

## CLUSTER TOOL DESIGN COMPARISONS VIA SIMULATION

Kyungsu Park  
James R. Morrison

Korea Advanced Institute of Science and Technology  
Department of Industrial and Systems Engineering  
335 Gwahangno, Yuseong-gu  
Daejeon, 305-701, REPUBLIC OF KOREA

### ABSTRACT

The anticipated transition to 450 mm diameter wafers provides the semiconductor manufacturing industry an opportunity to consider new equipment designs that address issues associated with small lot sizes and high mix production. One candidate design is the linear cluster tool. Compared to traditional circular cluster tools, linear cluster tools have advantages such as high flexibility and greater productivity. In this paper, we develop a simulation of cluster tools with realistic parameters, which incorporates rolling setups and wet cleans. We use the simulation to study the effect of rolling setups and wet cleans with different lot sizes and train levels. For the simulation based on data from a BlueShift cluster tool in production, the linear cluster tool has 5.22% and 4.09% greater throughput with rolling setups and both rolling setups and wet cleans, respectively.

### 1 INTRODUCTION

The potential shift to 450 mm diameter wafers gives the semiconductor industry an opportunity to rethink existing equipment designs (Pillai 2006). In particular, new cluster tool architectures have been proposed that may be more appropriate for the high mix production environment anticipated in the 450 mm wafer era. For existing fabrications (fabs), new designs should also be considered; they may lead to competitive advantage. In this paper, we strive to assess the capacity of two different cluster tool architectures: circular cluster tools and linear cluster tools.

Circular cluster tools consist of process modules, an input/output port and wafer transport robots grouped into a circular chassis. Such tools are commonly used in semiconductor wafer fabrication. In contrast, linear cluster tools were recently developed to provide high flexibility and increased productivity. Such tools consist of a collection of several links and one or two input/output ports. Each link has its own wafer transport robot and thus the tools may achieve higher throughput. This configuration simplifies robot scheduling and reduces the footprint with two input/output ports. Yi et al. (2007) analyzed throughput and robot schedules in steady-state. Since links are completely isolated and possess their own vacuum environments, linear tools allow dynamic *rolling setups* from link to link. Radloff et al. (2009) studied first wafer delay (FWD) and setup times. They recommend rolling setups as a method to improve fab cycle and equipment productivity. Meulen (2007) discussed various other beneficial features of linear tools.

In an effort to quantify the practical capacity differences between these architectures, we conduct simulation studies. Using process parameters from real production tools, we develop flow line models for the two architectures. As we will assume that our tools are process bound, it is reasonable to assume that there is a robot control policy giving steady state takt time equal to the process time of the bottleneck module plus any unavoidable robot overhead. Examples of such policies are the backward sequence in a single-arm cluster tool with constant travel time (Dawande et al. 2007). As this robotic overhead is small

in comparison to the process times we consider, we will simply ignore it. For transient behavior, there has been little research on cycle time and robot scheduling. We use the same approach and ignore the robot. Such flow line models have been claimed to estimate the throughput of production clustered photolithography tools to within 1% (see, Morrison (2011)). They are thus acceptable models. We simulate three cluster tools to quantify the effect of rolling setups and wet cleans. As we are ignoring robot overhead in both classes of tools and linear cluster tools have more robots (and thus we anticipate less average robot overhead), our results will be a conservative estimate of the advantage enjoyed by the linear configuration.

Our simulation study shows that the linear cluster tool's rolling setups can enable more throughput than for circular tools. For models of practical cluster tools, when the train level (the average number of lots between setups) is three and there are 24 wafers per lot, the linear configuration with rolling setups has a 5.22% throughput advantage. We also conduct simulations to estimate the effect of wet cleans. Since wet cleans consume on average six hours every time a clean is required (on average every 750 hours), the effect of wet cleans is not as prominent. With both setups and wet cleans, the linear configuration has a 4.09% higher throughput. This advantage is more pronounced when the setups occur frequently and there are fewer wafers per lot. For 12 wafers per lot, the rolling setups of the linear configuration improve throughput by 10.03%. With both setups and wet cleans, the linear tool has a 9.15% higher throughput.

As a transition to 450 mm diameter wafers will provide an opportunity to rethink existing tool designs, our results suggest that the linear configuration be considered for future fabs. Similarly, existing fabs may benefit from the use of this relatively new architecture.

