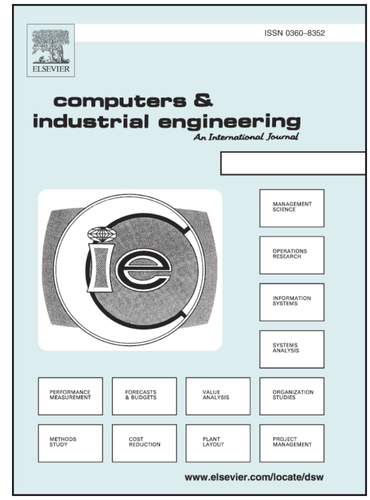
Accepted Manuscript



Semiconductor FAB Layout Design Analysis with 300-mm FAB Data: "Is minimum distance-based layout design best for semiconductor FAB design?"

Junghoon Kim, Gwangjae Yu, Young Jae Jang

PII: S0360-8352(16)30039-0
DOI: http://dx.doi.org/10.1016/j.cie.2016.02.012
Reference: CAIE 4259

To appear in: *Computers & Industrial Engineering*

Received Date: 8 April 2015
Revised Date: 16 October 2015
Accepted Date: 15 February 2016

Paper Title: Semiconductor FAB Layout Design Analysis with 300-mm FAB Data: "Is minimum distance-based layout design best for semiconductor FAB design?"

Junghoon Kim[*,1], Gwangjae Yu[*,2], and **Young Jae Jang[*,†,3]**

[*] *Department of the Industrial and Systems Engineering*
Korea Advanced Institute of Science and Technology, Daejeon, 305-701, KOREA
Phone: +82-42-350-3170,

[†] (Corresponding author) Department of Industrial & Systems Engineering
Korea Advanced Institute of Science and Technology

291 Daehak-ro, Daejeon, 305-701, KOREA
Phone: +82-42-350-3130, Fax: +83-42-350-3110, E-mail: yjang@kaist.ac.kr

[1] Junghoon Kim received his M.S. degree in Industrial and Systems Engineering from the Korea Advanced Institute of Science and Technology (KAIST) in 2011 and a B.S. degree in Industrial Engineering also from KAIST in 2008. He is currently a PhD candidate in the Industrial Engineering and Management Research Institute at KAIST. His research interests include semiconductor layout design, semiconductor operation analysis, and stochastic system modeling.

[2] Gwangjae Yu received his M.S. degree in Industrial and Systems Engineering from the Korea Advanced Institute of Science and Technology (KAIST) in 2013 and a B.S. degree in Industrial Engineering also from KAIST in 2011. He is currently a Graduate Research Assistant in the Industrial Engineering and Management Research Institute at KAIST. His research interests include semiconductor layout design, complex systems modeling, and optimization algorithm design.

[3] Young Jae Jang received his Ph.D. degree in mechanical engineering from Massachusetts Institute of Technology (MIT) in 2007 and a double M.S. degree in mechanical engineering and operations research from MIT in 2001. He received a B.S degree in aerospace engineering from Boston University in 1997. He is currently an Assistant Professor in the Industrial and Systems Engineering Department at the Korea Advanced Institute of Science and Technology (KAIST), South Korea. He is also affiliated with the Cho Chun Shik Graduate School of Green Transportation at KAIST. His current research includes stochastic modeling of complex systems and optimizations in transportation and logistics systems. Before he joined KAIST, he worked at Micron Technology, Inc., VA, USA, where as project manager, he led a global initiative to improve the efficiency of the manufacturing facilities located worldwide. While working at Micron, he was also involved in activities supporting the company's operational and strategic decision making using quantitative modeling and analysis techniques.

Semiconductor FAB Layout Design Analysis with 300-mm FAB Data: "Is minimum distance-based layout design best for semiconductor FAB design?"

