FabTime Cycle Time Management Newsletter

Volume 15, No. 1 January 2014

Information

Mission: To discuss issues relating to proactive wafer fab cycle time management

Publisher: FabTime Inc. FabTime sells cycle time management software for wafer fab managers. We are currently working on a new JavaScript charting engine. This engine will be available in Patch105, due to be released shortly.

Editor: Jennifer Robinson

Keywords: WIP Management, Dispatching, Simulation, Lot Size, Financial Impact of Cycle Time

Table of Contents

Welcome

Welcome to Volume 15, Number 1 of the FabTime Cycle Time Management Newsletter! We hope that you all had a relaxing holiday season, and we wish you a productive and profitable 2014. We have one announcement in this issue, about the upcoming Fab Owners Association Collaborative Forum. Our tip of the month is about identifying the lot that has been on hold the longest in your fab (the first of a new series of step-by-step skills instructions that we are working on). We have two submissions from subscribers, one about foundries and cycle time, and the other about the impact of increasing lot size.

In our main article this month, we present the results of a series of simulation experiments conducted by FabTime's Mike Krist and Frank Chance. The goal of this experiment was to evaluate the impact of several commonly-used dispatch rules on linearity of shipments. The results showed that, of the rules tested, some had problems with WIP bubbles. This was initially less apparent under critical ratio dispatching than the other rules. However, in the presence of late lots, WIP bubbles were also observed under critical ratio. We believe that these results call for further investigation of line balance-focused dispatching. As always, we welcome your feedback.

Thanks for reading – Jennifer

FabTime

Tel: (408) 549-9932
Fax: (408) 549-9941
www.FabTime.com
Sales@FabTime.com

Community News/Announcements

Fab Owners Association Collaborative Forum: February 5-6

The FOA's second annual "Collaborative Forum" will be held in Santa Clara, California on February 5-6, 2014 at the Biltmore Hotel. It will include collaborative case studies submitted by our Associate members and will include participation from our Device Makers. More details, including a tentative agenda, can be found here. The Forum is only available to FOA members and associate members, or invited guests. FabTime will be attending the Forum, but will not be presenting this year.

FabTime welcomes the opportunity to publish community announcements, including conference notices and calls for papers. Send them to newsletter@FabTime.com.

FabTime User Tip of the Month

Identify the Lot that Has Been On Hold the Longest

We are in the process of developing step-by-step instructions for skills that we think may be especially useful to FabTime users. Today's skill involves identifying the lot that has been on hold the longest. To do this:

1. Search for "WIP Lot List" from the FabTime Charts list (under WIP Charts) and press the "Go" button next to the name of the chart.

2. Near the bottom of the main set of filters to the left of the chart, find the "Hold:" drop-down (just above the first "Go" button). Click the arrow on the drop-down list and select "Hold". Also make sure that the "Age:" drop-down displays "Current Opn". Press the "Go" button just below the "Que:" drop-down.

3. The resulting chart will display all yellow bars. The tallest bar will correspond to the lot that is on hold and that has been at its current operation the longest. If there are too many lots on hold to see the lot numbers on the chart, you can use the data table. Simply make sure that the data table is sorted by "OperationTimeIn" without the "Descending" box checked. Top row of the data table will correspond to the lot that is on hold and has been at its current operation the longest.

4. Click on "History" in the data table to see more details about the lot, including what operation it was at, or had just moved out of, when it went on hold.

Alternatively, if you would like to see the lot that had the most total hold time over a particular time period (e.g. during the past week), you can use the Hold Time List chart. Set the "From:" and "To:" dates to your time period of interest and hit enter. There will be bars (or data table rows) for all lots that were on hold during the time period. The height of the bar represents the total time that the lot spent on hold during the time period. Yellow bars indicate lots that are still on hold; green bars denote lots that are now off hold.

We hope you find these step-by-step instructions useful. If you have things that you would like to know how to do with FabTime, or things that you think are key FabTime skills that new users should learn, please send your questions or suggestions to Jennifer.Robinson@FabTime.com. Thanks!

If you have questions about this item, or any other FabTime software questions, just use the Feedback form inside FabTime's software. Subscribe to the separate Tip of the Month email list (with additional discussion for customers only). Thanks!

Subscriber Discussion Forum

Issue 14.03: Why Should Foundries Care about Cycle Time?