The paper is organized as follows. In Section 2, we provide a description of two different cluster tool architectures: circular cluster tools and linear cluster tools. Flow line models for simulation are discussed in Section 3. In Section 4, we conduct a simulation study to compare circular cluster tools and linear cluster tools. Concluding remarks are presented in Section 5.

## 2 DESCRIPTION OF TOOL ARCHITECTURES

Here, we review the features of circular and linear cluster tools.

### 2.1 Circular Cluster Tools

Circular cluster tools consist of a collection of process modules, input/output ports, load locks, and a wafer transport robot housed in a single chassis; refer to Figure 1. When wafers arrive as a lot inside a Front Opening Unified Pod (FOUP), the FOUP is placed in an input/output port. Wafers are transported from the FOUP into load locks. More than one load lock may exist in cluster tools to facilitate rapid transition from lot to lot. The circular cluster tools in Figure 1 possess two load locks each. A wafer transport robot transfers the wafers to and from the load locks and process modules. Wafers receive service from all required processes in the desired order. After all processes are complete, wafers return to their original slot in the FOUP. The completed FOUP is delivered to its next stage of production via the automated material handling system (AMHS).

While such tools are common in semiconductor wafer fabrication, their architecture imposes certain manufacturing inefficiencies due to two major limitations: vacuum isolation and a central wafer handling system.

Manufacturing processes in semiconductor fabrication are generally sensitive to atmospheric conditions and vacuum isolation is employed to prevent cross contamination. Thus, almost all activities inside a cluster tool should be conducted in vacuum isolation. However, since circular cluster tools share a central vacuum chamber, there are various drawbacks. First, the tools have limited configuration options. Certain process must be isolated from other processes to avoid cross contamination. For example, a single cluster tool can either provide CVD or PVD processes, not both. Second, it is difficult (if not practically impossible) to reconfigure a tool once it has been installed. Third, when changing from one class of lot to an-

other, a setup may be necessary that requires the entire tool to empty before the setup can begin. Fourth, some tool/chamber cleaning processes and failures may require an empty tool.

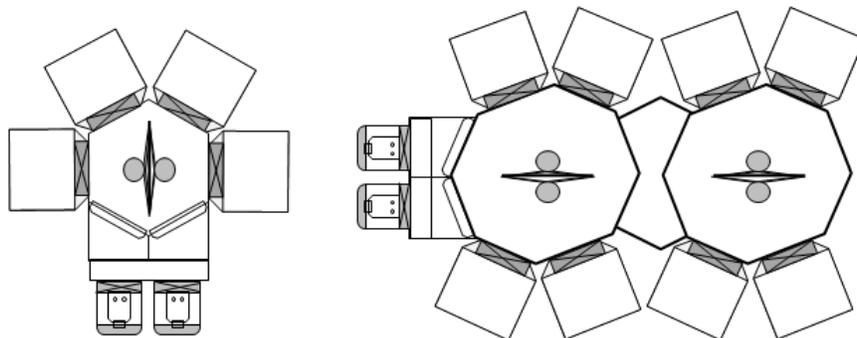


Figure 1: Circular cluster tools

The central wafer handling system also has drawbacks. First, due to the circular configuration, the tools have a large footprint (Meulen 2007). Second, since one robot is used for wafer transport, optimized robot control algorithms are required to ensure maximum throughput.

## 2.2 Linear Cluster Tools

In contrast, linear cluster tools are relatively new to semiconductor wafer fabrication; see Figure 2. They consist of a collection of paired process chambers (“links”), each link attended by its own wafer transport robot, connected in a linear flow; see Meulen (2007) and Yi et al. (2007). Figure 2 shows a 3-link linear cluster tool. The most prominent improvement of the linear configuration is that the links possess independent vacuum isolation.

