a computerized file that allows downstream assembly operations to discriminate between die which should be scrapped and die which should be processed further.

After the first Sort operation, an additional Bake operation is required to further stabilize the electrical charges. After the second Bake step, the wafers go through a second sorting operation (Sort 2) which not only checks die functionality but also programs each die with the necessary memory logic.

The final operation, Pack and Ship, prepares the wafers for shipment to the appropriate ADI site.

18

4.1.2 Process Technology P804

Figure 4.1.2 shows the sorting process flow for the P804 technology products. Cross-site wafers arrive at Fab 23 ready for processing at the Sort 1 operation. Two subsequent Bake-Sort processing loops are then required to fully stabilize, program, and test die functionality. Again, the same 63 sorting testers have the flexibility to handle any of the Sort 1,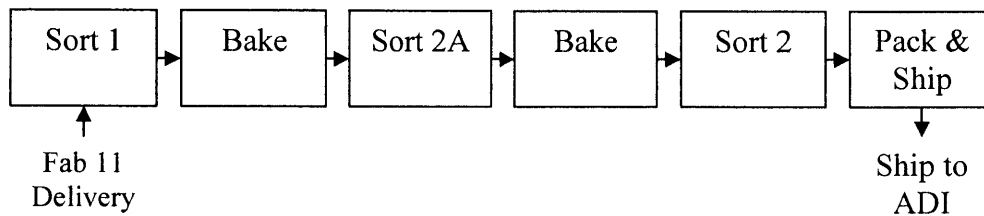 Sort 2A, or Sort 2 operations. The final operation, Pack and Ship, is identical with both the P803 and P804 technologies.

Figure 4.1.2 Sort Process Flow for P804 Technology

```
┌─────────┐   ┌─────────┐   ┌─────────┐   ┌─────────┐   ┌─────────┐   ┌─────────┐
│ Sort 1  │ ► │  Bake   │ ► │ Sort 2A │ ► │  Bake   │ ► │ Sort 2  │ ► │ Pack &  │
│         │   │         │   │         │   │         │   │         │   │  Ship   │
└─────────┘   └─────────┘   └─────────┘   └─────────┘   └─────────┘   └─────────┘
     ▲                                                                      ▼
  Fab 11                                                                Ship to
  Delivery                                                                ADI
```

4.1.3 The Product Mix of P803 and P804 Processing

Fab 23 is currently a single-technology fab, supplying Sort with a varying suite of between three to nine products within the P803 process technology. Although Fab 11 manages more than one process technology in their operations, they only supply Sort 23 with between two and three flash memory products from their P804 technology process.

Although the product mix can change over time, a larger majority of Sort's throughput volume is for the P803 flash products (approximately 82% of the wafers output during March and April of 2004 were P803). Note that although the throughput volume is smaller for the P804 products, the additional sort operation required for 804 technology leads to nearly 50% of the sorting tester capacity allocated to P804 on any given week.

In a way, Sort separates two very different manufacturing environments at Intel. As Sort's *supplier* of production material, both wafer fabrication facilities principally manage their product flow in lots which are segmented by process technology. Intel's Assembly and Test facilities have the most-immediate need for product from the Sorting operation; these *customers* require delivery of functional die segmented by product to feed their ADI levels. Sort operations then face a weekly challenge of connecting a downstream demand for product-specific die with an upstream supply of process technology-centric wafer lots.

4.1.4 Lot Handling & Changeovers

All wafer lots are loaded and moved between equipment sets manually. To facilitate manual handling, there are WIP racks nearby each of these processes in order to stage incoming lots.

The E-test and Sort tester operations require that equipment be setup differently for each product type. When a different product is run at these operations, several activities must be performed:

1. The correct software program, or 'Test Tape', must be loaded
2. The unique hardware interface, or probe-card, must be available and installed onto the equipment
3. The tester must be brought to the correct operating temperature.

The product-specific test programs can be substantially different due to programming logic and number of die per wafer, resulting in differing processing times across products. Typically, product changeovers require between 7 and 13 minutes of tester downtime to complete depending on the range and direction of temperature change (heating the tester cell is faster than cooling). Changing a tester to operate on a different process technology is essentially the same as a product-to-product changeover in this regard.

In the case of purely flexing a Sort tester over to a different sorting operation (i.e. from Sort 1 to Sort 2 with no change in product-type), a probe-card change is not required. However, the required temperature change between sorting operations is typically more significant, raising the required tester downtime to between 5 and 45 minutes depending on the direction of the necessary temperature change.

4.2 Organizational Structure

The Sort 23 organization is a relatively small department, consisting of a Department Manager, two Operating Managers who directly manage 4 shifts of Manufacturing Technicians, and two teams of Process and Equipment Engineers (see Figure 4.2). Sort is also supported by several Automation & Industrial Engineering resources who normally work during the daytime-shifts, Monday through Friday.

Figure 4.2: Sort 23 Organization
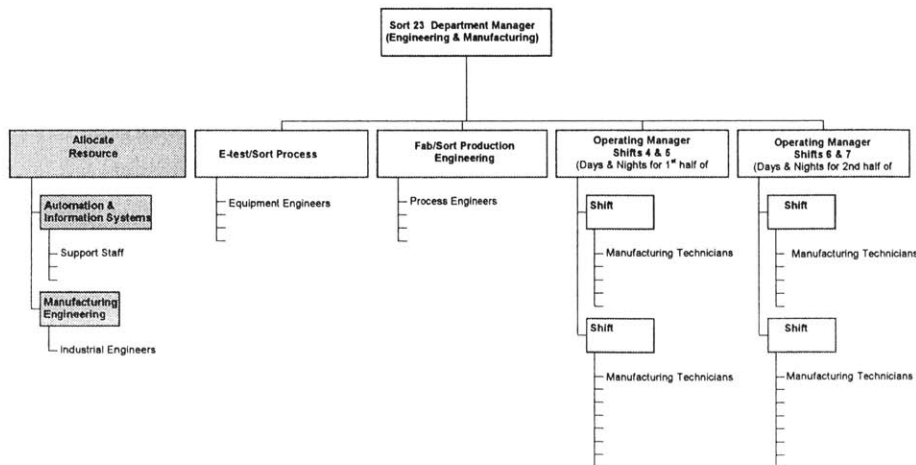
The Operating Managers (OMs) split production responsibilities across the first half (front-half) and second half (back-half) portions of each week. Each OM formally covers both the day and nighttime shifts but tend to split the majority of their 12 hour work-days with the day-shift. About eight Manufacturing Technicians (MTs) are assigned per day shift and five per night shift. All the MTs share responsibility for operating and maintaining the production equipment on the floor as well as the manual moving and loading of lots. Within these teams, one or two experienced MTs are designated as Area Coordinators (ACs) with additional responsibilities of coordinating daily operating activities both within a shift as well as across other shifts.

4.3 Sort Economics & Metrics

Sort's major sources of cost are three-fold:

1.) Payroll
2.) Probe-card inventory purchases
3.) Depreciation of capital from both equipment and building investments

Other operating costs, such as energy consumption, clean room support, and repair tools are not significant expenditures relative to these categories. Payroll and probe card inventory make up approximately $1/3^{rd}$ of the total cost of running Sort 23, whereas non-cash depreciation of equipment and building make-up approximately $2/3^{rds}$.

Intel benchmarks their internal operations on a dollar per wafer cost basis. For Sort 23, this means their aggregate costs of payroll, probe card purchases, and depreciation are divided by the volume of good wafers sorted each month to generate a summary performance measure. This cost per sorted-wafer metric is then used to rank Sort 23 among other sort operations within Intel. Notably, measures of WIP inventory and cycle time performance are not key benchmarking metrics within the group.

The heavy emphasis on the floor's cost-per-wafer drives tester utilization as a critical success indicator for Sort 23, as maximizing this bottleneck resource is a sound approach to maximize output volume. Additionally, Intel coordinates its production volumes across a network of distributed fab/sort manufacturing capacity. Across this Virtual Factory, or VF, Sort tester capacity is also designed as a bottleneck constraint; Sort 23 is encouraged to maintain high tester utilizations for this reason as well.

Working with the fab operations, Sort 23 must sort wafers according to a production schedule. Each week, the Fab/Sort collaboration strives to deliver all the planned product volumes (and hence scoring a 100% LIPAS metric[4]).

---

[4] Line Item Performance As Scheduled (LIPAS) is a binary (two-state) performance metric used in reference to product schedules which are either being met on time or not.

4.4 Key Takeaways

Clearly, Sort 23 is responsible for a complex workflow, as managing the delivery of multiple products across multiple re-entrant processes through a flexible pool of sorting testers is a puzzling challenge. In contrast to this operational complexity, the group has a very simple, unifying goal: maximize tester utilizations subject to meeting all the weekly product schedules. As we will see, this focused coordination has enabled Sort 23 to out-perform its Intel peers in terms of the cost per wafer-sorted metric - making the prospect for further waste identification somewhat daunting.

However, the principle of directly observing the work towards waste discovery can be a powerful method to see outside existing paradigms of success and raise the awareness of latent inefficiencies. Thus, our next chapter will be an attempt to assess and quantify an objective reality of the Sort 23 operations.

5. DIRECT OBSERVATION OF THE CURRENT REALITY

"There are many examples of waste in the workplace, but not all waste is obvious. It often appears in the guise of useful work. We must see beneath the surface and grasp the essence."
- Shigeo Shingo (Shingo and Robison, 1991, p. 83)

Since everyday data and metrics are abstractions from a complex reality, a central principle of mX is to continually re-examine the work in question in order to discover and understand the latent wastes within the system. mX advocates the use of two discovery tools which facilitate this kind of deep and direct observation:

- Activity Mapping
- Product/Process Mapping

These tools are helpful in channeling direct observations into a greater context of "what usually happens" in a given process flow and then looking for opportunities to eliminate waste. However, the tools fail to include an understanding of the significance and sources of *process variability*. Hence, additional statistical data covering various characteristics of the Sort operation is also presented here, including:

- Variability in Cycle Times
- Tester Utilizations
- Rework Rates
- Product Delivery Rates from Fab 23
- Production Scheduling Dynamics

Finally, descriptions of the current methods for Tester Capacity and WIP Management are provided.

5.1 Activity Mapping

The Activity Map defines who is involved in the manufacturing process and how long each of their activities takes. Through direct time study, a typical map shows the time each employee must contribute to a process item or product as the item progresses through a series of standardized steps. For example, mapping a typical lunch-order process in a restaurant might show 1.) a 15 minute block for the customers to decide on their selections, then 2.) a 5 minute block for the waiter to receive their order and take it to the kitchen 3.) a 10 minute block for the cook to receive the order and prepare the food, and finally 4.) a 5 minute block for the waiter to deliver the food back to the customer.

Applying the Activity Map methodology to Sort reveals a much more nimble work environment for the Manufacturing Technicians; the number and variety of the daily MT activities in Sort vary constantly depending on equipment-related problems, number of tester loads or changeovers required that day, and number of lots awaiting E-test or shipping. A snapshot of three Technicians shows about 80% of their total shift was

consumed by direct manufacturing activities such as moving lots, firefighting problems, and equipment maintenance. The remaining 20% of time was dedicated towards administrative tasks, lunches, and breaks. Notably, the timing and coordination of these activities were mostly unstructured and keyed by equipment malfunctions and lot-completion events. However, Manufacturing Technicians are allowed a high-degree of autonomy in scheduling their daily tasks around these events, and most Technicians take pride in managing their time productively.

5.2 Product/Process Mapping

The Product/Process map depicts what typically happens to a product during the production process. With different processing times and process steps across the two different technologies, two different maps are generated to capture the current reality. Figure 5.2 shows the mappings for both P803 and P804 process technologies based on median operating times for 20 weeks in early 2004.

Figure 5.2: Product/Process Maps

Median Cycle Times
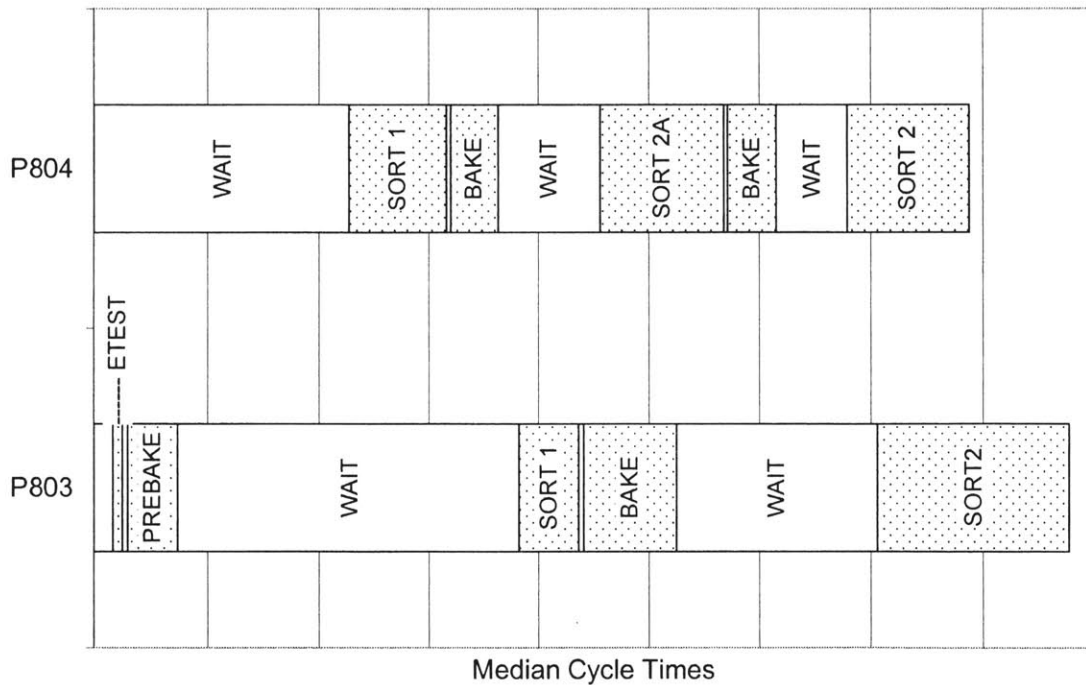
The maps show the most significant wait times occur before the Sort operations, reaffirming tester capacity as the floor's bottleneck constraint. Also, approximately half of the total cycle time is from waiting for an available sort tester (56% for P803 and 49% for P804).

5.3 Variability in Cycle Times

A statistical analysis of cycle times was conducted for all the lots processed at Sort 23 for 20 weeks. Figures 5.3.1 and 5.3.2 show the 1st, 2nd, and 3rd quartile for this sample data for each process technology.