Abstract

In this paper, one of the most common layout design approaches in general factory design – minimum flow-weighted distance (min-distance) layout – is verified for semiconductor fabrication facility (FAB) design. In min-distance layout design, machines or workstations with high flow rates are allocated close to each other to minimize the total distance of the material flow. This approach is widely used in various industries including semiconductor manufacturing. However, some research also discusses the drawbacks of this approach due to flow congestions and heavy traffic in material handling systems. We validate the min-distance approach with actual data from a modern 300-mm DRAM FAB. We logically generate 18 different cases with different layouts, overhead hoist transport track configurations, and production scenarios. The effectiveness and drawbacks of the min-distance approach in FAB design is investigated with these cases. From the simulation analysis using the Design of Experiments method, we logically show that the performance of material delivery is sensitive to the production volume in the min-distance layout. Also, the performance of the min-distance layout is significantly degraded when the volume is heavy; however, performance can be improved considerably with a few modifications to the bays. We also provide practical tips for an effective layout design method from the insight gained from the simulation analysis. This paper contributes to the critical analysis of the conventional layout design method and the identification of its effectiveness and limitations by using actual FAB data.

Index Terms

Semiconductor FAB design, AMHS, factory layout design, and material flow design

I. Introduction

Semiconductor wafer fabrication processing is known to be one of the most complex of manufacturing systems. A memory chip, which is the typical product processed in the wafer fabrication process, requires hundreds of processing steps. Multiple products are processed simultaneously in a semiconductor fabrication facility, also known as a FAB. Each step requires specific processing machines, which are commonly called *tools* in the semiconductor industry. Therefore, hundreds of different types of tools and large flows between the tools are required to process wafers.

The transportation of these wafers is accomplished by the Automated Material Handling System (AMHS), which is an automated system that moves, stores, and controls the parts and inventory throughout the manufacturing process. In a semiconductor FAB, the AMHS consists of transportation units such as overhead hoist transport (OHT) vehicles, tracks, and storage systems such as overhead bins and stockers. With the increasing scale of wafer production, there are generally more than a hundred OHT vehicles to transport wafers in the typical memory chip FAB environment.

The performance measures of the AMHS includes delivery time, throughput, and utilization of vehicles[1]. The key design parameters in the AMHS are track configurations, OHT vehicle parameters such as velocity, loading and unloading time, controller logic and algorithms. Since the primary role of the AMHS is to transport a lot from one tool to another, allocation of the tools is directly related to AMHS performance. Therefore, it is common in the semiconductor industry that the tool allocation process and the AMHS design process are performed simultaneously during the FAB design stage[2].

Tool allocation and AMHS design are critical steps in the FAB design process. Once they are decided, modification is difficult. Also, an inefficient design may result in serious performance loss and additional costs. For example, it is widely known in the industry that in the mid-2000s one of a leading DRAM manufacturers (let us name it “FAB-K”) had to dismantle the existing AMHS in one of its FABs and reinstall a new one while ramping up to target production volume. The original AMHS, which had been designed during the initial FAB design, did not have enough capacity to move the lots, and it became the bottleneck as production volume increased. Eventually, the engineers concluded that with the original AMHS design, the target production volume could not be achieved, and they decided to replace the AMHS with a completely new system. The replacement took several months. The direct replacement cost and the indirect cost from loss of production were enormous[2].

In the FAB-K case, a primary cause of the underperformance of the original AMHS was OHT vehicle congestion in a concentrated area within the FAB. Although this traffic congestion occurred in a limited area within the FAB, the congestion resulted in the delay of vehicle movements throughout the entire FAB when the production volume was increased. The engineers determined that the primary cause of the congestion was tool allocations. They allocated the *bays*, which refers to the sections in the FAB containing similar types of tools, with a high inter-bay flow close to one another to achieve fast lot delivery. Unfortunately, the heavy flow volume between the bays resulted in substantial vehicle congestion and eventually created a bottleneck for all vehicle traffic in the FAB.

The research described in this paper is motivated by this incident at FAB-K. The method of allocating the bays with high inter-bay flow close to each other is called *minimum-distance layout design* or, simply, *min-distance* layout. The criteria of the min-distance approach are widely used in the allocation of tools and the design of the AMHS in semiconductor FAB design[3][4][5]. The survey in [2] also indicates that majority of layout designers in major semiconductor companies in Korea use the min-distance as the primary criteria in layout design. However, some research has addressed the weaknesses of the min-distance approach in the general context of manufacturing[6][7][8].