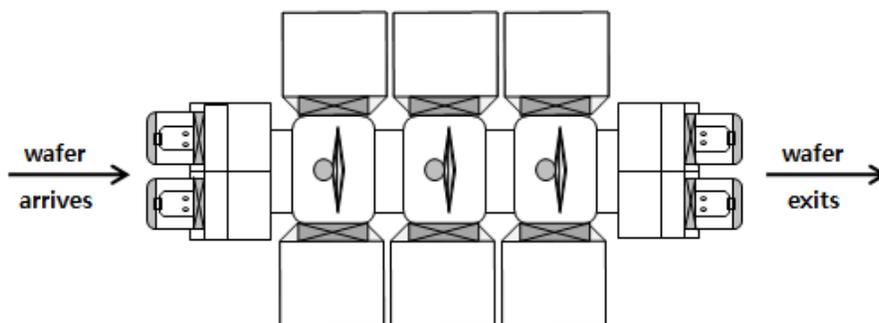


Figure 2: Linear cluster tools

Such a linkable, linear system has many advantages. First, the chambers can be readily mixed and matched. Since links provide independent vacuum isolation, the tools have less limitations on configuration. For example, a tool can provide both CVD and PVD processes. In addition, links can be readily connected or disconnected. Thus, the tools can be reconfigured even after it has been installed. Therefore, tools be reconfigured for a well-balanced process configuration, which may result in increased tool throughput. Second, a dynamic rolling setup can be conducted in linear tools. In linear tools, setups can proceed from link to link without the need to completely flush the tool. For example, a setup for link 1 can start when the link is empty even if a wafer is in service in link 2. The rolling setups allow cascading of lots and thus reduce first wafer delay (FWD). Consequently, it reduces fab cycle time and increases productivity; see Radloff et al. (2009). Third, chamber cleaning can be conducted independently in each chamber. When the particle level inside a cluster tools increases, process chambers may require cleaning

to avoid contamination. A wet cleaning process requires the process chamber to be empty. For circular cluster tools, the whole tool must be empty if one chamber needs to be cleaned. In contrast, linear cluster tools need not stop to clean one process chamber of parallel process chambers.

Linear cluster tools have other advantages as well: reduced floor space use due to rectangular footprint and simplified robot scheduling due to the separated input/output port and wafer transport robots in each link (Yi et al. 2007).

### 3 FLOW LINE MODELS

For tractability, we abstract the tools to flow line models that do not include wafer transport robots inside the tools. In Morrison (2011), flow line models have been claimed to estimate the throughput of production clustered photolithography tools to within 1% when not robot bound. With such flow line models, we can effectively model the operation of cluster tool features except wafer transport robot control. That is, we can study setups requiring an empty tool, rolling setups from link to link, cleaning processes and failures. In addition, our simulation study focuses on linear cluster tools with rolling setups and wet cleans; we are not concerned with the possible benefits derived from an increased number of robots and simpler control needs. Flow line models are reasonable for our objectives.

In general, it is known that if a cluster tool is process bound there exists a robot policy giving the bottleneck throughput in steady state. For example, the throughput time of a single-arm cluster tool with constant travel time is just the bottleneck process time plus pick, move and place time (Dawande et al. 2007) and the throughput of a dual-arm cluster tool with constant travel time is the bottleneck process time plus swap operation time (Park, Ahn, and Morrison 2011). If robot bound, the robot cycle time dictates the throughput. Since we assume process bound, adding robot behavior time to the bottleneck process time will ensure the same steady state throughput. In the transient state, the issue is not as simple as for steady state. However, we treat it as the steady state case.

A flow line model consists of a sequence of processes, from which each wafer must receive service in turn. Several equivalent process modules may be devoted to each process. One module can hold at most one wafer. When wafers arrive as a lot, they wait until the first process is ready in an infinite capacity queue. Wafers enter each process in a FCFS manner as soon as a module devoted to that process is available and the process has been setup to provide the required service (setup). When a wafer completes the last process, it exits the tool. Such a flow line model is depicted in Figure 3.

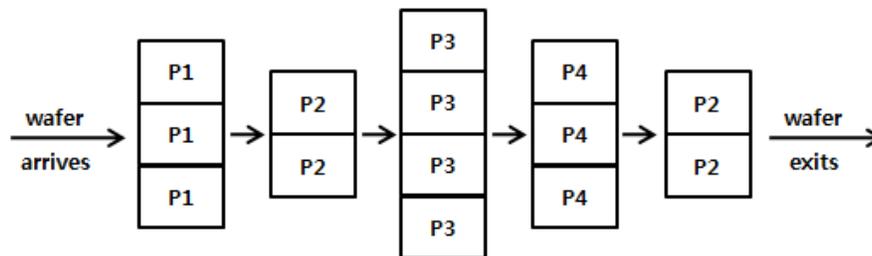


Figure 3: Flow line models

### 4 SIMULATION RESULTS

We now conduct numerical experiments for the two cluster tools architectures over an array of system parameters. For each desired setting of the system parameters, we simulate 6000 lots and use data from the last 5000 lots. Our warm-up period is thus 1000 lots. For all simulations, we assume that lots are always available at the tool entrance (so called, just in time JIT lot arrivals) and enter the tool in a FIFO order. We conduct 10 replications for each case. We simulate with MATLAB.