Figure 5.3.1: Cycle Time Variability for P803

□ Median   • 1st Quartile   • 3rd Quartile
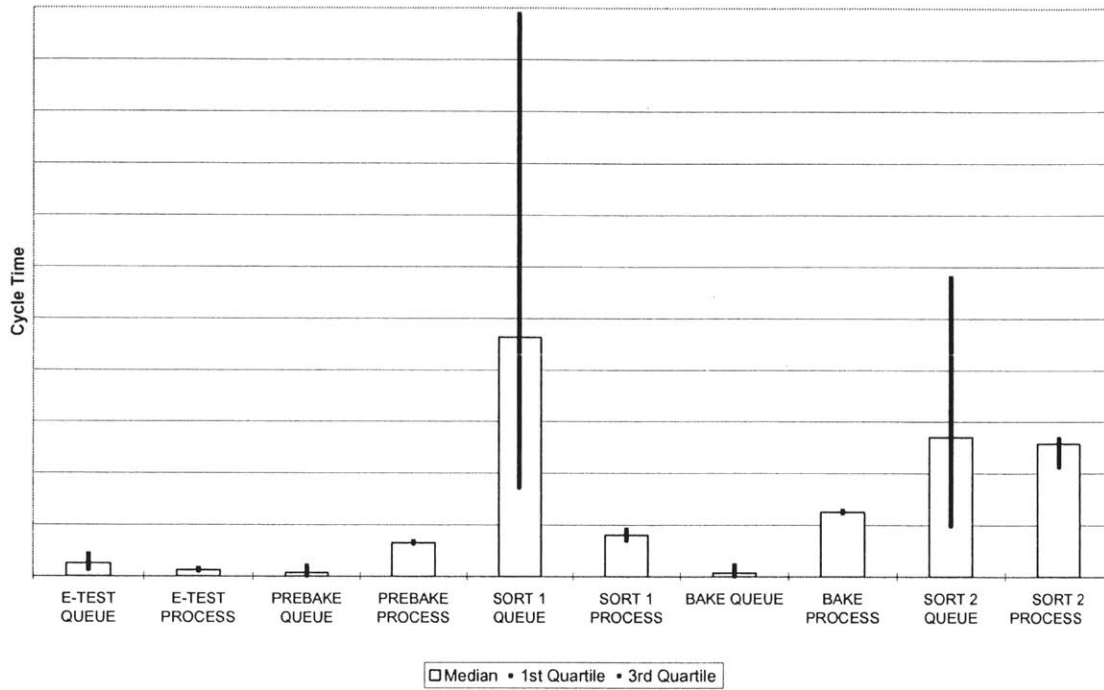
Figure 5.3.2: Cycle Time Variability for P804

□ Median   • 1st Quartile   • 3rd Quartile
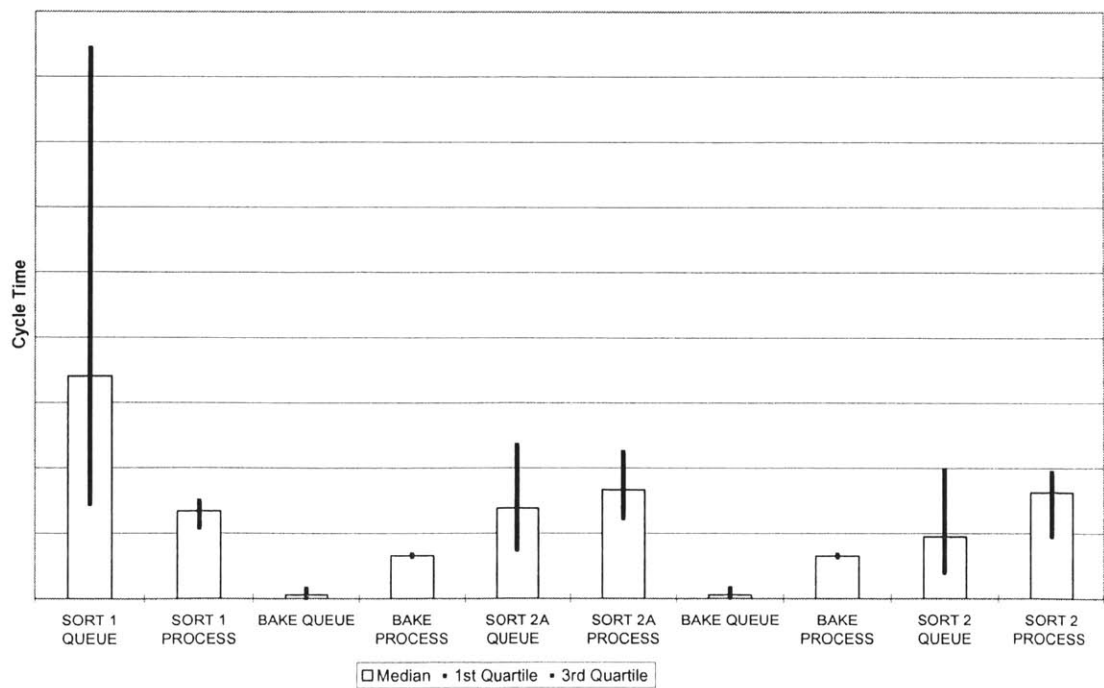
Significantly, the largest source of variability appears in the queue times before the sort operations. The variability in wait time before Sort 1 appears to be especially significant, suggesting further variability in the arrival of product from the upstream fab operations. Note the distribution of Sort processing time can be principally attributed to focused differences in processing time per product family, as shown in Figures 5.3.3 through 5.3.5.

Figure 5.3.3: Variability in Sort 1 Process Time
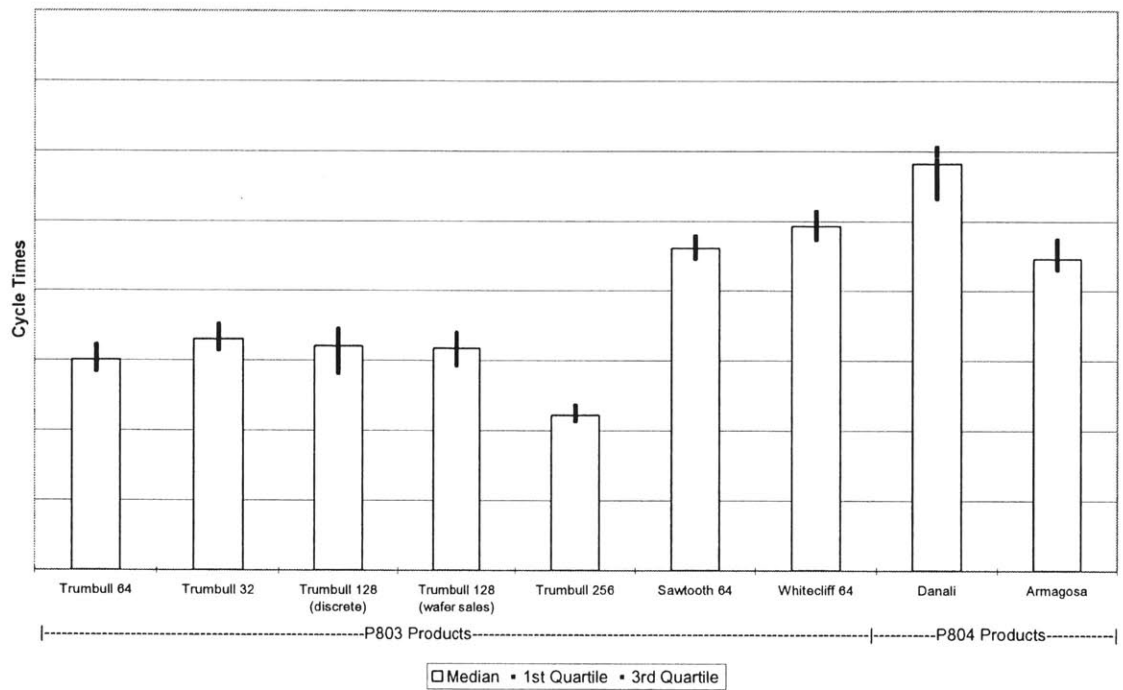
Figure 5.3.4: Variability in Sort 2A Process Times

Cycle Times

Danali

Armagosa

P804 Products

☐ Median   1st Quartile   3rd Quartile
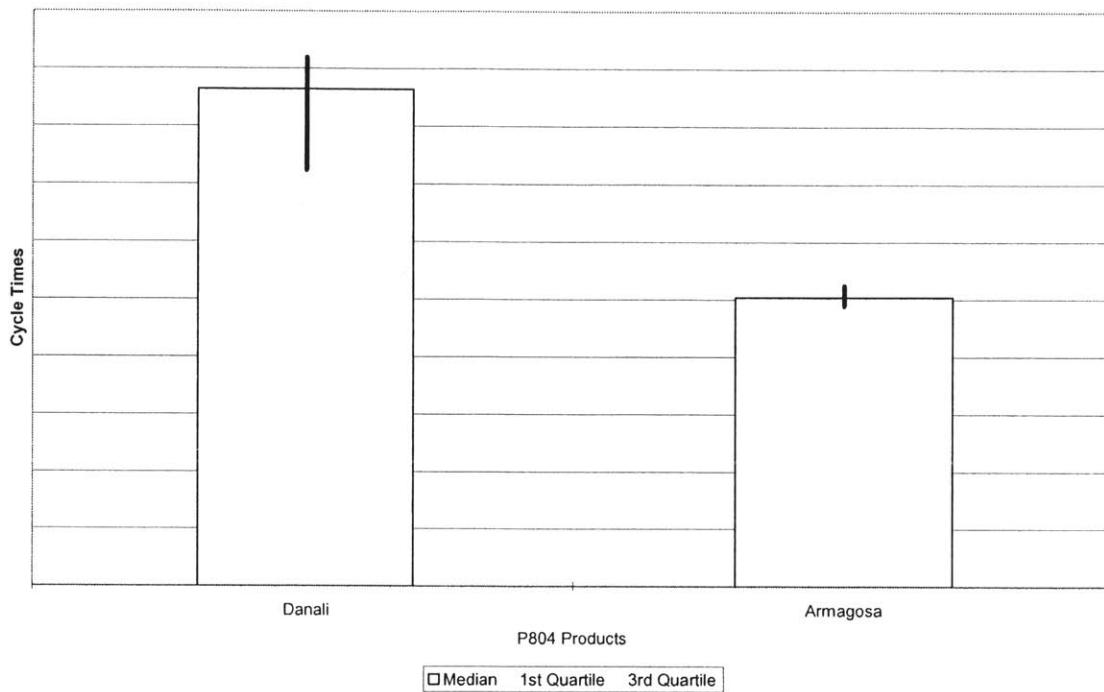
Figure 5.3.5: Variability in Sort 2 Process Times

Cycle Times

Trumbull 64   Trumbull 32   Trumbull 128 (discrete)   Trumbull 128 (wafer sales)   Trumbull 256   Sawtooth 64   Whitecliff 64   Danali   Armagosa

|------------------------------------P803 Products--|----------P804 Products-----------|

☐ Median • 1st Quartile • 3rd Quartile
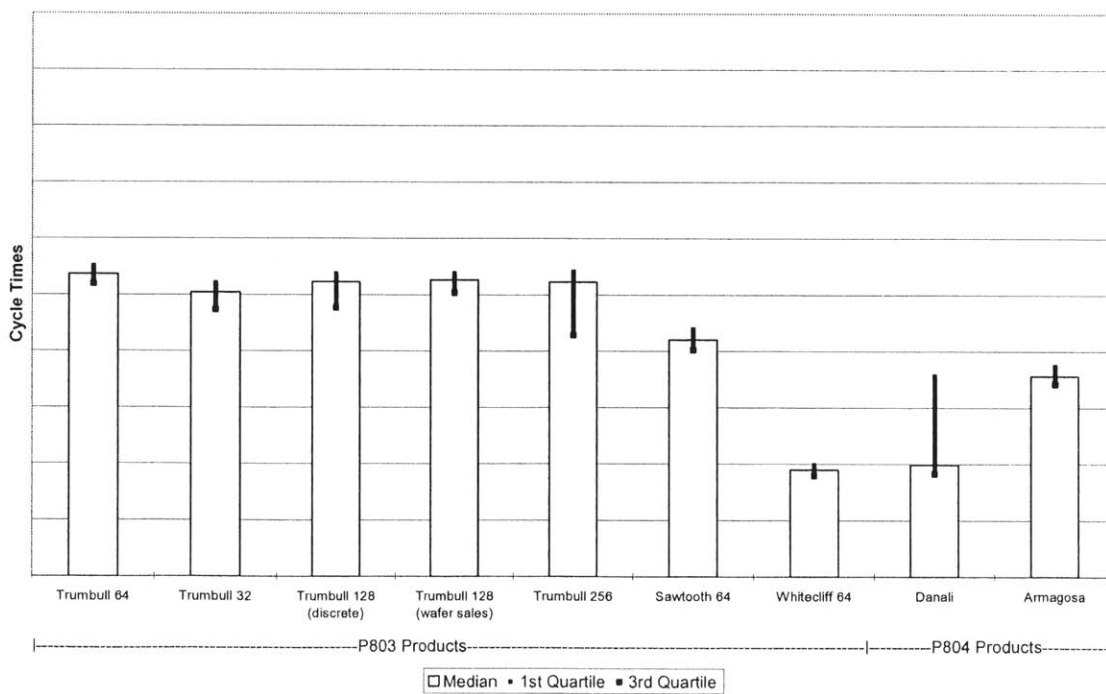
5.4 Tester Utilizations

Sort 23's utilization measure is defined by the amount of time the floor's sorting tester equipment is processing a lot (excluding setup times) divided by the total time the equipment is qualified[5] to run the product, or arithmetically:

$$Utilization = \frac{Time\,Sorting\,Wafers}{Total\,Time}$$

For each of Sort's tester models (VHP and AGT), the average weekly utilization metric is calculated as:

$$Weekly\,Utilization = \frac{\sum_{i=1}^{Number\,of\,testers} \sum_{j=1}^{\#Wafers\,processed} Time\,Sorting\,each\,Wafer}{Number\,of\,Testers \times 168\,hours\,/\,week}$$

Examining weekly utilization levels (indexed against internal benchmarks, where 1.0 equals target performance) shows consistently high metrics for its fleet of 60 VHP testers but sporadic dips for its three AGT testers (Figure 5.4.1). Clearly, this metric is sensitive to the total number of testers within a model category. For example, in WW26 one of the AGT testers failed a maintenance diagnostic and experienced considerable down time from delayed troubleshooting and re-qualification. This single incident caused the AGT utilization to plummet to around 0.75, yet an equivalent downtime on a VHP model tester would only register as a 0.02 drop in that model's weekly utilization (indexed).

Figure 5.4.1 Weekly Utilizations for VHP and AGT Tester Models
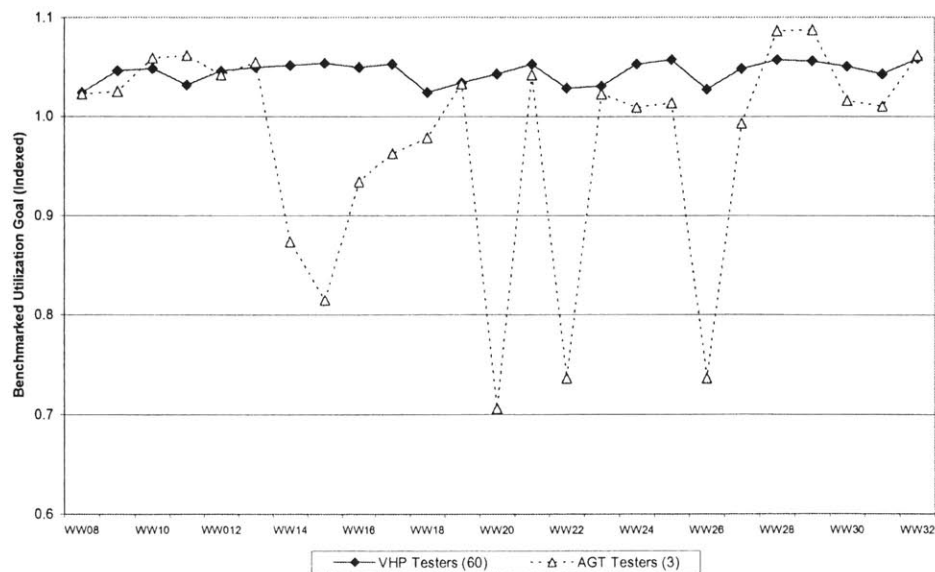
[5] Qualification is an engineering process that certifies individual testers as capable of performing particular processes on particular products. Only equipment which has been qualified may be used to produce product for shipment to customers and hence is included as part of utilization.

One of the ways Sort maximizes its utilization metric is by minimizing the lost production time from tester changeover activity. For example, assume a sort tester is setup to run product A with product B already waiting to be tested. If a product A arrives and there is no imminent need for more of product B, the new product A will be loaded onto the equipment ahead of the waiting product B to save a setup changeover from occurring. Similarly, a sort tester processing Sort 1 will continue to operate on Sort 1 until a pressing need requires a change to Sort 2A or Sort 2.

As mentioned before, changeovers can vary from 5 to 45 minutes depending on the range and direction of the operating temperature change. With only a handful of changeovers occurring each shift, the lost tester time typically amounts to between 0.1 % and 0.4 % of the total tester availability (see Figure 5.4.2 for complete breakdown of tester availability).

Figure 5.4.2: Breakdown of Logbook Data by Tester Status

RUNNING PRODUCT	94.35%
DOWN	2.80%
IN REPAIR	1.38%
MAINTENANCE	0.45%
WAITING PRODUCT	0.31%
ENGINEERING	0.26%
CHANGEOVER	0.21%
WAITING PART	0.13%
WAITING TECHNICIAN	0.06%
MAINTENANCE OVERDUE	0.04%
Total Tester Time	**100%**

(Data taken from all VHP Testers for 1st 14 weeks of 2004).

5.5 Rework Rates

Rework at Sort 23 is defined by the redundant processing time spent re-sorting a wafer. Figure 5.5 shows the percentage of rework time over total processing time for a period of nearly 5 months. Intel plans for 0.3% rework rates in each of their Sorting operations, which approximately matches the mean values of the observed data.

Figure 5.5 Sort Rework Rates
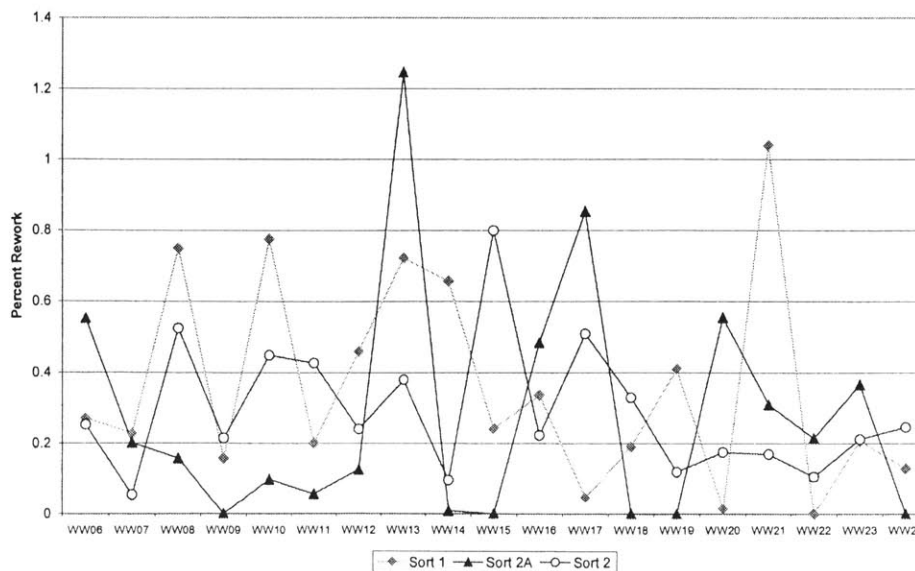
5.6 Fab 23 Planning and Product Delivery to Sort 23

Each week, Fab 23 starts manufacturing a certain number of wafers per product with a planned lead time of 8 weeks before delivery to Sort. Sort then has a two-week lead time in which to process these wafers and deliver the scheduled die quantities to ADI. Hence, the fab wafer starts are always based on what ADI will need 10 weeks in the future; die-schedules oftentimes change during this lag period. Indeed, schedules may change even before a few days of being due. So as the last set of operations before ADI delivery (and LIPAS scoring), Sort has a de-facto responsibility to monitor and meet these changing die schedules on a weekly basis.