In response to Issue 14.03, an anonymous subscriber wrote: "Another input about foundries and cycle time is this. Offering lower average selling prices (ASPs) can compensate for not having competitive cycle time. Getting both at the same time is a bonus. In order to provide very competitive ASP, the production cost must be lowered. Part of the costs are utilities, chemicals, gases and test wafers. For those that expire based on a period of time, cost can be improved by running a higher number of product or passes through each tool. This increases the utilization of the fab and thus also increase the queue time and the overall cycle time.

However, if a foundry takes on a bigger share of the market and is able to control that market for a specific technology, then they may be able to achieve higher ASP because their customers are willing to pay for the technology. For more mature technology this may not be the case."

Impact of Increasing Lot Size

Another anonymous subscriber wrote: "We are planning to increase the lot size in our fab. Before we do that I want to identify the pros and cons and do some analysis. I understand that increasing lot size will increase cycle time but will it be possible to measure all the relevant metrics? What should I be looking for in terms of benefits and negative effects that increasing the lot size will have in the wafer fab?"

FabTime Response: We wrote about this quite a long time ago in the newsletter – see Issue 2.02 for details. It's a complex question. Some things improve when you increase lot size, and some things get worse. It's difficult to model because you have to consider things like what happens to the process time at per-lot tools. (Does it stay the same?) But our newsletter article has a short list of impacts to consider, so hopefully that will help. We welcome feedback from other subscribers. If this topic generates a response, we will republish and augment that prior article in our next issue.

FabTime welcomes the opportunity to publish subscriber discussion questions and responses. Simply send your contributions to Jennifer.Robinson@FabTime.com.

Dispatching and Line Balance

Introduction

Last year, FabTime undertook a dispatch simulation study. The purpose of the study was to determine which of three common dispatch rules would yield the most consistent shipments from the fab (i.e., the most balanced line). The results showed that each of the tested rules led to WIP bubbles, although this was initially less pronounced under critical ratio dispatching than under FIFO or shortest remaining process time. However, in the case where a significant portion of the lots were late at that start of the simulation, critical ratio also led to significant WIP bubbles. We did not specifically study a Line Balance dispatch method. However this study illustrates the potential benefit of researching line balance-focused rules in more detail.

Methodology

We compared three different dispatch rules: first in first out (FIFO), shortest remaining process time (SRPT), and Critical Ratio (CR). We looked at two different variants of CR, as well as at different compositions of late lots at the start of the runs. We used FabTime as a simulation tool because it provided the capability of showing WIP levels at hourly intervals. We viewed the changing WIP levels via a series of animated GIFS (though it is not possible to show this animation in the newsletter).

We did two primary sets of simulations, using a small test case in FabTime. We initially simulated the three dispatch rules with no late lots (Simulation 1). We later did a second set of experiments (Simulation 2) that included late lots. All experiments were done using a variant of the critical ratio rule called CR2 (defined in detail below). Key results were also validated against another critical ratio variant called CR1. (This naming reflects the naming of the rules in FabTime's dispatch module).

Assumptions:

- We pseudo-randomly assign due dates around the expected finish time of the lots.

- Arrivals come in once every two hours, with up to two hours of variability uniformly distributed.

- The process flow has 22 steps across 3 reentrant layers.

- 4 tools are modeled in detail as single-tool constraints, with lower utilization delay steps in between (to model intermediate operations). Key tools are visited 2-3 times each. Process time variability is uniformly distributed. All processing is per-lot.

- 48 lots are processed in each simulation.

Each scenario was simulated 2 separate times, and the results were charted in FabTime each hour until all lots finished processing. The total Theoretical Cycle Time was 68.65 hours for each lot.

The sample process flow is shown at the top of the next page. In the table, Arrival Rate the total arrival rate to the tool (reflecting the number of visits). No downtimes were modeled for these experiments.

Simulation 1: Results

For the first set of simulations, looking at FIFO, SPRT, and CR2, with no late lots, we did two replications for each dispatch rule. Variability between replications was minimal.