### 4.1 Systems

We conduct simulations of three different tools. The tool data is depicted in Table 1. Process times are given in seconds.

Table 1: System parameters

		Process	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
System 1	Morrison (2010)	Redundancy	2	3	1	2	1	1	1	3	2	1	3
		Process Time	92.00	126.00	33.00	65.00	47.00	100.00	29.00	141.00	50.00	23.00	96.00
System 2	Meulen (2007)	Redundancy	3	1	1	1	1	1					
		Process Time	300.00	120.00	60.00	60.00	60.00	60.00					
System 3	Geismar et al. (2006)	Redundancy	1	1	1	1	1	1	1	1	1	1	1
		Process Time	50.00	52.24	59.00	70.13	52.24	25.00	70.13	52.24	86.00	70.13	52.24

- System 1 is from Morrison (2010) and similar to real clustered photolithography tools. There are five pre-scan processes, one bottleneck scan process and five post-scan processes. In Table 1, “Redundancy” means the number of identical process models devoted to that process. For example, two modules are devoted to process P1 in System 1.
- System 2 is from Meulen (2007) at Blueshift Technologies. It is a multi-cluster tool with both CVD and PVD processes. Process P1 has three CVD modules and processes P2 to P6 have one PVD module for each process.
- System 3 is from Geismar, Dawande, and Sriskandarajah (2006). It is data from an FSI International, Inc. tool. There is no redundancy. That is, all the processes have only one module.

We conduct a simulation study with these three systems. For brevity, we only describe the results for System 2. Systems 1 and 3 provide similar behavior; their results are relegated to the appendix. Our focus is on rolling setups and wet cleans.

### 4.2 Effectiveness of Rolling Setups

We first compare the throughput of circular and linear cluster tools in the face of setups. Circular cluster tools (CT) must be empty before a setup can begin. Linear cluster tools (LT) only require that the links currently undergoing setup be vacant. With a transition to 450 mm wafer sizes and customization trends, it is essential to study small lot sizes and high mix production environment. Thus, we study various train levels and lot sizes. Our simulation conditions are described as follows.

- **Train level:** There are three classes of lots and a setup is needed when the class of lots changes. The train level is the average number of lots between setups. We consider three train levels ( $T = 1, 3, 5$ ). Thus, when the train level is equal to 1 ( $T = 1$ ), a setup is required for every lot.
- **Lot size:** We consider 7 different lot sizes with wafers per lot  $W = 6, 9, 12, 15, 18, 21, 24$ .
- **Setup duration:** Setup duration follows a uniform distribution with a range of 150 to 300 seconds (Meulen 2011). For circular cluster tools, setups are conducted for every chamber at the same time. For linear cluster tools, rolling setups are conducted from link to link. The setups for each link take independent random durations.

The simulation results for the throughput (THPT) of System 2 are depicted in Figure 4. The results of System 1 and System 3 are in Appendix A and B. The “no setup” case ( $T = +\infty$ ) gives the same throughput for both the circular tool and linear tool. This is because we neglect wafer handling robots (otherwise the linear tool would be marginally superior). Here, we examine the improvement by rolling setups. As shown in Figure 4, circular cluster tools are more sensitive to lot size and train level. For smaller train

level and lot size, the throughput of the tool decreases substantially. In contrast, the throughput of the linear tool changes significantly less than the circular tool. Since rolling setups allow cascading of different classes lots, there is a reduction in first wafer delay.