Figures 5.6.1 and 5.6.2 show the delivery of Fab 23's wafers into Sort 23 during each shift as well as aggregated over a weekly horizon for a period of 17 weeks. Wafer arrivals appear to be highly-variable from shift-to-shift yet roughly consistent on a weekly tally. Notably, fab operations emphasize the execution of wafer delivery, or 'fab outs', on a weekly basis.

Figure 5.6.1 Fab 23 Delivery Rates to Sort 23
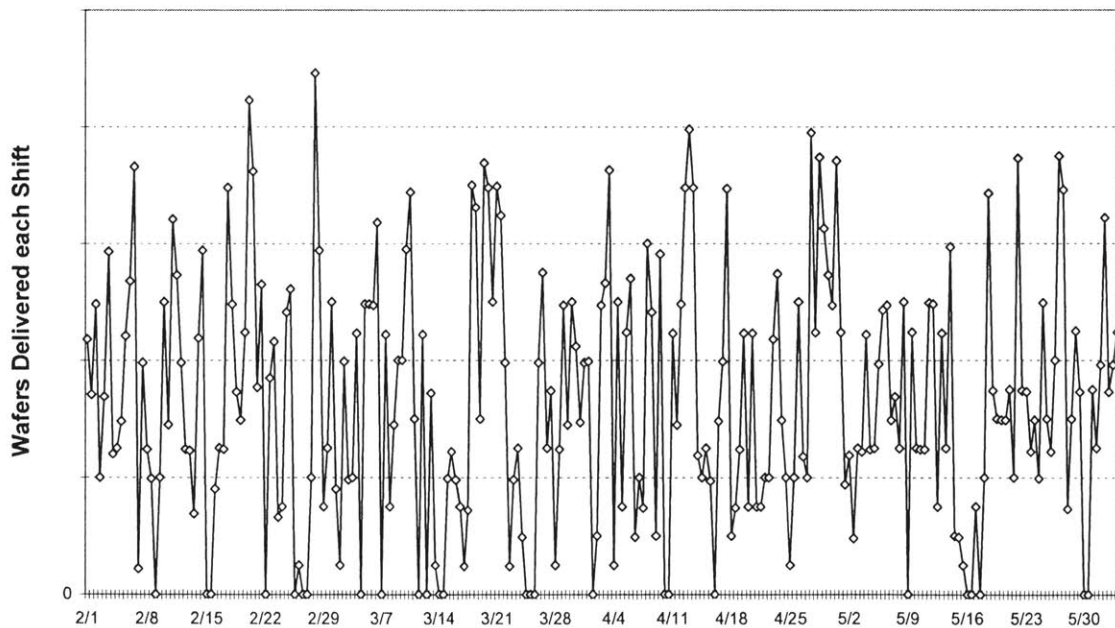
Figure 5.6.2 Fab 23 Delivery Rates to Sort 23
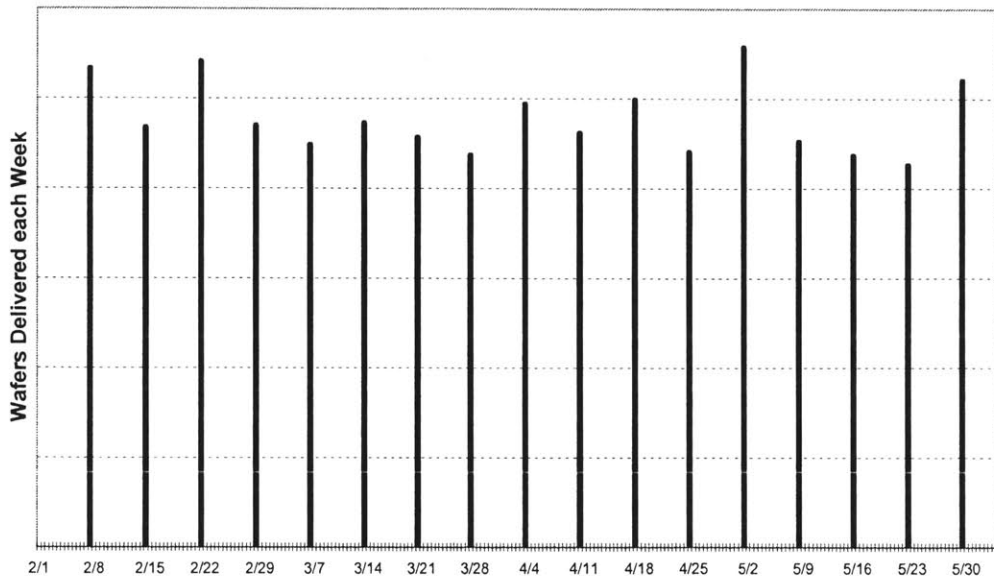
In order to satisfy its weekly die-out targets, Sort 23 must allocate their tester capacity by product type and sorting operation. Figure 5.6.3 shows how the delivery of wafers from Fab 23 corresponds to the product-specific loadings necessary for Sort 23 to meet their schedules within their two week lead time. As a customer, Sort sees large swings in the delivery of products from their fab supplier – a dynamic one technician describes as 'feast or famine'.

Figure 5.6.3 Fab Delivery Performance by Product*
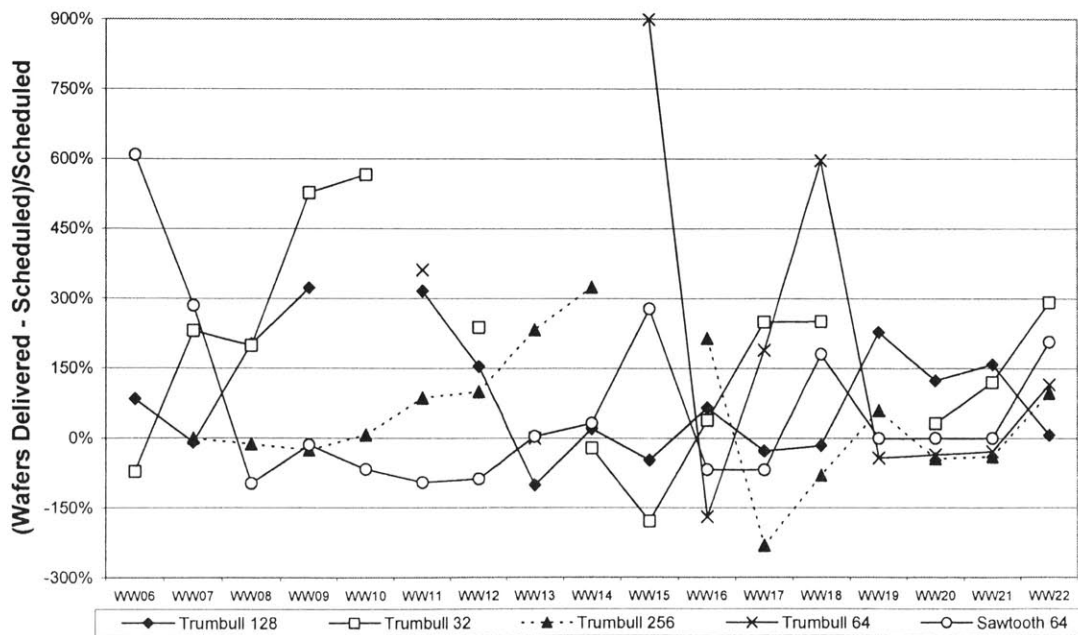
* Since only a subset of products are scheduled on any given week, many product volumes are intermittently absent.

While both Fab and Sort Operations Managers reference common information reports detailing daily product delivery status from the fab, the communication pathways between the groups are numerous (emails, phone calls, and meetings), complex (involving dozens of Technicians, several Operating Managers, Planning personnel, and Industrial Engineering liaisons), and crippled by delays; at the point in time when a Sort Technician communicates a need for more Trumbull 32, the Fab managers might easily require 3 to 7 weeks to adjust the delivery rate for this product. This gap in customer-supplier connection leads not only to variability in product inventory levels on the Sort floor but also frustration on the part of Sort personnel to maintain predictable flow to fulfill die schedules.

For the time interval examined, there appears to be more product volume delivered weekly from the fab than is actually needed by Sort's downstream customer, ADI. As an extreme example, the fab delivered approximately ten times the amount of Trumbull 64 in WW15 than was necessary to fulfill schedules (only to be balanced by a deficit of 2.5x the next week)! As we will see next, these excessive inflows may cause surges in Sort inventory levels in the near term yet support a steady practice of overproducing beyond scheduled amounts over the longer term.

Since the size of ADI inventories remains manageable at Intel despite this overproduction, one might argue that the final die schedules are systematically underestimating the real demand of the downstream Assembly operations on ADI. Given the time interval and scope of this thesis research, however, this conclusion is difficult to confirm.

5.7 Production Scheduling Dynamics

Twice a week, each Sort Operations Manager and daytime Area Coordinator meets with the department's Industrial Engineer and a representative from the Fab 23's planning group to discuss anticipated production volumes and product mix over a 2-7 week horizon. Schedule changes from ad-hoc customer requests, changes in fab yields, run times, and equipment availability are all coordinated across these groups during these meetings. Information on the latest schedule is then communicated to the floor by both the OM and AC through email, shift-to-shift pass-down meetings, and informal discussions.

As an example of these changes, Figure 5.7.1 shows the production schedules for the Trumbull product categories over a period of 18 weeks. The planned product volumes (in wafers, indexed) are broken down into three sequential snapshots in time:

  1.) The schedules as they appear two weeks before LIPAS is due
  2.) The same schedules once 'fixed' a week before LIPAS is due, and finally
  3.) The final volumes delivered to ADI

Figure 5.7.1 Schedule Changes and Final Production Volumes

Trumbull 128 Discrete

Legend: —O— Sort Output   —■— Schedule @ 1 week to LIPAS   —▲— Schedule @ 2 weeks to LIPAS

Trumbull 32

Legend: —O— Sort Output   —■— Schedule @ 1 week to LIPAS   —▲— Schedule @ 2 weeks to LIPAS
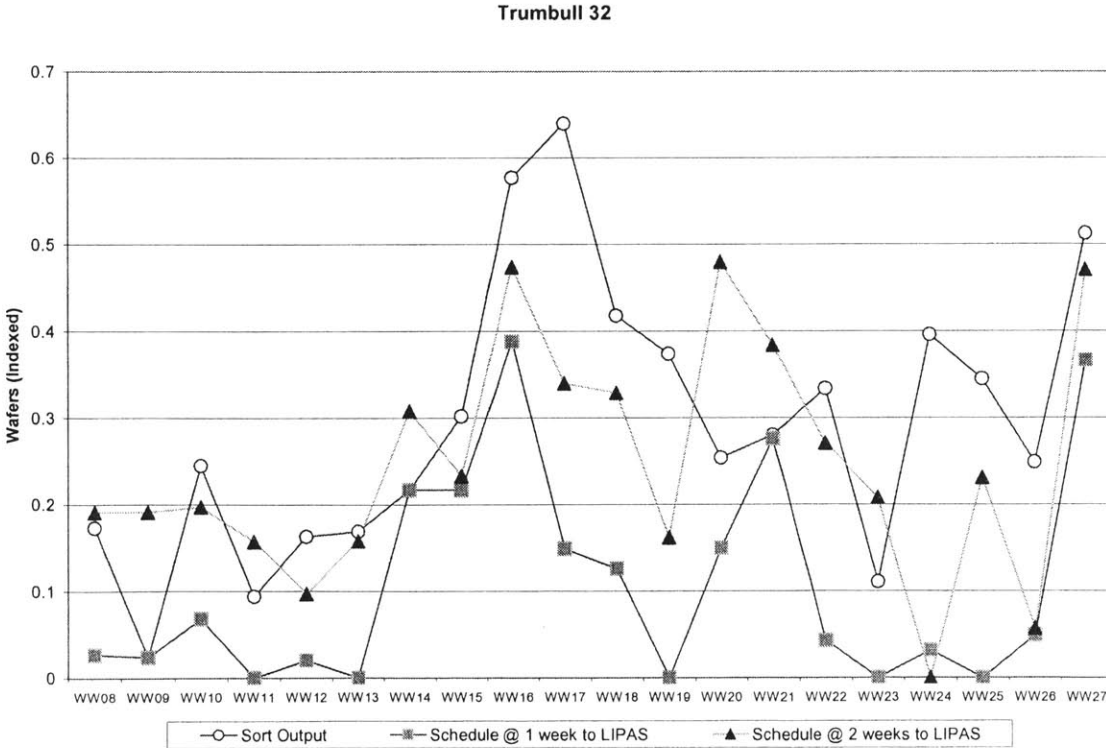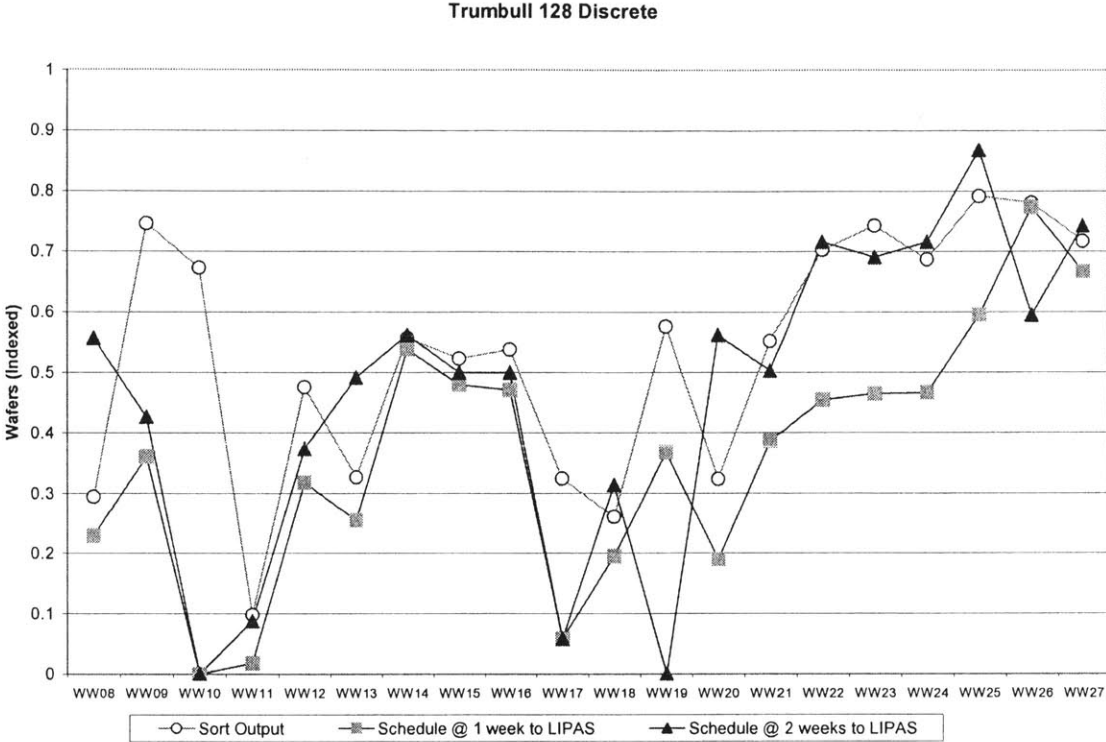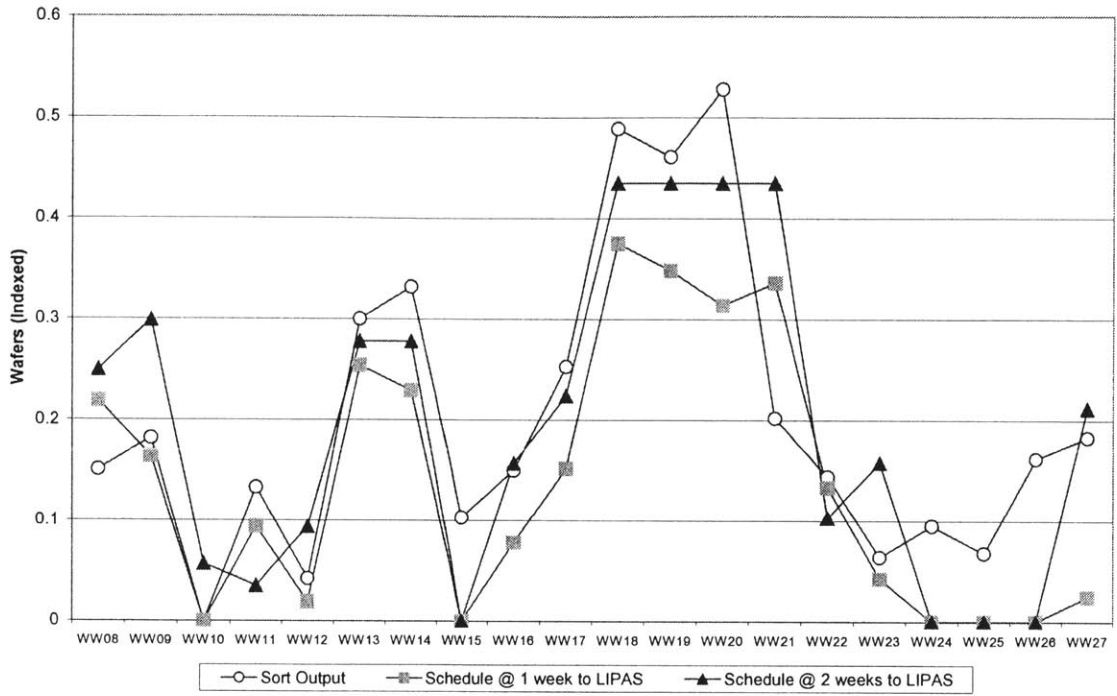
Trumbull 256