Operation	Photo	Delay	Etch	Delay	LiftOff	Delay	Dep	Delay
1	0.65							
2		6						
3			0.5					
4				5				
5					0.6			
6						6		
7							0.7	
8								5.5
9	0.65							
10		6						
11			0.5					
12				5				
13	0.65							
14		6						
15			0.5					
16				5				
17					0.6			
18						6		
19					0.6			
20						6		
21							0.7	
22								5.5
Total Time	1.95	18	1.5	15	1.8	18	1.4	11
# of Tools	1	15	1	15	1	15	1	10
Arrival Rate	1.5	1.5	1.5	1.5	1.5	1.5	1	1
Service Rate	1.54	2.5	2	3	1.67	2.5	1.43	1.82
Theo Utiliz	98%	60%	75%	50%	90%	60%	70%	55%

Figure 1. Simulated 22-Step Reentrant Process Flow

Observations from Simulation 1

■ FIFO dispatch results in the WIP being clumped together throughout the process.

■ SRPT has a large amount of WIP held at operation 01 until the WIP is drained from the end of the flow. This is because all WIP at later steps will always have a shorter remaining process time than WIP at operation 01.

■ CR2 has the most balanced dispatch of the three dispatch methods (though not perfect).

Conclusion: Under ordinary circumstances (without an excess of late lots), CR2 will result in a more balanced line than the other two methods. We also did other experiments with another version of critical ratio, CR1, and the results were similar.

A snapshot showing the WIP profile in the factory during the mid-point of the simulations is shown at the top of the next page, in Figure 2

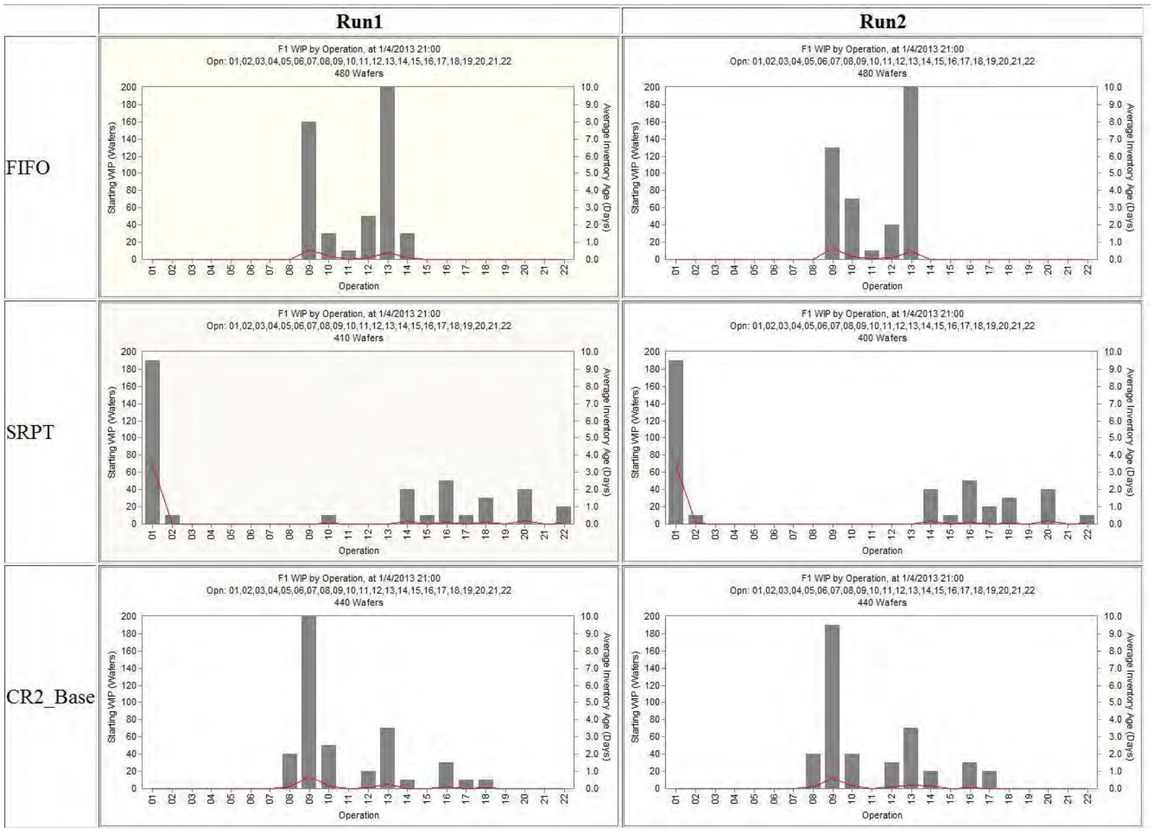

Figure 2. Snapshot of Results from Simulation 1, Mid-Point of Simulation

How Critical Ratio Is Supposed to Work, and Why It Can Lead to an Unbalanced Line