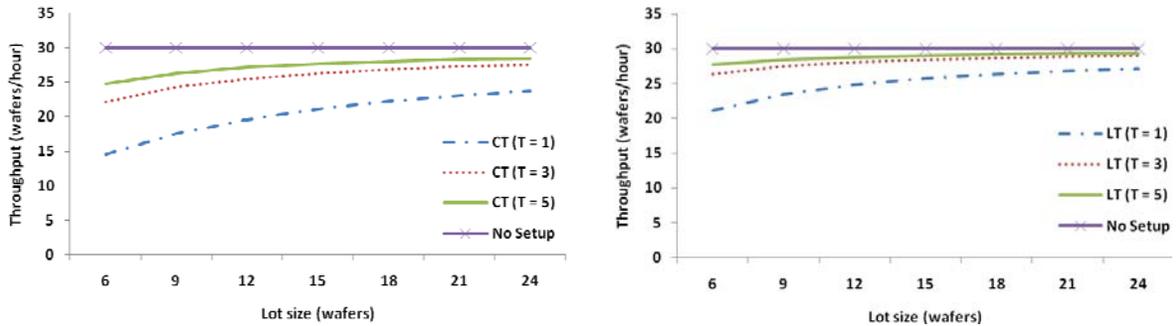


Figure 4: Throughput results for System 2

Table 2 and Figure 5 show the throughput improvement attributed to rolling setups. When lot size is 24 ( $W = 24$ ) and the train level is 3 ( $T = 3$ ), practical conditions for current manufacturing environments, rolling setups provide 5.22% increased throughput compared to that of standard circular cluster tool setup. For 12 wafer lots, the throughput difference is 10.03%.

Table 2: Rolling setups improve throughput in System 2

		W=6	W=9	W=12	W=15	W=18	W=21	W=24
T=1	THPT (CT)	14.55	17.56	19.59	21.05	22.15	23.01	23.71
	THPT (LT)	21.19	23.49	24.84	25.72	26.35	26.82	27.18
	Improvement	45.66%	33.75%	26.80%	22.17%	18.92%	16.52%	14.64%
T=3	THPT (CT)	22.15	24.29	25.50	26.26	26.80	27.25	27.56
	THPT (LT)	26.35	27.45	28.06	28.43	28.68	28.85	29.00
	Improvement	19.00%	13.01%	10.03%	8.25%	7.03%	5.88%	5.22%
T=5	THPT (CT)	24.76	26.26	27.15	27.66	28.03	28.30	28.50
	THPT (LT)	27.71	28.44	28.80	29.03	29.18	29.31	29.38
	Improvement	11.91%	8.29%	6.08%	4.94%	4.10%	3.58%	3.08%

### 4.3 Effectiveness of Wet Cleans

We next address wet cleans. For circular cluster tools, the entire tool must be stopped when one chamber needs to be cleaned. If a serial process chamber need to be cleaned, the tool engineer prohibits new wafers from entering the tool. Wafers inside the tool, excluding the wafers before the chamber needing to be cleaned, complete their remaining processes. If one of a collection of parallel process chambers are to be cleaned, the tool continues processing and completes the current lot. Then, the tool starts the wet clean process. For example, as process P1 in the linear architecture of System 2 is a parallel process, the tool can keep processing even during the wet clean process of one chamber of process P1. Simulation conditions are as follows.

- **System 2:** Lot size is 24 ( $W=24$ ). We only conduct simulations of System 2; the wet clean data was specifically obtained from Meulen (2011) which is our source for System 2. We have no clean data for Systems 1 and 3.

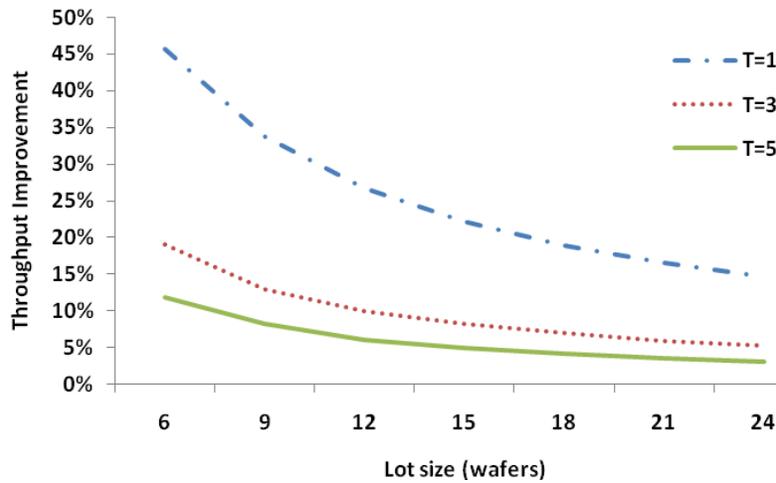


Figure 5: Rolling setups improve throughput in System 2

- **Time between wet cleans:** The time between wet cleans follows a uniform distribution from 500 to 1000 hours. In our simulation, each module has its own clock to indicate the time to next wet clean. For circular cluster tools, the clocks of all the modules reset when wet cleaning is conducted. On the other hand, only the clock for the module to be cleaned is reset in linear cluster tools.
- **Wet cleaning duration:** The wet clean duration follows a uniform distribution with a range of 4 to 8 hours.
- **No Setup:** We only focus on wet cleans. We do not consider setups here ( $T = \infty$ ).