Trumbull 64
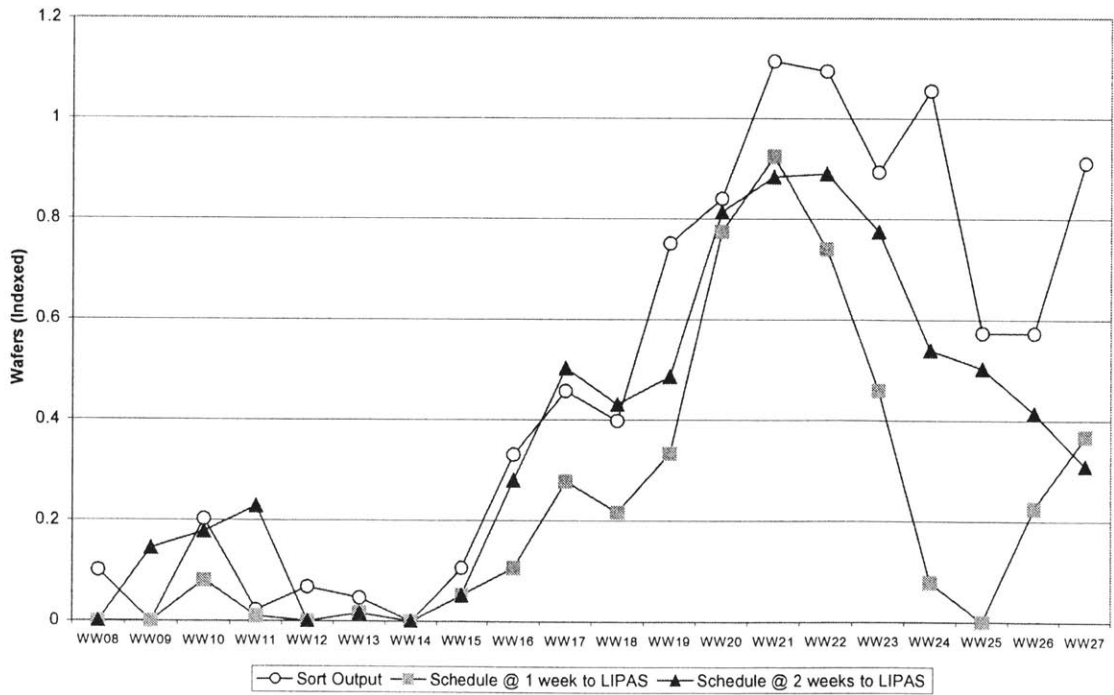
Two interesting trends emerge from this viewpoint. First, as delivery dates become more immediate (i.e. the horizon closes from 2 weeks to 1 week), the scheduled volumes seem to drop across many of the product categories and weeks. However, comparing these final 'fixed' schedules to the volumes that Sort actually delivers to ADI shows a systematic trend in Sort consistently delivering wafers beyond the final schedule amounts.

One key reason for this dynamic stems from a policy of accruing all *overproduction* as credit in future week's schedules. For example, if Sort ships 100 die of Trumbull 128 beyond their weekly target, then planning will lower their production schedule by 100 units the subsequent week. On the one hand, the practice of shipping ahead of schedule allows Sort to smooth large peaks in scheduled demand through early accruals, mitigating large fluxes in tester changeovers. On the other hand, overproduction is a clear signal that schedules are lax relative to the output performance of the floor. Notably, Sort only missed a final die schedule three times over the course of 18 weeks studied; if Sort were held to the die schedules set two-weeks away from the delivery due dates, they would have missed their schedules 56 times!

Since overproduction is a normal practice, a tier catalogue is used to coordinate the product priorities from delivery to ADI across all of Intel's Fab/Sort operations. Figure 5.7.2 shows how tier assignments optimize output levels across product categories subject to scheduling volumes.

Figure 5.7.2 Tier Priorities for Product Delivery to ADI

Priority	Description
Tier 1	Maximize output. Subordinate other schedules to overship if possible.
Tier 2	Hit schedules. Once schedule is hit, do not subordinate Tier 3 schedules to overship schedules.
Tier 3	Maximize performance to schedule. But not at expense of Tier 1 & 2 expectations.
Tier 4	Prioritize behind Tier 1-3. Only run if you have idle testers or you are on track to make all Tier 3 schedules.

Hence, an over-shipment of Tier 3 product while missing a Tier 2 schedule would constitute a misallocation of tester capacity for Sort 23. It is difficult to measure how strictly Sort performs to these guidelines given the complex variables that determine each product's overproduction; the tier catalog prioritization must be considered alongside any future peaks in schedule volume (spurring early accruals and smoother production volumes across multiple weeks) as well as limits from the delivery of product from the fab (constraining the amount of accrual possible).

As a rough gauge of performance, Figure 5.7.3 shows the tier priorities and Sort's over-shipments during weeks 15 and 16 across all products. From this limited perspective, Sort 23 appears to be lax in adhering to the tier priorities rules; the excess delivery of Armagosa 32 above schedules is especially suspicious when ranked against the higher-priority, lower-volume delivery of Trumbull 32.

Figure 5.7.3 Overproduction across Tier Priorities in Work Week 15 and 16

Work Week 15

Work Week 16
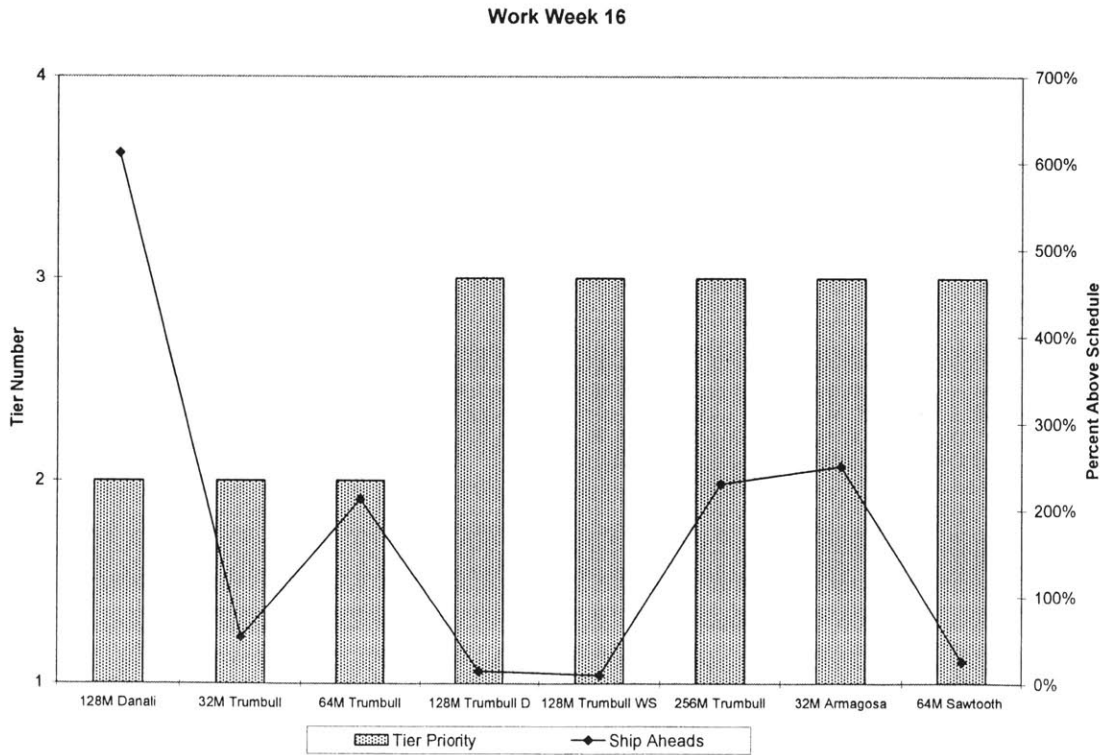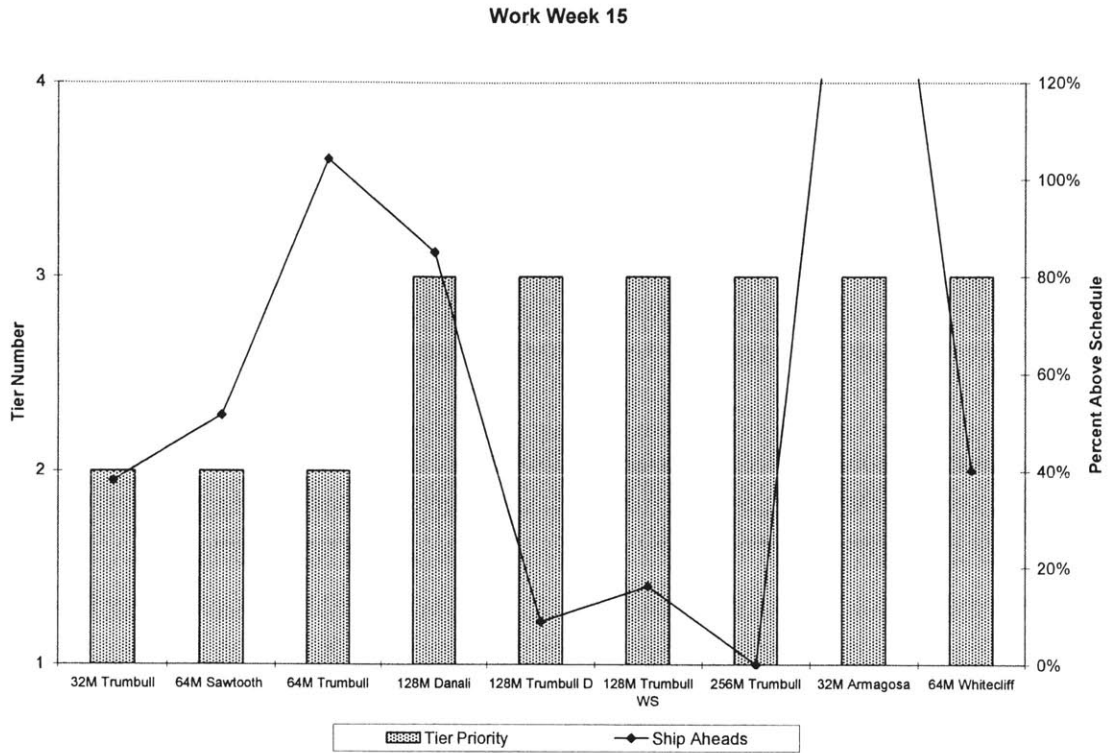
5.8 Capacity Management Systems

The Sort Area Coordinators have developed spreadsheets that estimate the minimum number of testers that must be setup on each product and each sort operation to fulfill the current mix of die schedules. Given a series of sorting operations i and products j, the tester estimates for any given week can be calculated as follows:

Let

Q_j = the remaining quantity of scheduled die for product j

R_{ij} = process time required to sort each wafer of product j at sorting operation i

D_j = the fixed die-per-wafer ratio

Y_j = the historical die yield for product j

T = the amount of time left before the delivery due date

The whole number tester estimate S_{ij} for each sorting operation i and product j is

$$S_{ij} = ROUNDUP\left(\frac{Q_j R_{ij}}{TD_j Y_j}\right)$$

where the total sum of available testers must not be exceeded, or

$$\sum_{i=1}^{Sorting\ Operations} \sum_{j=1}^{Total\ \#\ of\ Products} S_{ij} \le \left(total\ sort\ testers\right)$$

Notably, fractional tester quantities are always rounded up to promote over (versus under) production. The capacity estimates also assume 100% tester utilizations, 0% resorting rates, and deterministic run rates and die yields. Given these uncertainties, the model is iteratively updated with quantities Q_j throughout the week as production due dates advance and time horizon T decreases.

For example, let's assume product A has 200 die per wafer, product B has 750 die per wafer, and that each product has an historical yield of 85% and 95%, respectively. Furthermore, let's assume we know that the run times for the Sort 1 and Sort 2 operations are 1 and 2 hours per wafer for product A and 1.5 and 3.5 hours per wafer for product B. Finally, at a time of 168 hours before schedules are due we have 625,000 die of product A and 435000 die of product B yet to deliver:

Product (j)	Die per Wafer (D_j)	Historical Die Yield (Y_j)	Run time per wafer for Sort 1 $(R_{i=1})$	Run time per wafer for Sort 2 $(R_{i=2})$	Remaining Die Due (Q_j)
A	200	85%	0.75 hours	1 hour	625,000
B	750	95%	2 hours	3.5 hours	435,000

Sort would then allocate testers to produce products A and B as follows:

$$S_{1A} = ROUNDUP\left(\frac{625,000\,good\,die \times 0.75\,hours\,/\,wafer}{168\,hours \times 200\,die\,/\,wafer \times 0.85\,good\,die\,/\,die}\right) = 17\,testers$$

$$S_{2A} = ROUNDUP\left(\frac{625,000\,good\,die \times 1\,hours\,/\,wafer}{168\,hours \times 200\,die\,/\,wafer \times 0.85\,good\,die\,/\,die}\right) = 22\,testers$$

$$S_{1B} = ROUNDUP\left(\frac{435,000\,good\,die \times 2\,hours\,/\,wafer}{168\,hours \times 750\,die\,/\,wafer \times 0.85\,good\,die\,/\,die}\right) = 8\,testers$$

$$S_{2B} = ROUNDUP\left(\frac{435,000\,good\,die \times 3.5\,hours\,/\,wafer}{168\,hours \times 750\,die\,/\,wafer \times 0.85\,good\,die\,/\,die}\right) = 13\,testers$$

Here, the total number of fully-allocated testers required to fulfill the schedules for product A and B is 60 – three testers below Sort's capacity constraint of 63. Similarly, when scheduled demand is below tester capacity, some quantity of testers in Sort will not be accounted for in this model. In this case, the Area Coordinators then work with the Operating Manager to allocate these spare testers based on a judgment of any low WIP levels in front of the Sort 1, Sort 2, and Sort 2A operations or a preference to strategically over-produce a given product.

Notably, the time horizon used for these spare allocation decisions is inconsistent across the shifts. As a rule, the front-end of the week does not look beyond the current week's schedule in making allocation judgments. By comparison, the shifts on the back-end of each week take both the current *and the following week's schedules* into account when allocating spare testers. For example, the Trumbull 32 volume might jump by 30% in the subsequent week's schedule. In this case, the back-end of the week would allocate spare tester capacity on Sort1 and Sort2 operations for Trumbull 32 in order to over-ship during the current week and thus lower next week's spike in volume. The front-end of the week would most likely ignore this upcoming spike and allocate the spare tester capacity on a Tier 1 or Tier 2 product.

5.9 WIP Management Policies

Within a product category, Sort follows FIFO processing at all operations. Expediting lots outside this discipline is rare and makes up a negligible percent of total output volume. However, during extreme schedule pressure for a specific product, the floor will split lots into smaller batches in order to accelerate the sort processing times. This practice is strongly discouraged, as it leads to more ad-hoc work for the Manufacturing Technicians and an increased risk of mis-processing a lot.

In order to manage WIP across the process flows, the Sort Area Coordinators use a customized spreadsheet which summarizes the number of lots and wafers at each process operation by individual product family. The tool promotes a reactionary approach to WIP

management, as noticeable dips and peaks in current WIP levels are continually 'smoothed' through the aforementioned re-allocation of spare tester capacity.