We try to answer the following question with this study: *"Is min-distance layout design always best in semiconductor FAB design?"*

The primary weakness of the min-distance layout approach is the congestion of material flow. This weakness is somewhat clear and intuitively understandable. However, there has been little research to rigorously verify the limitations of this approach in the area of the semiconductor FAB design. Given the fact that there are hundreds of OHT vehicles moving around on the uni-directional OHT tracks in a FAB, heavy flow in a concentrated area may cause heavy traffic and delivery delay, eventually creating bottlenecks in the AMHS.

In this paper, we examine the effects of material flow on the OHT track with various flow scenarios using data of the actual processing steps from a 300-mm DRAM FAB. We analyze the effects of vehicle congestion on the performance of the AMHS using simulation. We try to answer the following additional questions:

- What are the weaknesses and limitations of the min-distance approach in FAB design?
- How do various track configurations and traffic volume influence traffic congestion on the OHT track?

To logically understand the effectiveness of the min-distance layout, we generate a *max-distance* layout, which is the opposite extreme of the min-distance approach, and compare the behaviors of the two layouts. We hypothesize that if the performance of the max-distance case is better than that of the min-distance case under certain conditions, we can conclude that the min-distance approach may not be the best approach for the given conditions.

Eighteen different configuration/layout cases are generated for the study: min-distance and max-distance layouts, with each layout having the combination of three different flow-volume scenarios and another three different OHT-track configurations, as shown in Figure 1. The simulation is used to analyze the cases. With the results generated from the simulation, we analyze the performance of the AMHS, particularly the delivery time, using the Design Of Experiments (DOE) method. From the insights gained from the simulation study, we provide practical implications and tips for layout design process.

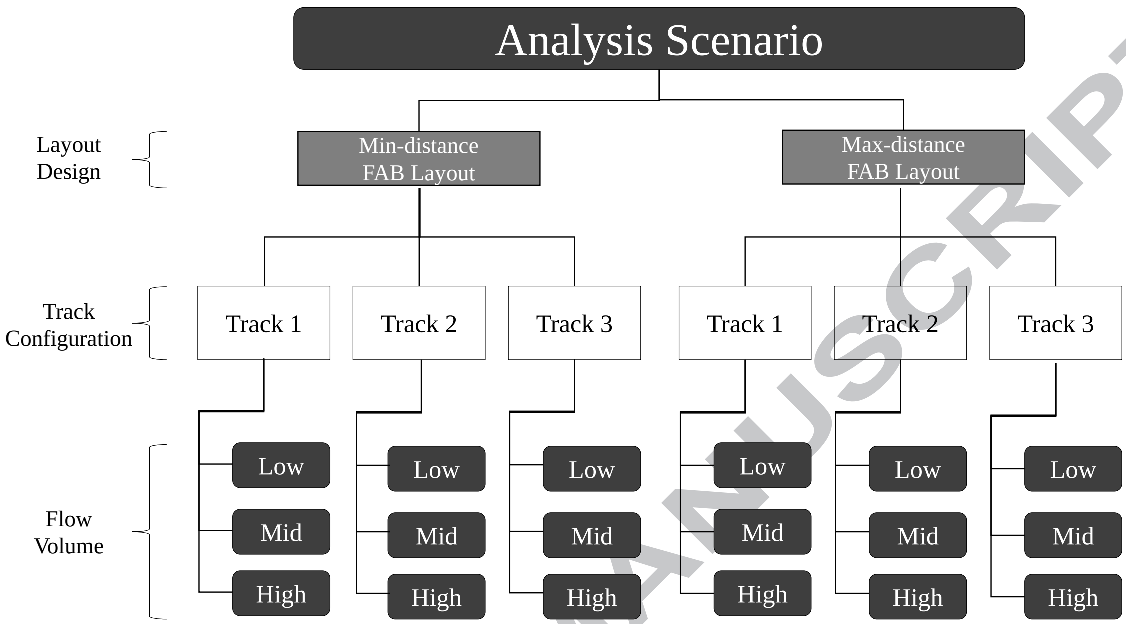


Fig. 1. **Analysis scenario of the three factors: Layout design, Track configuration, and Flow volume**

The paper is organized as follows. Section II summarizes the previous research on layout design methodologies and congestion issues. Section III defines the problem. Then, we generate FAB layouts and input data for the simulation experiments. In Section IV, DOE is performed to analyze the simulation results of the AMHS for each layout. The simulation results are analyzed and insight is provided. **In Section V, we suggest tips for layout design from the insights gained from the analysis.** In Section VI, we summarize this study and make suggestions for future research.