The simulation results for System 2 are given in Table 3. The effect on throughput of wet cleans is relatively small since wet cleans occur very rarely though the duration is long compared to process times. The throughput difference between the circular tool and the linear tool is only 1.41%. However, contrary to our expectation, the throughput of the linear tool is worse than that of the circular tool. This is because the number of scrapped wafers is increased in the linear tool.

Table 3: Effect of wet cleaning

	THPT (wafers/hour)	Number of Scrapped Wafers	Scrapped Wafer Rate (wafers/lot)
No Failure	30	0	0
Circular	29.69	26.80	0.0054
Linear	29.28	90.40	0.0151

The number of scrapped wafers increases from one to four in 5000 wafers. Scrapped wafers are those wafers in the tool that cannot complete processing due to a wet clean. For example, consider that all the chambers are filled with wafers in System 2. When process 3 requires a wet clean, all the wafers in processes P1 to P3 are scrapped. Scrapped wafer require rework or are simply destroyed. The reason for the increase in scrapped wafers is due to the decreased actual time between wet cleans. For linear cluster tools, one wet cleaning process resets only the chamber cleaned. Therefore, wet cleaning processes occur more often. If the cost of wafers is expensive, then the tradeoff between throughput improvement and cost for scrapped wafers should be considered.

Though the effect on throughput of wet cleans is relatively small, there may be interaction with setups. Therefore we conduct a simulation study with both setups and wet cleans in Section 4.4.

#### 4.4 Effectiveness of Linear Cluster Tools

Simulation conditions are as follows.

- **System 2:** Train level is 3 ( $T=3$ ). Setup duration follows a uniform distribution with a range of 150 to 300 seconds.
- **Lot size:** We consider 7 different lot sizes with wafers per lot  $W = 6, 9, 12, 15, 18, 21, 24$ .
- **Wet Clean Time:** The time between wet cleans follows a uniform distribution with a range of 500 to 1000 hours. The wet clean duration follows a uniform distribution with a range of 4 to 8 hours.

Figure 6 shows the throughput results for circular and linear tools. As is expected from Sections 4.2 and 4.3, the linear tool provides significantly improved throughput. Interestingly, when the lot size is 24 ( $W=24$ ), the throughput improvement is 4.09%. This is smaller than the improvement when only rolling setups exist (5.22%). We conjecture that the frequent wet cleaning disturbs the cascading of lots and setups, resulting in an impediment to throughput improvement.

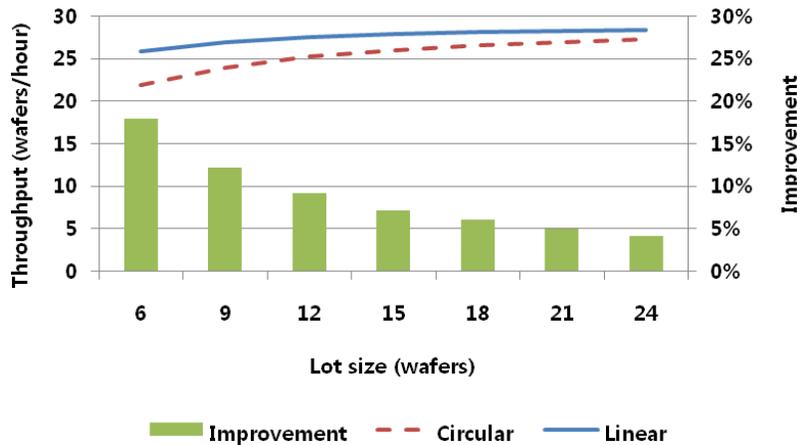


Figure 6: Throughput comparison between circular tools and linear tools

## 5 CONCLUDING REMARKS

Linear cluster tools enjoy advantages on account of their potential for reconfiguration and isolated links. We evaluated the effectiveness of linear cluster tools via simulation of their flow line approximations. In this paper, we focused on rolling setups and wet cleans. The effect of the linear configuration in the face of wet cleans is relatively low. This is due to the infrequent occurrence of wet cleans. However, the linear configuration had significantly higher throughput when faced with setups. As rolling setups allow cascading of lots and reduction of FWD, the linear tools have a pronounced strength when train level increases and/or lot size become smaller. The advantages of such tools, including resilience to setups and wet cleans, ease of reconfiguration, simplified robot scheduling and reduced foot print, argue for their careful consideration as next generation equipment.