Indeed, the cycle time variability at the Sort 1, Sort 2A, and Sort 2 operations can be seen as consequences of this unstructured management activity. Little's Law is an equation that shows the relationship between average WIP, cycle time, and the throughput of any production system (Hopp and Spearman, 1996):

$$WIP = (Throughput) * (Cycle\ Time)$$

Since Sort fills the capacity of their testers to the point of over-production, their throughput levels remain nearly constant; the ad-hoc control of WIP levels at Sort1, Sort 2A, and Sort2 can then be mathematically linked to cycle time variability shown in Figures 5.3.1 and 5.3.2.

When pressed for a formal methodology to WIP management, one Sort technician declared, "We try to push as much inventory into Sort2 to hedge ourselves against any surprise increases in [production schedules]. That way, our product is in a position to be delivered rapidly. In reality, we are always faced with changing mix of production volumes and sporadic wafer delivery from the fab, making any structured approach to WIP management extremely difficult."

5.10 Key Takeaways

In summary, Figure 5.10 breaks down the major directly-observed areas of waste alongside Intel's current management systems and the alignment of that waste's *eradication* with the fab's strategic objectives.

Figure 5.10 Comparison of the observed sources of waste at Sort 23

Area of Waste	Relative Level	Management System	Current Metric	Strategic Alignment
Direct Labor Productivity	Medium	MT Autonomy	Annual evaluations	Lower cost per wafer
Production Waiting Times	High	Industrial Engineering Models	Throughput times	Cycle time reduction
Equipment Utilization	Low	Base Tester Allocations	Weekly Utilization	Lower cost per wafer
Rework	Low	Industrial Engineering Models	Resort Rates	Lower cost per wafer
Delivery from the fab	High	Pull Station Reporting	Days behind (schedule)	Cycle time reduction
Overproduction to ADI	High	Tier priority catalog	None	Cycle time reduction
Inventory	High	Spare Tester Allocations	None	Cycle time reduction

Directly examining the operational characteristics of Sort reveals some key aspects of their past success. For example, the group's performance in minimizing rework and maximizing tester utilization has kept their operating costs well below their nearest internal rivals. And despite the variable influx of product volumes from the fab, Sort leverages base capacity estimates to continually allocate their tester resources towards shifting production schedules, enabling the Fab 23-Sort 23 collaboration to rarely miss a LIPAS goal.

However, the observed data also shows a significant waste of overproduction from lax scheduling. Decisions on allocating the resulting spare testers are unstructured across the shifts, and the management of inventory flow is largely reactionary. In evaluating the opportunities for mX improvement, *these sources of waste stand out*; cost per wafer is a deeply-entrenched measure of success at Sort 23, but cycle time reductions lie outside most of the current management systems (as well as the group's paradigm for success). The next chapter will critically examine these latent sources of waste through the use of mX thinking and principles.

6. HYPOTHESES FOR CHANGE TOWARDS THE IDEAL STATE

"It's easy to read these [lean rules] and think, "We've already done that. We have a book of standards; we've developed process maps for the flows; we know the customer of every process – so what's new?"
- Jamie Flinchbaugh (Flinchbaugh, August 2004, p. 3)

In striving to be the lowest-cost sorting operation at Intel, Sort 23 already typifies the focus on waste elimination and scientific improvement inherent in lean systems. Accordingly, many Sort personnel have often asked the question, "So what about mX is new?"

The answer to this challenge lies in a difference in the level of *depth* to which mX rules must be applied. Most of Sort's current improvements fall comfortably within current paradigms – that is, most everyone in Sort would agree and immediately recognize the value that these projects generate for their department. For example, it's relatively straightforward to see how an effort to speed up process times (or diagnose an equipment malfunction, or even rush to feed an idle tester) is valuable towards eliminating wasted capacity, raising throughput, and hence lowering the cost per wafer sorted. In contrast, the value in the following set of activities may not be immediately obvious:
- Experimenting with a strict WIP flow even when it requires an immediate increase in changeovers, more direct labor, and an overall decrease in available tester capacity.
- Adhering to a consensus production plan when yield unexpectedly drops and an ad-hoc effort would enable a quick boost in the delivery of a critical product.
- Strict adoption of another shift's maintenance protocol when there's a conviction that your shift does it more efficiently.

When applied regularly to the thousands of daily scenarios and decisions within a factory, the mX rules can generate a powerful shift in how a manufacturing organization sees itself and its quest for improvement. Structures, connections, and flows become thinking norms which ultimately demand experimenting with the kind of disciplined challenges and non-intuitive improvement activities outlined above.

In this chapter, we present and discuss a series of improvement suggestions for each mX rule; many of these hypotheses fall outside of Sort 23's paradigms for improvement yet are essential for reducing the cycle time of the manufacturing system. Critically, these ideas all create palpable tension towards an ideal state for Sort, where they "deliver what their customer wants, when they want it at the price they want with zero waste and everyone safe." (Flinchbaugh, March 2004).

6.1 mX Rule #1: Structure Every Activity

As an observed current reality, there is inconsistency in how testers are allocated; the unstructured nature of this decision activity leads to undisciplined WIP levels and ship-ahead executions which vary across product priorities. In the ideal state, the right tester allocation is unambiguous and obvious; the number of testers at Sort 1, Sort 2A, Sort 2 for each product leads to a predictable result, namely the consistent flow of product to steadily fulfill unambiguous production goals.

In order to conceptualize a more-disciplined approach to tester allocations and resulting flow, a few realities of Sort's operational context must be considered. Figure 6.1.1 is a depiction of the flow of P803 products in terms of a familiar "water & bath-tub" system. Nozzle #1 (i.e. the highly-variable wafer delivery from the fab) is outside of Sort's control, whereas Nozzle #2 (i.e. the ratio of testers on Sort1 & Sort2) and Nozzle #3 (i.e. production volume) fall within Sort's direct and indirect control. Without reliable information about fab delivery, however, Sort is limited to achieving improved flow times *only after product enters the Sort1 operation.*

Figure 6.1.1: Sort's Inventory Flow and Points of Control for the 803 Process Technology

Nozzle #1: Fab Delivery
- outside of control
- highly-variable

Nozzle #2: Ratio of Sort1-to-Sort2 Testers
- direct control

Nozzle #3: Production Volume
- moderate direct control with over-ship accruals
- indirect control with collaborative planning

Sort1 WIP

Sort2 WIP
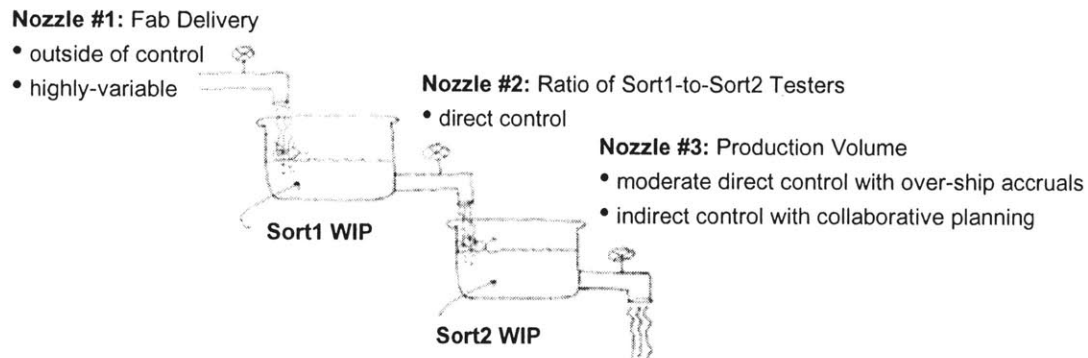
With possible inventory control bounded by the start of Sort 1, several pull methodologies can be considered for the structured release of work flow through downstream operations. As one of the simplest forms of WIP control, Figure 6.1.2 shows a conceptual Constant Work in Process, or CONWIP (Hopp and Spearman, 1996) system. Unlike a kanban system, where movement from one operation to the next is strictly dictated by a finite release structure, a CONWIP system allows the release of a unit of new material into the first operation when a unit of material completes the last operation of the process. This additional flexibility can be easier to manage than a pure kanban system, especially with the complexity of Sort's re-entrant process flows.
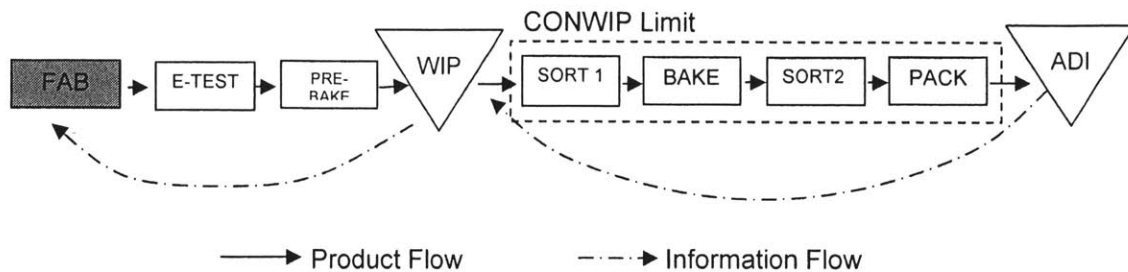
In general, there are many benefits from a control of WIP (pull) versus a control of throughput (push) system. From a flow perspective, Little's Law dictates that controlling WIP levels will translate into steady cycle times. In most cases, the introduction of WIP discipline forces operators to undergo activities and improvements to maintain a consistent flow of correct product mix towards fulfilling production goals; production problems, such as testing failures or re-sorting, result in the visible stoppage of progress and must be dealt with immediately.

In the case of Sort 23, such reductions in cycle time variability would not only help to align the group's execution cycle (currently 14 days) with the companies' larger planning cycle (every 7 days) but would also propel them as improvement leaders within the Fab/Sort organization. Strategically, total cycle time through both the fab & sort operations *must be reduced* to meet the increasingly-demanding lead time requirements of the communications market.

From a tactical perspective, the maintenance of constant WIP levels (versus throughput) can also be easier to manage and coordinate across shifts. Unlike capacity estimates, WIP movement is directly observable and can thus be powerfully enforced through visual management techniques. Secondly, Sort's capacity estimates depend on numerous uncertainties such as die yield, run times, tester utilizations, and changes in product mix. These uncertainties are hard to characterize and hence make precise capacity estimates problematic (Hopp, Spearman, and Woodruff, 1990, p. 883).

6.2 mX Rule #2: Clearly connect every customer-supplier relationship

This mX principle advocates the proliferation of *binary* customer-supplier connections where "there is only one way to make a request, and any request can only mean one thing". In this way, responses to requests are direct, immediate and exact, and the connection should be self-diagnostic; if a request fails, the person responsible for the link is immediately signaled (Flinchbaugh, March 2004).

Through this ideal state lens, a significant opportunity exists to clearly connect the product supply requirements of Sort with the fab production priorities. In exploring the possibilities for binary request signals, the inventory before Sort 1 becomes very

strategic; as an effective buffer between fab delivery and Sort production, this WIP can be viewed as a supermarket shelf which is steadily drawn upon to feed Sort's production flows. Of course, consistent downstream flow through the remaining Sort operations must be maintained in order to retain this connection; undisciplined flow would render the level of Sort1 inventory meaningless as a communication device.

A more subtle application of the customer-supplier principle extends to Sort's obligation to deliver their output in die quantities, not wafers. In most cases, Sort has little to no impact over the die yield on the wafers; they merely process wafers through standardized operations and report the fallout die and line yield data back to the fab operations. Yet Sort's delivery performance is measured in terms of fulfilling die schedules (LIPAS), so that any downward trends in product yield encourages Sort to boost that product's output to compensate for the unexpected losses. As one Sort Technician boasts with pride, "We really go to extremes down here to 'save the day' for the fab. If die yield is down, we can end up splitting lots, performing massive tester changeovers, and jumping through all kinds of hoops to make a production target." Reciprocally, a fab manufacturing manager once conceded that, "the Sort floor is the last opportunity to make up for any kind of low production. If they have the leverage to meet LIPAS, then they'd better deliver to it [for the benefit of both Fab 23 and Sort 23]."

But as a supplier of quality information, Sort has a major responsibility to provide clear and directed feedback to the manufacturing processes of the fab; otherwise, the true causes of defects can be obscured and the yield inefficiencies may last. The practice of holding Sort accountable for die schedules clearly disrupts this customer-supplier relationship, as Sort's ongoing efforts to compensate for variable fab yields diminishes the criticality of their quality feedback and enables downward trends to endure. In fact, a major premise of the mX rule of customer-supplier relationships dictates that this feedback relationship must be self-diagnostic; if Sort fails to supply its fab customer with appropriate yield data, the consequences should manifest themselves immediately in a causal result, namely a lower delivery of die to ADI.

Notably, W. Edwards Deming's philosophies of continous product and process improvements center on this same kind of strict alignment of measurement and behavior. In fact, Deming's passion for creating "constancy of purpose" and a climate free of finger pointing (which blocks cooperative identification and solution of problems) ring powerfully true today when viewing the current reality of the Fab-Sort relationship (Deming, 1982, p. 17).

Hence, a measurement scheme which rewards Sort for consistent delivery of wafers, not die, seems appropriate. In turn, fab delivery performance could be evaluated by what its downstream customers clearly request, namely, the arrival of product-based wafers as Sort requires them and the consistent delivery of high-yielding die to effectively replenish ADI stock levels.

6.3 mX Rule #3: Specify and simplify every flow

In reviewing the observed activities of three Manufacturing Technicians, I found that a considerable amount of their time and energy is dedicated to monitoring and coordinating WIP flow across a common fleet of 63 sorting testers. Figure 6.3.1 shows a representation of the current networked product flows across Sort's fleet of testers. The network flow of materials, people, and information becomes very complex and requires electronic notes, spreadsheets, email reminders, and daily pass-down meetings to ensure the 'best product' is processed at the right sorting operation on each tester as they become available.

Figure 6.3.1: Current networked product flows across Sort's sixty VHP testers.
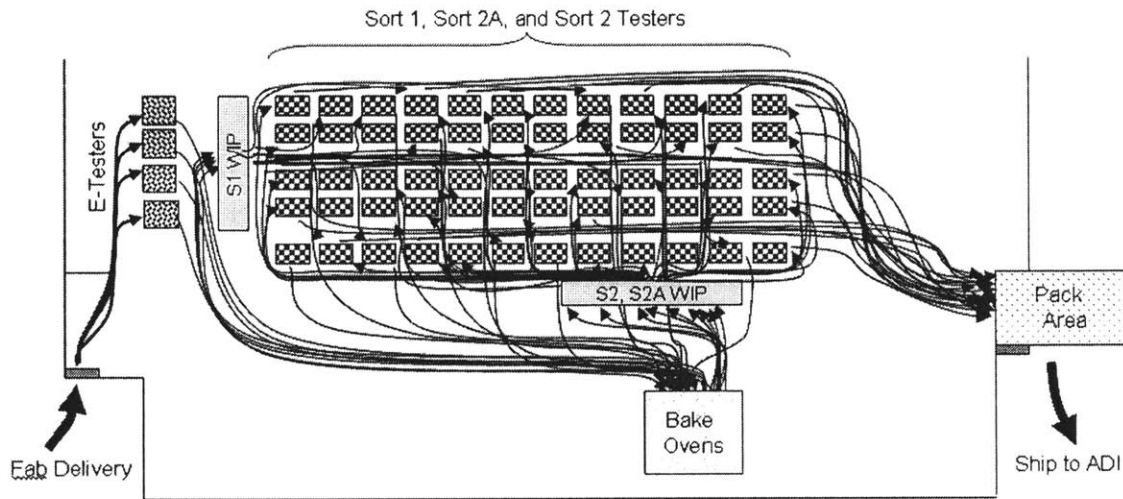
By contrast, one might imagine a production flow in which each product follows a standardized path across testers dedicated for Sort 1, Sort 2A, and Sort 2 operations. Figure 6.3.2 is a visualization of how one such layout would work given a fixed location of the baking ovens. Segregating testers by sorting operation simplifies Technician traveling & coordination activities as well as provide visual clarity on capacity allocations and the production status of the floor. Of course, differences in run times and product mix will vary the number of testers required for each sorting operation on any given week. Hence, a small cell of flexible testers is necessary to accommodate these current realities.
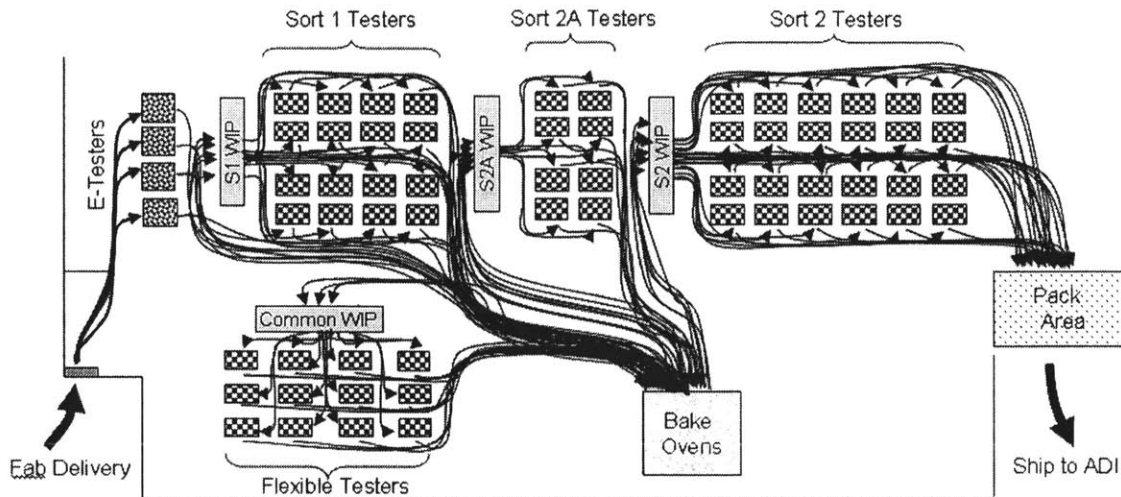
Since a majority of the testers will no longer require changeovers in sorting operations, tester availability should also improve (the average changeover for a product requires $1/3^{rd}$ the downtime of the average changeover in sorting operation).