When a long downtime leads to a large number of lots being equally late, we see different behavior under CR. The lots will be (roughly) equally late because CR will have kept them each close to the planned schedule prior to the downtime. Here we again apply the CR2 dispatch rule from FabTime. Under CR2:

Slack = Lot Due Date – Now = Days Until Late

Raw score = -1 * CR2

If (Remaining Planned CT Days) = 0, then CR2 = 0

Otherwise, CR2 = Slack / (Remaining Planned CT in Days)

If all lots are equally late, and have similar slack values, then lots will be prioritized in reverse order of remaining planned CT. This is similar to SRPT (in which lots are ordered by the remaining total process time, across all future steps, with the lots closest to completion run first).

As long as there are any late lots, the fab will continue processing those in shortest remaining process time order (pulling from the end of the line). Once there are no late lots, however, if slack times are relatively similar, CR2 will favor lots with the longest remaining process time. Lots at the beginning of the flow will get pushed forward, until all of the lots again have similar CR2 values (with the goal of all lots running exactly on schedule). The simulation that we get from this ends up looking very similar to the one that we get when we run SRPT (see below).

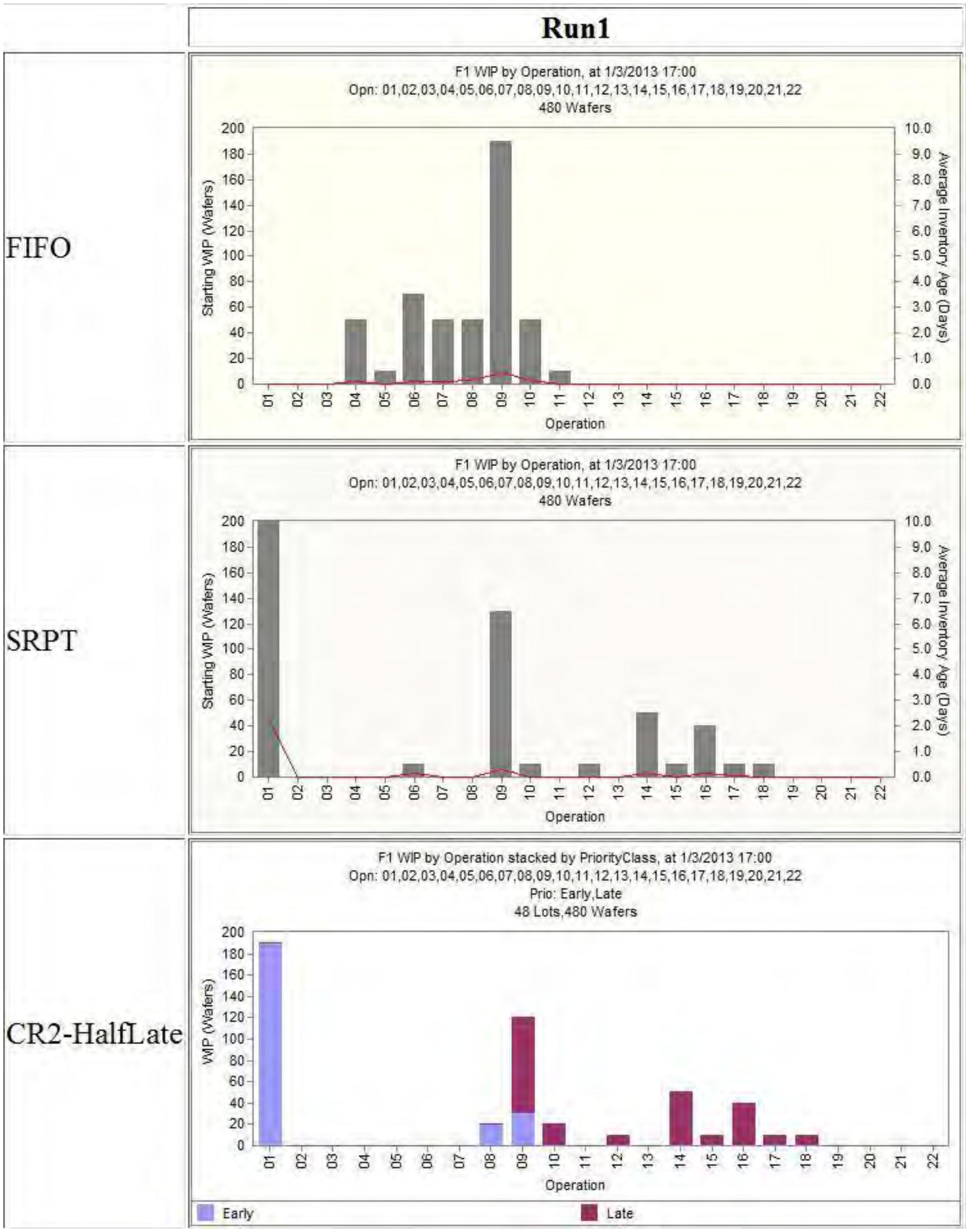

Figure 3. Snapshot of Results from Simulation 2, Mid-Point of Simulation

Simulation 2: Results

We repeated our experiments, but with one change. A significant portion of the lots are late at the start of the simulation. As described above, this mirrors the common real-world scenario in which a fab is recovering from a significant bottleneck downtime event. We looked at various percentages of late lots, with consistent results. The chart at the top of this page shows a snapshot of a simulation in which half of the lots are late at the start of the simulation run. (Due to the low run to run variability, we show only a single replication here). Purple represents late lots, blue represents early lots. Whether the lots are early or late is not relevant for FIFO and SRPT, and so is not shown.

Observations from Simulation 2:

- When half of the lots start out as late, what we see under CR2 is that the late lots get prioritized ahead of lots that are ahead of schedule. This is what the CR rule is supposed to do, but the result is unbalanced WIP.