## ACKNOWLEDGMENTS

This work was partially supported by National Research Foundation of Korea (NRF) Grant N01100263.

**A APPENDIX: THROUGHPUT RESULTS FOR SYSTEM 1 AND SYSTEM 3**

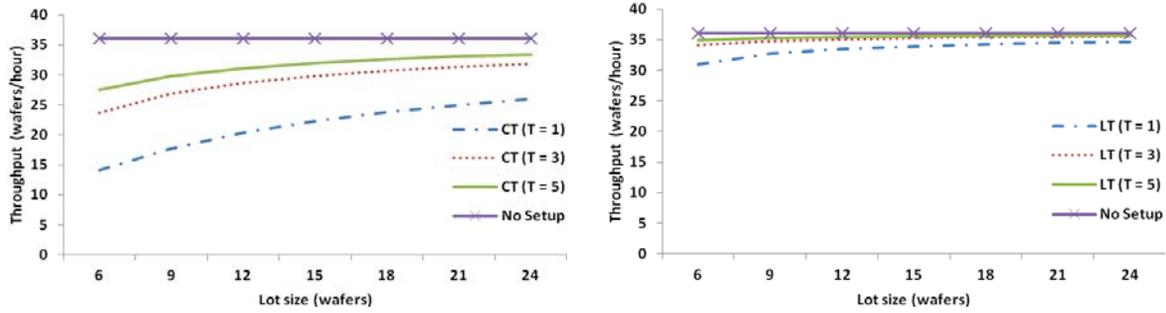


Figure A-1: Throughput results for System 1

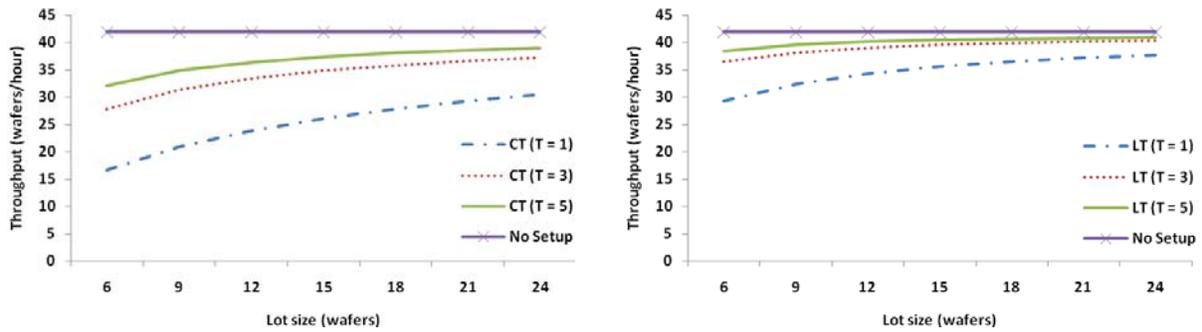


Figure A-2: Throughput results for System 3

**B APPENDIX: IMPROVEMENT BY ROLLING SETUPS IN SYSTEM 1 AND SYSTEM 3**

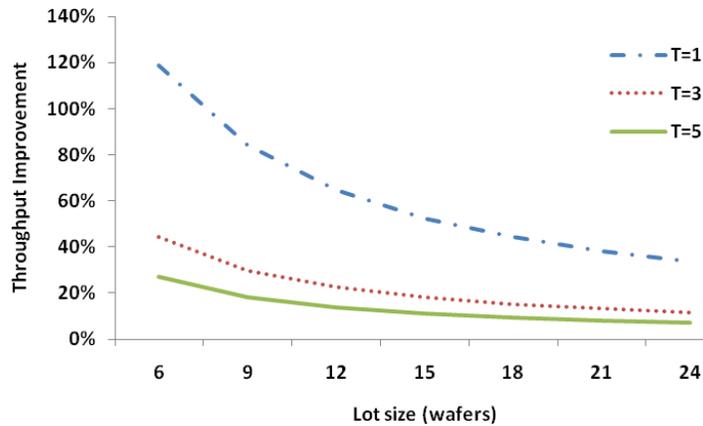


Figure B-1: Rolling setups improve throughput in System 1

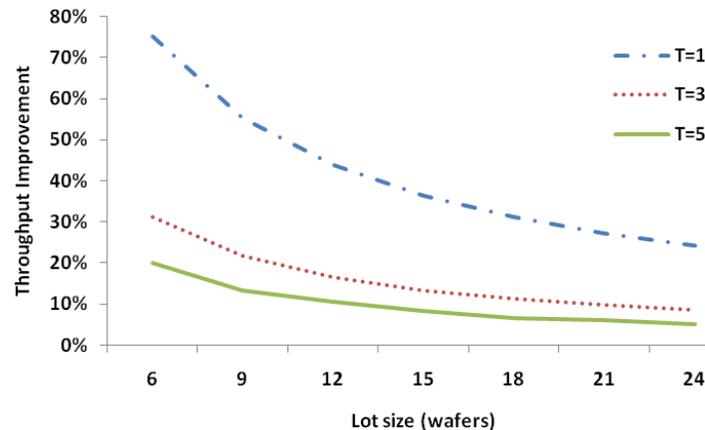


Figure B-2: Rolling setups improve throughput in System 3

## REFERENCES

- Dawande, M., H. N. Geismar, S. Shethi, and C. Sriskandarajah. 2007. *Throughput Optimization in Robotic Cells*. NY: Springer.
- Geismar, H. N., M. Dawande and C. Sriskandarajah. 2006. "Throughput Optimization in Constant Travel-Time Dual gripper Robotic Cells with Parallel Machines." *Production and Operations Management* 15(2):311-328.
- Morrison, J. R. 2010. "Deterministic Flow Lines with Applications." *IEEE Transactions on Automation Science and Engineering* 7(2):228-239.
- Morrison, J. R. 2011. "Multiclass Flow Line Models of Semiconductor Manufacturing Equipment for Fab-Level Simulation." *IEEE Transactions on Automation Science and Engineering* 8(1):81-94.
- Meulen, P. van der. 2007. "Linear Semiconductor Manufacturing Logistics and the Impact on Cycle Time." In *Proceeding of the 18th IEEE/SEMI Advanced Semiconductor Manufacturing Conference*, 111-116. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Meulen, P. van der. 2011. Private communication.
- Park, K., Y. Ahn, and J.R. Morrison. 2011. "Performance Models for Dual-arm Cluster tools." In *Proceeding of the 2011 World Congress on Intelligent Control and Automation*, 816-821. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Pillai, D. 2006. "The Future of Semiconductor Manufacturing: Factory Integration Breakthrough Opportunities." *IEEE Robotics & Automation Magazine* 13(4):16-24.