6.4 Key Takeaways

The application of the mX rules to Sort's current reality creates tension for many new possibilities for cycle time improvement and associated change towards to the Ideal State. Of course, managing change is difficult in any organization, and the adoption of many of these ideas is highly contingent on all the Sort shifts valuing the new common way or process more than they value their current practices.

Next, we will see how high agreement around experimentation, with an example of the implementation of CONWIP control, enables actionable first-steps towards change while promoting systems thinking and tension for further improvements.

7. EXPERIMENTATION & IMPLEMENTATION

"How often have we made a change, whether technical or organizational, and not known whether it actually solved a problem or better served our customer? How often have we rolled out a change as THE fix? How much more powerful would it be to roll out all changes as controlled experiments, acknowledging that we will continue to improve upon them, or even totally abandon them if something better is discovered? Would people be more accepting of change if they knew that all change is an experiment, and that each new activity, connection or flow will be scrutinized for effective and efficient delivery of results? And that they were empowered to make *changes to the change*?"
- Clay Carlson, mX Core Team Leader

As one of the major leaders of the mX initiative, Clay Carlson has firsthand experience with the challenge of implementing lean improvement ideas on the floor. By framing proposed changes as *experiments*, though, mX leaders can not only create a failsafe environment for learning but also promote a bias for immediate action. At Fab 23, improvements are often framed by a scientific approach (i.e. the Plan, Do, Check, Act Cycle), but then never leave the 1st stage due to fruitless efforts to plan a perfect course of action. Tactical experimentation (and mX Rule #4) is seen as the pathway which propels teams out of the planning stage and into the realm of active learning.

7.1 mX Rule #4: Improve through experimentation at the point of activity

Here, an experiment is defined by an investigative change with an expected result that steps incrementally closer to the Ideal State of production. Structured by the scientific method, experiments at Fab 23 are defined by:

1.) An understanding of the current reality
2.) Clear instructions and clear expectations for the change
3.) Actual results that are formally checked (through data or reflection) against expectations, and then
4.) Captured learning and the continuous refinement of the experiment.

This approach was instrumental in implementing one of the aforementioned hypotheses, namely the tactical control of cycle times through a CONWIP release of work. With a common understanding of Sort's current reality, a cross-functional team of Manufacturing Technicians, an Operations Manager, a Planner, and an Industrial Engineer first collaborated on the instructions and expectations necessary for a controlled experiment of a pull control of flow through Sort.

7.2 Clear instructions and expectations for the change

As a first goal, the team targeted more-consistent cycle times from the start of Sort 1 through Pack/Ship for all their Tier 3 & 4 products within the P803 process flow. The realities of scarce delivery of Tier 1 & Tier 2 product and the additional processing complexity of P804 products defined this experimental subset as a necessary starting

point. *More-consistent* was defined as a range of 25% above or below the operational cycle times used by the Industrial Engineers and Planners to set Sort's delivery schedule.

With product categories and target goals selected, Little's Law proves useful again in structuring CONWIP limits. If we assume that adequate capacity is available to meet the scheduling demands of all products, then we can define a target WIP level Wj for each product j that is in proportion to its throughput Tj by the sum of the operational cycle times Ci across i operations:

$$W_j = T_j \sum_{i=1}^{\#\,of\,Processes} Ci$$

The definition of each product's throughput Tj becomes interesting in Sort's environment of over-production. Formally, all ship-ahead volumes should be allocated by the Tier Priority scheme; informally, Sort's back-end shifts will run spare testers on Tier 3 and Tier 4 product if a volume peak is imminent on those line item schedules. To accommodate this self-smoothing practice, the target throughputs Tj can be defined by the maximum scheduling volumes seen over a two-week horizon for each product. For example, Figure 7.2.1 shows how the throughput $T_{Trumbull\ 128}$ would be defined given Trumbull 128's schedule volume (forecasted for 7 weeks).

Figure 7.2.1 'Smoothed' Throughput Targets across a typical Product Schedule (Trumbull 128)
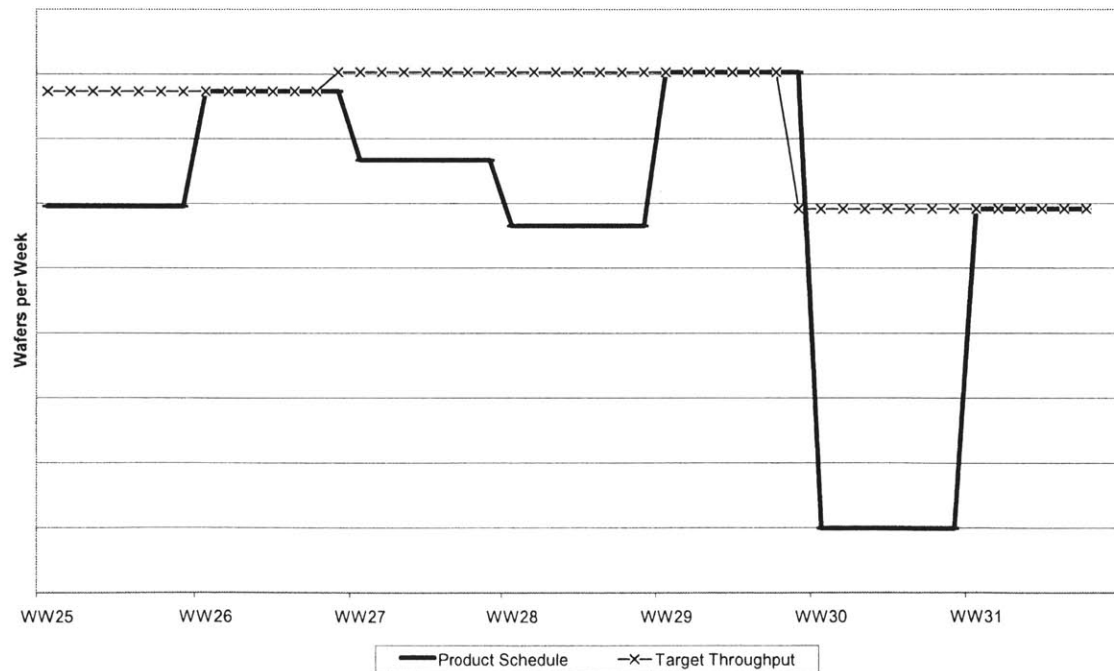
As an example of calculating target WIP levels, let's assume that the above analysis shows $T_{Trumbull\ 128}$ to be 100 wafers per week in WW29. We know that Trumbull 128 product requires four process operations within the CONWIP control (Sort 1, Bake, Sort

2, and Pack/Ship) and we also have target cycle time data for each of those operations from the Industrial Engineering Division (changed here for confidentiality):

Product	Cycle Time for Sort 1 (process only)	Cycle Time for Bake (wait and process)	Cycle Time for Sort 2 (wait and process)	Cycle Time for Pack/ Ship (wait and process)
Trum128	4 hours	8 hours	7 hours	1 hour

In this case, the target WIP level $W_{Trumbull\ 128}$ can be found for WW29 as follows:

$$W_{Trumbull128} = (100\ wafers\ /\ week) \times (1\ week\ /\ 168\ hours)(4\ hours + 8\ hours + 7\ hours + 1\ hour)$$
$$W_{Trumbull128} = 12\ wafers$$

With target throughputs Tj and corresponding target WIP levels Wj defined for all the P803 products, the Sort Area Coordinators were given a structured list of operating rules to dictate the regular enforcement of the CONWIP control across the Sort1, Bake, Sort2, and Packing operations (Figure 7.2.2).

Figure 7.2.2 Operating Rules for CONWIP control

If total WIP[1] is approaching an upper control limit for a product, perform the following actions as each Sort 1 operation on that product finishes:
- Change and load that finishing tester to perform any sorting operation on higher tier product (subject to available probe cards and WIP).[2]
- If no probe cards or WIP is available for a higher-tier product, change and load tester to a Sort 1 operation for any other product with equal or lower tier priority which has a total WIP level lower than its upper control limit.
- If all total WIP levels are approaching upper control limits across these equal or lower tier products, change and load that tester to perform Sort2 on that same product.

If total WIP is approaching a lower control limit for a product, perform the following actions:
- Changeover the next available spare tester that is currently performing a Sort 1 operation on a different product with a high total WIP to the Sort1 operation of the product with a low total WIP.
- If no spare testers are available, change the next available tester from running Sort 2 to run Sort1 on the same product.

[1] Total WIP is defined as the sum of inventory being processed in Sort 1, awaiting and in-process at Bake, awaiting and in-process at Sort 2, and awaiting and in-process at Pack/Ship. Total WIP excludes inventory awaiting Sort 1 and inventory awaiting and in-process at E-Test.
[2] Always maintain a minimum number of testers (1) at Sort1 for each product to provide uninterrupted yield feedback to fab.

As an example of how these rules work, let's assume we have a product A that is a tier-2 product (high-priority) as well as products B and C that are both tier 3 products (low priority). Imagine now that the amount of total WIP for product A is almost zero (as

every available lot from the fab has already been expedited), the amount of total WIP for product B is approaching its upper control limit, and the amount of total WIP for product C is well within its control levels:

Product	Tier Rank	Total WIP Level (relative to upper and lower control limits)
A	2	None
B	3	Very high
C	3	Medium

As the next tester operating Sort 1 on product B finishes, an Area Coordinator must decide how to re-allocate this capacity so as to reduce the inflow of WIP into product B's CONWIP band. First, she would attempt to allocate this tester to sort more of product A - yet the zero WIP levels for A make this action impossible. Instead, the AC would examine product C's total WIP; since C has medium WIP, the AC would change the finishing tester (Sort 1, product B) to support this same-tier product (Sort 1, product C).

These operating rules were carried out for a period of 5 weeks (July '04) to allow adequate collection of cycle time data. Critically, product tier priorities remained constant during this period.

7.3 Results versus Expectations

Figures 7.3.1 shows how these operating rules translated into structured WIP levels within the calculated CONWIP target bands. The 25% bandwidth control allowed the Area Coordinators flexibility in dictating rates of inventory change (through the 'nozzle' of Sort1-to-Sort2 tester ratio) and hence autonomy over changeover timing. Each excursion, or movement of WIP outside the control band, was investigated for root-cause, and nearly every case was succinctly corrected with an appropriate changeover.

Figure 7.3.1 Pilot CONWIP Control (Trumbull 128)
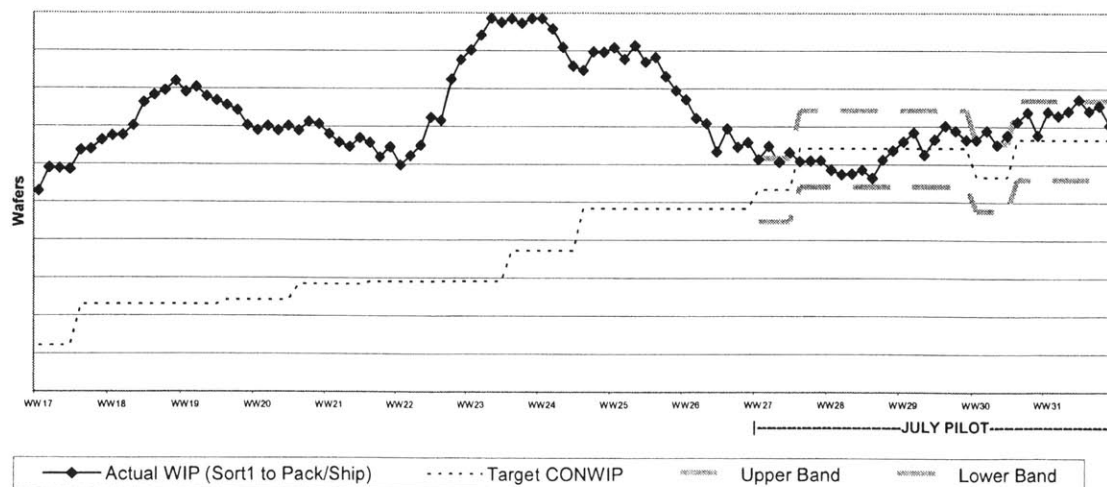
50

Figure 7.3.1 Pilot CONWIP Control (Trumbull 64)

| ——◆—— Actual WIP (Sort1 to Pack/Ship) | · · · · · ·Target CONWIP | Upper Band | Lower Band |

Figure 7.3.1 Pilot CONWIP Control (Trumbull 32)

| ——◆—— Actual WIP (Sort1 to Pack/Ship) | · · · · · ·Target CONWIP | Upper Band | Lower Band |
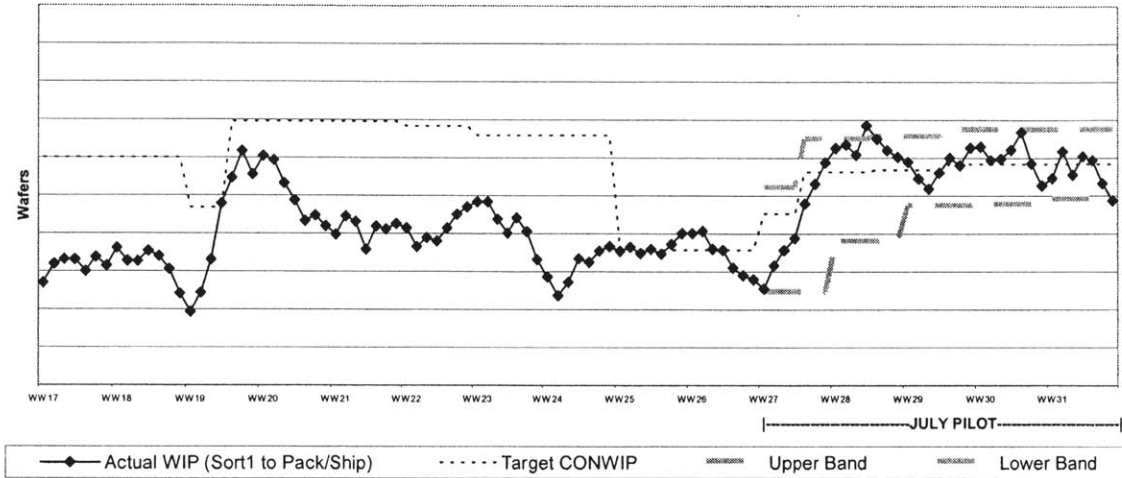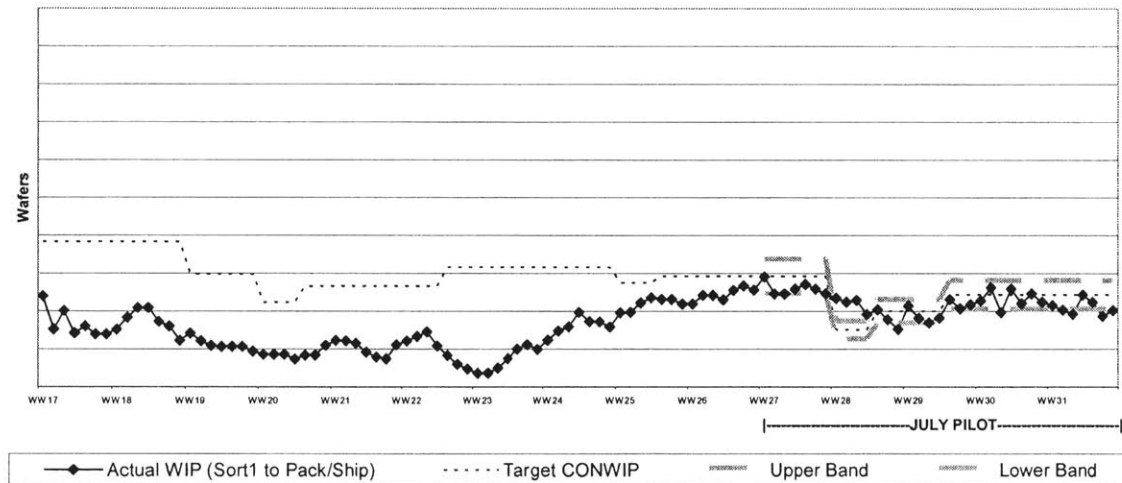
By structuring the level of inventory in proportion to throughput volume, Sort expected decreased cycle time variability for each of these product families. Figures 7.3.2 show a comparison of the cycle times for these products for both a pre-pilot period (June '04) and during the pilot (July '04). Examining the cycle times of the Sort operations before and during the pilot CONWIP control reveal a few key points:

- The major differences in cycle times across the operations stem from the variable queue times before Sort 1 and Sort 2. In the case of Trumbull 128, the total cycle time through Sort seems to have remained unchanged; decreased wait times before Sort 2 have been substituted for increased wait times in front of Sort 1. Note that evidence of this zero-sum trade-off is only clear when fab delivery-rates are roughly consistent.