- WIP piles up at the first couple of operations, and doesn't move from there until the late WIP drains from the end of the line. After that, the early WIP starts to move, but stays relatively clumped together for the rest of the flow. In real-life, what could happen next would be another downtime, causing additional lots to be late, and repeating the process.

Conclusions

For fabs that care about on-time delivery, Critical Ratio is an appealing rule, because it does prioritize lots that are behind schedule. In the absence of extended downtimes and reentrant flow, CR should work well. In the presence of extended downtimes, however, CR can lead to undesirable behavior – WIP bubbles that oscillate between the front and back of the line.

If you can instead focus on keeping your line balanced (roughly equal WIP throughout segments or sub-segments), you may find your fab easier to manage. You'll see fewer WIP bubbles, fewer starved bottlenecks, and lower cycle time variability. This lower cycle time variability will end up helping with on-time delivery. You can also still use CR as a secondary dispatch factor, of course, but we recommend that line balance be a primary goal of your dispatch policies. We anticipate future experiments that look at line balance-focused dispatch rules in more detail.

Acknowledgements

The experiments described here were conducted by FabTime's Mike Krist and Frank Chance. This article was derived from a summary report written by Mike Krist, modified for this publication by Jennifer Robinson.

FabTime would like to thank John Matthews and Richard Ledesma from Anadigics for discussions on dispatching and line balance that led us to conduct this analysis. No Anadigics data was used in the simulations.

Subscriber List

Total number of subscribers: 2798, from 444 companies and universities.

Top 20 subscribing companies:

- Intel Corporation (148)
- Micron Technology, Inc. (138)
- Maxim Integrated Products, Inc. (134)
- International Rectifier (121)
- Fairchild Semiconductor (100)
- GLOBALFOUNDRIES (84)
- Carsem M Sdn Bhd (73)
- Texas Instruments (72)
- ON Semiconductor (71)
- X-FAB Inc. (64)
- Western Digital Corporation (55)
- STMicroelectronics (54)
- Freescale Semiconductor (53)
- Analog Devices (52)
- IBM (52)
- Infineon Technologies (50)
- Skyworks Solutions, Inc.(49)
- Seagate Technology (42)
- Cypress Semiconductor (33)
- ATMEL (31)

Top 4 subscribing universities:

- Ecole des Mines de Saint-Etienne (EMSE) (12)
- Arizona State University (8)
- Nanyang Technological University (7)
- Virginia Tech (7)

New companies and universities this month:

- Alten Group
- Kovio Inc.
- Mestek Inc.