- Radloff, S., M. Abravanel, B. Rhoads, D. Steeg, P. van der Meulen and M. Petraitis. 2009. "First Wafer Delay and Setup: How to Measure, Define and Improve First Wafer Delays and Setup Times in Semiconductor Fabs." In *Proceeding of the 20th IEEE/SEMI Advanced Semiconductor Manufacturing Conference*, 86-90. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Yi, J., S. Ding, M.T. Zhang, and P. van der Meulen. 2007. "Throughput Analysis of Linear Cluster Tools." In *Proceeding of the 3rd IEEE Conference on Automation Science and Engineering*, 1063-1068. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

## AUTHOR BIOGRAPHIES

**KYUNGSU PARK** is currently a Ph.D. student in the Department of Industrial and Systems Engineering at KAIST, Republic of Korea. He received his bachelor's degrees in Information and Industrial Engineer-

ing and Chemical Engineering from Yonsei University, Republic of Korea. His research interest is on modeling and scheduling in semiconductor manufacturing. His email is [ks-park@kaist.ac.kr](mailto:ks-park@kaist.ac.kr).

**JAMES R. MORRISON** is an Associate Professor in the Department of Industrial and Systems Engineering at KAIST, Republic of Korea. He received his Ph.D. in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign. Beginning in September 2008, he served as a co-chair for the IEEE Robotics and Automation Society Technical Committee on Semiconductor Manufacturing Automation. He is active in IEEE. From 2000 to 2005, he was with the IBM Fab Operations Engineering department at IBM's Burlington Vermont semiconductor wafer fab. His email and webpage are [james.morrison@kaist.edu](mailto:james.morrison@kaist.edu) and <http://xs3d.kaist.edu> respectively.