51

- The CONWIP control dictates the cycle time from Sort 1 to Pack/Ship be consistent across all Tier 3 & 4 products. Here, the Trumbull 64 and Trumbull 32 queue times before Sort2 *increased* to match the common drumbeat. And although the variability in Sort2 queue time looks equivalent for Trumbull 64, we see a solid reduction in the same variability for the cases of Trumbull 32 and Trumbull 128.

- Since P804 products and Tier 1/ Tier 2 P803 products were outside the scope of the experiment, no change was expected for their cycle times. While there appeared to be no significant change in queue times in the P804 product flows, interestingly, Trumbull 256 (the only P803 product above a Tier 3 status) shows July queue times which are markedly lower than comparable June levels. It would seem the prescriptive operating rules (with a structured priority for allocating spare testers by Tier category) translated to faster flow of a priority product i.e. more testers were immediately placed on Trumbull 256 product versus building up excessive Sort 1 or Sort 2 inventory on products with lower Tier priority.

Figure 7.3.2 Reduced Variability in Sort2 Cycle Times
(Trumbull 128 in June '04 vs. July '04)
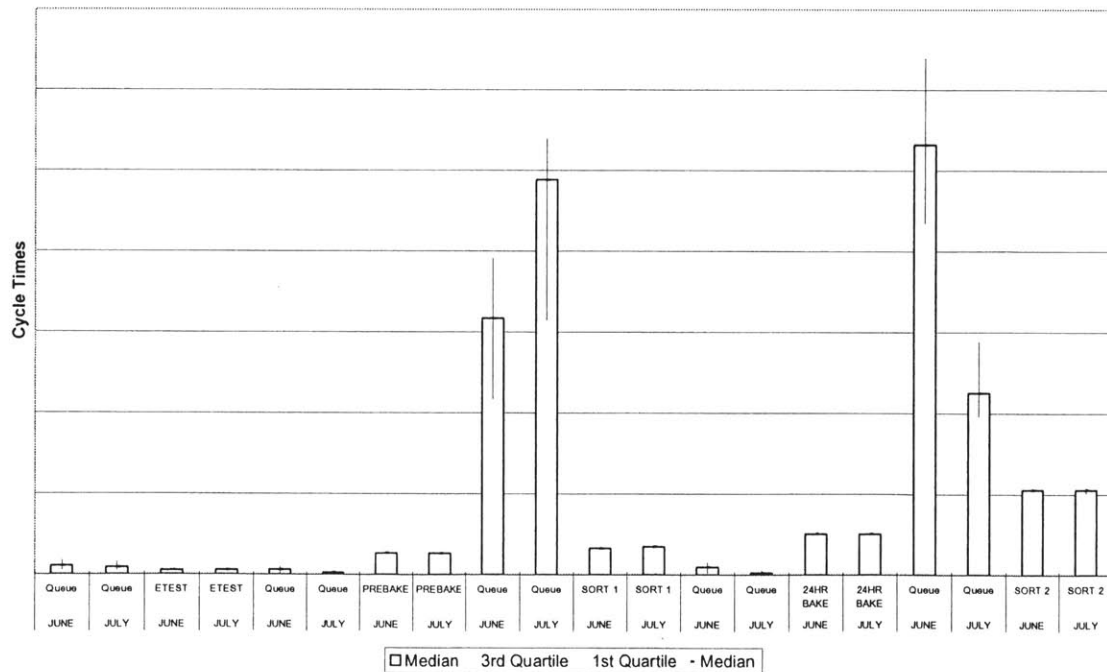
Figure 7.3.2 Reduced Variability in Sort2 Cycle Times
(Trumbull 64 in June '04 vs. July '04)
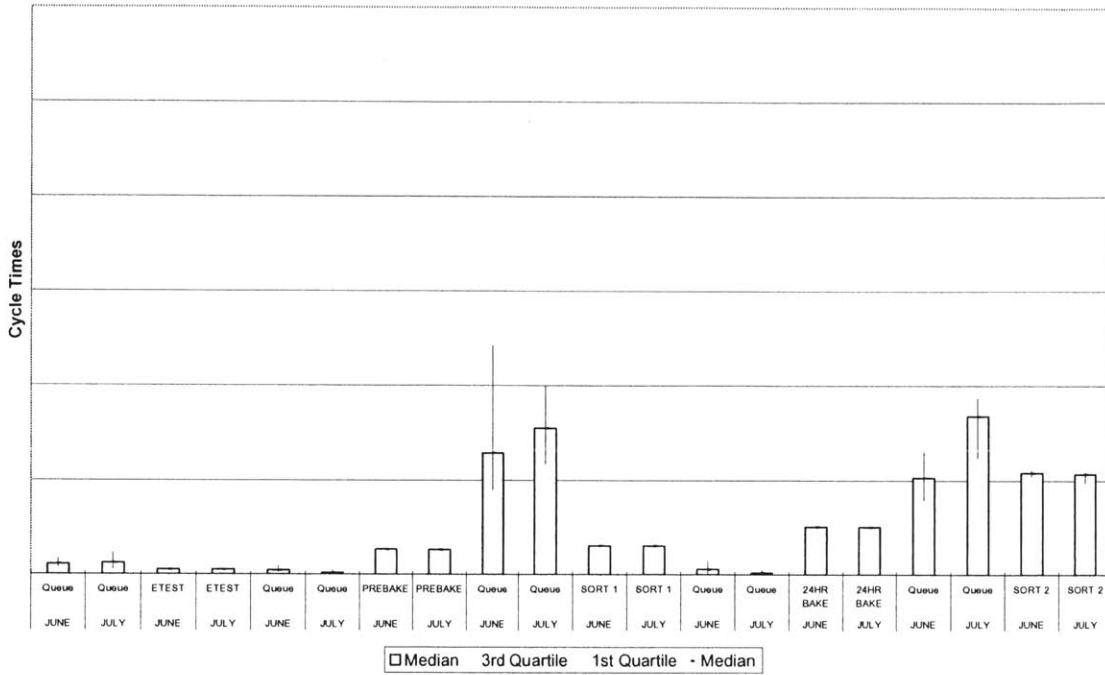
Figure 7.3.2 Reduced Variability in Sort2 Cycle Times
(Trumbull 32 in June '04 vs. July '04)
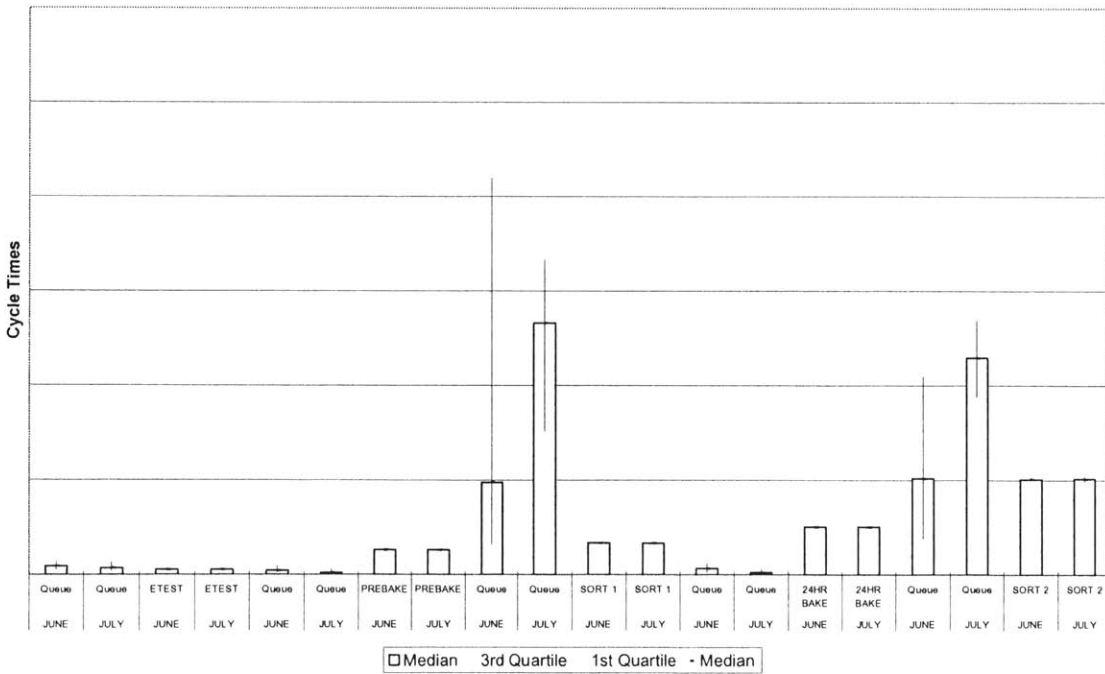
Figure 7.3.2 Reduced Variability in Sort2 Cycle Times
(Trumbull 256 in June '04 vs. July '04)

Queue	Queue	ETEST	ETEST	Queue	Queue	PREBAKE	PREBAKE	Queue	Queue	SORT 1	SORT 1	Queue	Queue	24HR BAKE	24HR BAKE	Queue	Queue	SORT 2	SORT 2
JUNE	JULY	JUNE	JULY	JUNE	JULY	JUNE	JULY	JUNE	JULY	JUNE	JULY	JUNE	JULY	JUNE	JULY	JUNE	JULY	JUNE	JULY

☐ Median   3rd Quartile   1st Quartile   - Median
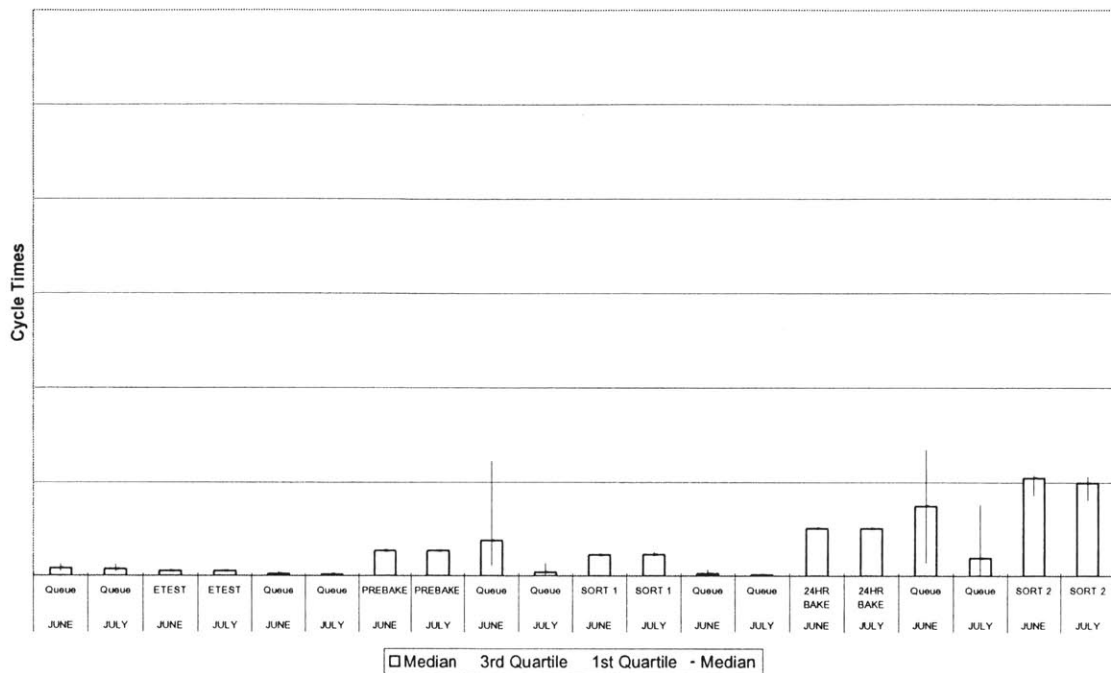
Per original expectations, the number of setups shows some increase during this pilot period (Figure 7.3.3). Product changes appear to make up a majority of the increase, consistent with the structured Area Coordinator operating rules (rules give priority in changing product over a sorting operation).

Figure 7.3.3 Setup Changes on VHP Testers

☐ Change in Product   ☐ Change in Sort Operation
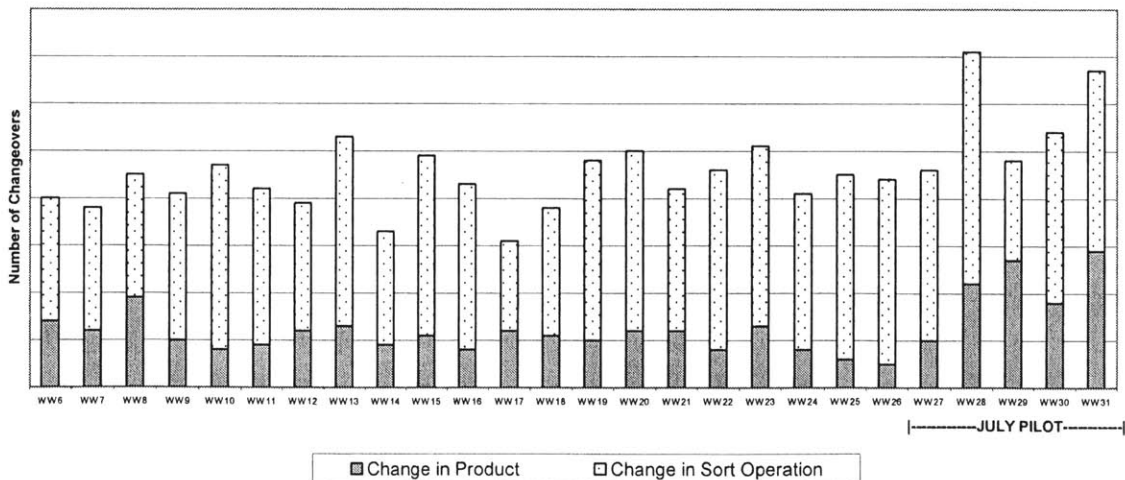
As a percentage of total availability, however, changeover time remained minimal. Sort showed consistently high utilization metrics well-above internal benchmarks during the period (Figure 5.4.1) and executed a LIPAS of 97% (out of the total 33 line items, only 1

product delivery was missed due to a disposition hold stemming from an error in a new testing program).

Notably, most Manufacturing Technicians viewed the increase in tester changeovers with criticism. Changeovers are currently cumbersome and time-intensive, and increases are commonly seen as a burden on Technician availability. As one Technician argued, "Where is the waste reduction in creating more flexes? If anything, we need to decrease changeovers [to minimize waste] – both to maximize utilization and reduce the labor content in the job."

7.4 Learnings and Refinement

As a first experiment in structured WIP control, the Sort floor executed the pilot change without terrible difficulty; one might argue that simply structuring these products flows to be consistently within the given planning times seems trivial. However, the real power of this change is not to be found in the absolute results but in how those results are interpreted, learned from, and directed towards the next level of refinement.

For Sort, there are several avenues available for continuous improvement upon this WIP control. A principle opportunity lies in extending the principles of pull across Sort's entire product suite. While Tier 1 and 2 products may be continually subject to expedited, variable flows, structuring the queues in P804 product flow seems straightforward. With one additional Bake-Sort process loop, the natural CONWIP level would begin at the start of Sort1 processing and extend to Pack/Ship, thereby controlling to total queue time in front of Sort2A and Sort2 to be in constant proportion to throughput.

Once the floor has demonstrated steady cycle times across its products, the natural next step is to start incrementally decreasing target cycle times below those anticipated by planning. In this manner, cycle time will systematically be lowered, forcing an increased diligence in WIP control and consensus throughput targets across front-end and back-end shifts.

Perhaps most challenging will be a final push to decrease the WIP control bandwidths. Set as a percentage of target throughputs, narrowing this range would increase the product changeovers, the labor content of the Manufacturing Technicians, and raise the criticality of each setup timeliness and quality.

7.5 Key Takeaways

As these next refinements are carried out, Sort may run into operating constraints that will lead them to better-coordinate their structured WIP flow with both back-end fab operations along with the current methods practiced by Planning and Industrial Engineers. In the longer term, all the operational areas at Fab/Sort 23 will feel pressure to lower their cycle times if they want to compete successfully in the communication market. So as cycle time becomes a more critical indicator of performance, Sort should

also be in a position to coach the rest of the fab organization from their experience with structured WIP management.