Sampler Set of Other Subscribing Companies and Universities:

- Affymetrix (1)
- Alpha-Sang (1)
- Arnstadt (1)
- Bourns (5)
- Dublin City University (3)
- Elettranova Engineering (1)
- EM Microelectronic Company (1)
- First Solar Inc. (2)
- Intersil (8)
- Littlefuse (1)
- Macronix International Co. (2)
- Mattson Technology (1)
- MIT Lincoln Laboratory (4)
- Nanophotonics Technology Center Valencia (NTC) (1)
- National University of Singapore (2)
- New Jersey Institute of Technology (1)
- OSRAM Opto Semiconductors (1)
- Selantek (1)
- Technical University of Eindhoven (2)
- University of Ulsan – S. Korea (1)

Note: Inclusion in the subscriber profile for this newsletter indicates an interest, on the part of individual subscribers, in cycle time management. It does not imply any endorsement of FabTime or its products by any individual or his or her company.

There is no charge to subscribe and receive the current issue of the newsletter each month. Past issues of the newsletter are currently only available to customers of FabTime's web-based digital dashboard software or cycle time management course.

To subscribe to the newsletter, send email to newsletter@FabTime.com, or use the form at www.FabTime.com/newsletter.htm. To unsubscribe, send email to newsletter@FabTime.com with "Unsubscribe" in the subject. FabTime will not, under any circumstances, give your email address or other contact information to anyone outside of FabTime without your permission.

FabTime® Software Capacity Planning Module

CP Configuration

We offer our capacity planning module for an additional monthly fee (on top of your regular FabTime subscription). This includes:

- Identification of the source of any additional data needed for the planning module.
- Automation of the process of importing the additional data into FabTime.
- Validation against client data.

Interested?

Contact FabTime for more information, or for a quote.

FabTime Inc.
Phone: +1 (408) 549-9932
Fax: +1 (408) 549-9941
Email: Sales@FabTime.com
Web: www.FabTime.com

Do you need to answer questions like:

- Given a target product mix, do we need any new tools?
- Given the tools that we have, and the products that we are running, how many wafers can we expect to produce?
- Given our existing set of products and tools, what happens if the product mix changes? Where can we expect bottlenecks?

Are you tired of maintaining a standalone capacity planning spreadsheet?

FabTime's capacity planning module leverages the data already stored in the FabTime digital dashboard software, to make it easier to build capacity planning scenarios. The only required manual inputs are:

- Weekly ships per product.
- Product line yield percentages.

FabTime uses route information from the fab MES and calculates UPH data (tool speed) based on actual performance. FabTime also uses tool uptime performance to estimate availability (though this can be overridden). These inputs are used to generate predicted utilization percentages for each capacity type. Detailed intermediate calculations (UPH, tool productive time, tool rework percentage, etc.) are also available (an example for one tool is shown below). All outputs can be easily exported to Excel.

Capacity Planning Module Benefits

- Eliminate the need to maintain offline capacity planning models.
- Automatically update capacity planning data to reflect new conditions (process flows, tool uptime characteristics).
- Quickly run scenarios to anticipate (and avoid) bottlenecks caused by product mix changes.

C Type	Output	Value	Notes
1XStep	Rework Moves/Week	21	2004-09-06 10:00:00 to 2004-11-15 10:00:00
1XStep	Total Moves/Week	12310	2004-09-06 10:00:00 to 2004-11-15 10:00:00
1XStep	Rework Ratio	0	Rework Ratio = Rework Moves / Total Moves.
1XStep	Productive%	61	2004-09-06 10:00:00 to 2004-11-15 10:00:00
1XStep	Availability%	76.26	Availability = Productive% + Standby%.
1XStep	Historic Utilization%	79.99	Utilization (Mfg efficiency) = Productive% / Availability%.
1XStep	Productive(Rework)%	0.1	Productive(Rework)=Productive% * ReworkRatio.
1XStep	Net Availability%	76.15	Net availability% = Availability% - Productive(Rework)%.
1XStep	Arrivals (Units/Hour)	79.36	Based on total plan WGR=2025
1XStep	Tool Quantity	8	1XStep#1 ... 1XStep#8
1XStep	UPH	15.02	UPH = (TotalMoves/ToolQty) / (Productive% * 168)
1XStep	Required Hours/Day	126.84	Required hours = 24 * HourlyArrivalRate / UPH
1XStep	Predicted Utilization%	86.75	Util = 100 * ReqdHours / (24 * NetAvail * ToolQty / 100)
1XStep	Max WGR	2334.22	MaxWGR = PlanWGR / PredictedUtilization
1XStep	Historic WGR	2457.8	(Non Rework Moves) / (OperationCount / ProductCount).