Of course, many barriers will likely be encountered along the way. It is helpful then to highlight many of the organizational barriers to 'lean change' experienced during the CONWIP implementation as representative of the larger culture and complexity of Fab 23.

8. BARRIERS TO MX AT INTEL

"My greatest challenge [is] "to lead the organization as if I had no power." In other words, shape the organization not through the power of will or dictate, but rather through example, through coaching and through understanding and helping others to achieve their goals. This, I truly believe, is the role of management in a healthy, thriving, work environment."
- Gary Convis, President, Toyota Motor Manufacturing Kentucky (Convis, July 2004)

In order to replicate the mX changes in Sort to other operational areas at Fab 23, a close examination of the barriers encountered with the CONWIP control change seems highly appropriate. In this chapter, these barriers will be described in the larger context of the company's organizational design and culture.

8.1 Organizational Design

We have already examined the case where Sort's operating metrics (i.e. weekly tester utilization, subject to delivery requirements) did not immediately align with the strategic goals of the mX project. Consequently, the latent wastes of overproduction remained safely outside Sort's paradigm of success. Notably, Intel has recently dedicated an internal team of resources towards re-examining the company's fab-wide metrics in an effort to re-align daily behaviors with the company's critical performance objectives.

Additionally, we can highlight some organizational barriers that are intrinsic to the fab's formal organizational structure and the design of the MT job.

8.1.1 Organizational Structure

Sort's formal organization (Figure 4.2) is representative of many of the other functional areas in the fab. A key aspect of mX is delivering improvements through standardized work rules that enable coordinated improvements across multiple shifts. However, the management structure of the floor is succinctly divided; oftentimes Operating Managers run their shifts in a unique (and sometimes contradictory) fashion, and rarely meet in person to facilitate coordination. Furthermore, structured communication between the front-end and back-end shifts is limited to emails, web-tools, and electronic post-it notes - leading each group to finger point when production problems occur (thereby blocking cooperative identification and solution of problems). As such, any attempts to standardize work will be sure to require some external resources to coach, reinforce, and coordinate the increased discipline. In many ways, I found myself fulfilling this need during the pilot implementation of WIP control – revealing a very palpable gap between the fab's organizational design and mX.

Another challenge for executing mX activities within this organizational structure is the bureaucracy between functional divisions. As an example, I was tasked with procuring and installing a second clean-room printer in the fab to help reduce MT travel distances. A process-based view of this work would show a Byzantine system of hand-offs,

redundant processing, and delays as simple decision-items get passed back and forth among the Manufacturing, Micro-contamination, Manufacturing Systems, IT, Industrial Engineering, and Automation groups. Notably, none of these departments questioned the validity or value of the proposed change – in fact, everyone I worked with felt this printer should have been provided for the MTs a long time ago. Despite everyone's willing cooperation, the job took 56 work days from assignment to completion with usually one communication, meeting, or activity occurring daily. Imagine the difficulty of change when the value of a similar cross-functional task is not as widely accepted or understood?

8.1.2 Job Design of the Manufacturing Technicians

In order to promote flexibility and higher labor productivity, MTs formally operate under a "single-job" concept. Essentially, all MT's are expected to share the same job description, tool skills, and proactive attitudes about taking ownership of problems on the floor. In reality, many MTs will differentiate themselves by developing highly-desired skills in equipment maintenance and diagnostic firefighting. These "Super-MTs" are sometimes rewarded with promotions and additional training. Other aspects of the Technicians job (diligent WIP movement, team coordination, housekeeping, etc.) are activities which have not been historically rewarded and attract marginal respect from management.

This tension between official "single-job" policy and actual specialized practices creates an interesting environment for mX change. In one sense, informal job specificity would imply more individualistic practices and greater barriers to standardized work. On the other hand, the ability to coordinate the few experts in each shift might be an effective team to spearhead structured practices. Certainly having people with *deep expertise combined with breadth of experience* is a terrific asset when presented with the systems-thinking challenges of mX. Nevertheless, the nondescript treatment of routine tasks presents a significant barrier when petitioning commitment for Visual Controls, 5S, or Standardized Work Instructions.

8.2 The Intel Culture

8.2.1 The Different Symbols of mX at Fab 23

The mX project holds different symbolic meaning for different people in the fab. In questioning the top leadership at F23, one hears that mX symbolizes great *promise* as it may help capture and extend the "special sauce" of recent cost-cutting successes to the rest of the factory operations and the MT environment. Some middle managers feel mX is about "working smarter, not harder" and is essentially an opportunity for them to help create more *work-life balance* among their subordinates (if I reduce the waste in their jobs, my people will have more time to spend outside of work). But if one were to directly ask those whom mX impacts the most – the MTs on the floor – one might hear two criticisms:

1.) mX is about *somebody taking credit* for what Technicians have been doing all along (suggesting a breakdown in communicating the mX message to the floor)
2.) mX is a *'flavor of the month'*; the initiative will never last without immediate, tangible results to warrant the resources

This pressure on results has created a very different symbol of mX within the mX project leaders: one of *battle*. In communicating the vision of mX at Fab 23, one mX leader writes, "Human beings are the value. Profit is the goal and our reason for existence. Waste is the enemy. *mX is the weapon.* In our continuous reach for the goal, we will safely pursue and vanquish the enemy with *the weapon of mX*." In the same spirit, all mX Champions wear badges that boldly display a sword emblem and the words "Waste is the Enemy" on them. Perhaps the symbol reflects the leaders' frustration in facilitating systematic waste elimination among the MTs; the natural difficulty and hard work of negotiating a change on the floor combined with the pressure for results could easily be abstracted into a fab-wide "battle" for improvement.

The inconsistency in symbols across different groups also suggests an ambiguity in the way mX has been communicated to the organization. Formally, mX is broadcast as one of Fab 23's three 'Must-Do' Strategic Objectives:

1.) Sustain operational excellence on [current production commitments]
2.) Flawless execution on [a new process technology ramp-up]
3.) Make mX a way of life

The language of the last objective is important; working towards a certain "way of life" can mean different things to different groups (just as current *ways of life* can obviously be very different around the fab). The ambiguity of this communication may not only lead to different interpretations of mX but also makes the alignment of daily behavior with the strategy somewhat difficult. If an MT works extraordinary hard over the course of a week to help standardize the best WIP staging policy across shifts, has he just made mX a normal part of his job? If an engineer works to diagnose a yield disposition (a kind of work she has done for years) and successfully improves quality- is that not an mX way of life for her? Which is the behavior that supports the new mX objective, and why doesn't the other behavior qualify?

8.2.2 The fit of mX within the Intel's Values

The introduction of any cultural initiative which falls outside of Intel's current value rubric seems somehow absurd, as Intel's corporate-wide norms are powerfully articulated across the organization. Like cultural road-signs, the values of

| Risk-taking | Great place to work | Quality |
| Discipline | Results-oriented | Customer Orientation |

are visually reinforced on ID badges, office posters, calendars, coffee mugs, and even mouse-pads. Although most of these core values can be easily related to the mX project

(lean principles reinforce a similar emphasis on quality, internal customer-supplier relationships, structured activities, and job fulfillment) it still seems peculiar to suggest mX is on par with these long-time pillars of Intel Excellence.

The only existing norm which requires a more-robust interpretation is the *results-oriented* pillar, as it typically represents Intel's emphasis on producing causal data or metrics as a proof of delivered value. The hazard for mX lies in the misalignment of delivery horizons; Intel expects *fast results* (usually as a means of survival in the electronics industry), whereas many lean improvements can take longer to materialize. The tension between "preoccuptations of today [without] ensuring improvements for tomorrow" (W. Edwards Deming, 1982, p. 18) is not a new phenomina, and Fab 23 surely struggles with this classic dynamic amid their rapid-clockspeed culture.

8.2.3 An aside about mX coaching

The role of those who may help coach mX thinking among their constituents is especially precarious given Intel's value of *fast results*. With a new production ramp-up imminent, nearly everyone at Fab 23 is very busy fulfilling the duties that are essential for this mission-critical success. However, one usually finds that coaching mX activities is in and of itself a full-time commitment; the energy and time required to teach and demonstrate new ways of thinking is certain to be exhausting and burdensome to those with a regular roster of other responsibilities. Worse, the return for this invested time isn't nearly as immediate as many of the other projects at Fab 23 (Figure 8.1.3 shows a conceptual delay with mX work relative to traditional Intel activities).

Figure 8.1.3 Performance delay for mX projects
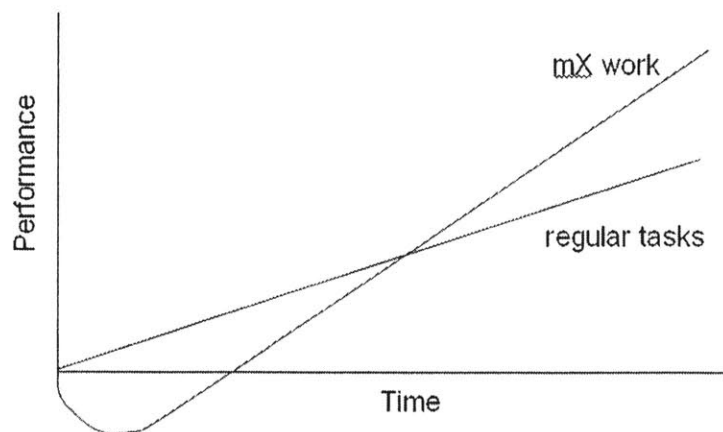
In order to escape the Catch-22 ("mX will free up my time through efficiency gains, but I don't have time right now to do mX"), mX coaches must somehow overcome the initial "add-on" barrier. Notably, Fab 23 does have mX Champions that are highly-effective at

being coaches and leaders of mX practices in addition to their full-time responsibilities. What is the "secret sauce" that enables them to be especially effective beyond the initial investment barriers?

8.3 Conclusions

Examining Fab 23 from both the organizational and cultural perspectives crystallizes some very clear challenges for the mX initiative. Central to these barriers may be Intel's critical need for hyper-speed results that align with their paradigms of success. As noted before, communicating mX as a "Must-Do Way of Life for 2004" sets a short-term expectation which is not only ambiguous (and thus difficult to execute-to, even *with* an alignment in performance metrics) but also asynchronous with the reasonable timeline for a lean transformation. As one planning manager predicted, "Everyone here is focused on individual projects completed within a year. Since there is no way lean is a one-year project, our reward system just won't support it." Perhaps even more insidious is the way speed can precipitate same-old thinking about problems. As Prof. Jan Klein writes, "The need for speed leads one to fall back on what one knows or has tried in the past." (Klein, , October 2004, Chapter 2 p. 7)

Furthermore, the *must-succeed* nature of this expectation, along with the heavy norm for results-oriented work, creates a very challenging atmosphere for lean exploration and learning; the value in practicing new ways of thinking is just not accepted without immediate data or results that move existing metrics. Worse, one manager expects that, "Lean will bring all sorts of benefits to Intel: higher quality, cycle time reductions, increased safety and labor productivity – it should move all our metrics." How can the critical reinforcing feedback for specific change efforts be targeted when it's being measured across such a diffused and grandiose set of metrics?

To be sure, changing ways of thinking (both one's own and others) is an enormous initial commitment that takes hard work and leadership savvy. At its essence, the demands of making mX change happen – the daily coaching, the inventing solutions as you go, the mix of tough negotiations with mutual cooperation – are challenges requiring both heavy systems thinking and personal persistence. If mX is to flourish at Fab 23, then leveraging the talent of the formal and informal leaders at the facility is most crucial. As one OM succinctly predicted, "If [our mX leaders] leave, then mX will fall apart with the rest of the middle managers…there will be no buy-in now or in the future without their sole leadership."

This page is intentionally left blank

9. TERMS AND ACRONYMS[6]

AC: Area Coordinators – the formal leaders within an MT team.

ADI: Assembled Die Inventory – an inventory staging point between Sort and Assembly operations.

CONWIP: Constant Work in Process – an inventory control methodology which caps the total amount of WIP that can enter a subset of operations.

Cycle time: Time elapsed between initiating and completing a process step or steps.

Device: One semiconductor product, such as a microprocessor or flash memory chip, which is manufactured on wafers. Each wafer can contain hundreds or thousands of devices.

Die: One individual device produced on a wafer.

Fab: A term for a fabrication facility, or factory, which manufactures semiconductor products.

Flash memory: Semiconductor devices which retain data after power is removed.

FSM: Fab and Sort Manufacturing – division covering all fab and sorting operations across Intel.

Lot: The batch size for wafer processing. The wafers at Fab 23 are generally processed in lots of 25 wafers each.

LIPAS: Line Item Performance As Scheduled –a binary (two-state) performance metric used in reference to product schedules which are either being met on time or not.

MT: Manufacturing Technician – worker responsible for lot handling, equipment maintenance, and housekeeping tasks within the fab and sorting operations.

mX: Manufacturing Excellence – representing all lean manufacturing principles at Intel

OM: Operating Manager – 1st-line supervisor for Manufacturing Technicians

P803: Process technology for Fab 23's flash memory products

P804: Process technology for a subset of Fab 11's flash memory products

---

[6] Many of the terms used in this section were adapted from the Semiconductor One Source Glossary at
http://www.semiconductorglossary.com/

Process technology: A standardized set of operations delivering a variety of products for Intel's customers.

Qualification: A process by which individual equipment is certified as capable of performing particular processes on particular products.

Sort: Process used to identify good die on each wafer and facilitate data collection for fabrication and test purposes.

Wafer: A silicon disc on the order of 1 mm thick and 200 mm in diameter. Wafers are used to form the substrate of a device. Each wafer may produce hundreds to thousands of devices, depending on the device size. Each individual device on a wafer is called a die.

10. REFERENCES

Andrey, Doug and Greenagel, John. "Global Semiconductor Sales Up 36.6% Year-over-Year." *Semiconductor Industry Association* (http://www.sia-online.org/). Press Release June 1, 2004.

Capelle, Michael J. "Improving Equipment Performance Through Queuing Model Applications." *LFM Thesis.* May 1995.

Convis, Gary. "Role of Management in a Lean Manufacturing Environment." *SAE International, Lean Monthly Column* (www.sae.org/manufacturing/lean/column/leanjul01.htm). Posted July 28, 2004.

Deming, W. Edwards. *Quality, Productivity, and Competitive Position.* Cambridge, MA: MIT, Center for Advanced Engineering Study, 1982.

Fackler, Gary. "Industry CEOs examine past to predict future." *Intel Circuit Intranet: Employee Communications.* Posted May 24, 2004.

Flinchbaugh, Jamie. Course Notes, *Lean Experience.* Novi, MI: Lean Learning Center. Taken March 15, 2004.

Flinchbaugh, Jamie. "Beyond Lean: Building Sustainable Business and People Success through New Ways of Thinking." *Lean Learning Center* (www.leanlearningcenter.com). Posted August 1, 2004.

Hopp, Wallace J, and Spearman, Mark L. *Factory Physics: The Foundations of Manufacturing Management.* Chicago: Richard D. Irwin, 1996.

Hopp, Wallace J, and Spearman, Mark L. "To Pull or Not to Pull: What is the Question?" *Maufacturing & Service Operations Management.* Vol. 6, No. 2, Spring 2004. pp. 133-148.

Hopp, Wallace J., Spearman, Mark L., and Woodruff, David L. "CONWIP: a pull alternative to kanban." *International Journal of Production Research.* Vol. 28, No. 5, 1990. pp. 879-894.

Intel in your Community, Colorado Page (http://www.intel.com/community/colorado/). Posted June 1, 2004.

Kanellos, Michael. "In a flash, Intel loses memory lead."*CNET News.com* (http://news.com.com/). Posted November 20, 2003.

Klein, Janice A. "A Reexamination of Autonomy in Light of New Manufacturing Practices." *Human Relations.* Vol. 44, No. 1, 1991. pp. 21-38.

Klein, Janice A. "The paradox of quality management : commitment, ownership, and control." Featured in *The Post-bureaucratic organization: new perspectives on organizational change* (Heckscher, Charles and Donnellon, Anne, editors). Thousand Oaks, CA: Sage Publications, 1994. pp. 178-194.

Klein, Janice A. *True Change: How Outsiders on the Inside Get Things Done in Organizations*. Jossey-Bass Inc., October 2004.

"Memory reloaded." *Electronicsnews.com* (www.electronicnews.com.au/articles). Released March 5, 2004.

Shingo, Shigeo and Robinson, Alan. *Modern Approaches to Manufacturing Methods: The Shingo System (Manufacturing & Production)*. Portland: Productivity Press, 1990.

Spears, Steven and Bowen, H. Kent. "Decoding the DNA of the Toyota Production System," *Harvard Business Review*. September-October, 1999. pp. 96-106.

Toyotaproductionsystem.net (http://www.toyotaproductionsystem.net/). Posted July 15, 2004.

Yoon, Jean and Sorid, Daniel. "Flash Memory Prices Slump as Market Showdown Looms." *Reuters* (http://www.reuters.com). Posted May 24, 2004.