II. Literature Review

Before considering the methodology for designing the FAB-specific layout, we examine first the approaches taken for general factory layout design. Recent surveys on this design are found in Singh and Sharma[9] and Drira et al.[10]. There are several methods to mathematically formulate the layout problems. The Quadratic Assignment Problem (QAP) and Mixed Integer Programming (MIP) are the most common modeling approaches to analytical layout design. In the case of modeling with QAP and variations of QAP, the layout is divided into rectangular blocks of equal area and shape, and then N machine groups are assigned to the M blocks[11][12][13]. In the MIP modeling approach, machines are assigned to the floor considering the centroid or the location of input/output coordinates of the tools[14][15]. Most QAP and MIP modeling approaches determine the location of facilities (or machines) to minimize the total material handling cost, which is a function of the flow-weighted distance[10][11][12]. Bukchin et al.[14] and Kim and Kim [15] also minimize the flow-weighted distance between input/output points of machines. With minimization of the flow-weighted distance, tools with high flow interaction are clustered together. In contrast to analytical modeling, there are also rule-based approaches in factory layout. Systematic Layout Planning (SLP), one of these rule-based approaches, places the relative location of tools that have heavy inter-flow interactions close to each other[16]. Tompkins et al.[17] studied tool assignment methods based on the material flow intensities and closeness ratings.

FAB has unique characteristics compared to those of other manufacturing systems. Its characteristics include a facility handling a large number of processing steps, re-entrance flows, and a bay-type structure[18]. Some approaches consider these FAB-specific characteristics. Meller[3] and Ignacio and Peters[4] proposed two steps of a mathematical model for designing the FAB layout to deal with bay structures. They consider assigning tools or departments to bays at first and then decide on the configuration of the bay. Chen et al.[18] suggested a procedure for FAB design with the following tasks: machine grouping, bay allocation, and inter-bay and intra-bay movement analysis. There are several approaches that consolidate the FAB layout and AMHS design. Peters and Yang[19], Yang et al.[20], and Ho and Liao[5] proposed methods to allocate tools and bays and decide the location of shortcuts in the AMHS simultaneously. Yang et al.[21] applied SLP to decide the relative location of the machine groups in the FAB.

Although special consideration must be given to FAB layout design, most of these methods try to minimize the flow-weighted distance as does general factory layout design. In Meller[3]and Ignacio and Peters[4], the objective function is to minimize the flow-weighted distance between tools in bays. Chen et al.[18] assigned tools to machine groups that comprise a large inter-flow intensity among the groups. Peters and Yang[19] used the shortest path algorithm to design shortcuts on the OHT track to minimize travel cost. In Yang et al.[20], the objective function is multi-objective terms, which are used to minimize the material handling cost and the number of turntables in an AMHS. The first term of the objective function also considers the flow-weighted distance to minimize material handling costs. Ting and Tanchoco[22] designed the main loop of a unidirectional circular loop layout for an AMHS to minimize the total flow-weighted distance between stockers. In Ho and Liao[5], the objective function is to minimize the total inter-bay flow distance. The method in Yang et al.[21] was SLP that considered flow-intensity between machine groups.

In this paper, we discuss congestion issues related to the FAB layout. Although little research has dealt with congestion in the FAB layout, several studies have addressed congestion in the context of general factory layout design. Some researchers mentioned that the distance-based layout has poor performance due to congestion caused by blocking, traffic jams, and vehicle interference. Benjaafar[6] found that the layout designed by the traditional distance-based QAP model can be infeasible due to accumulation of work in progress caused by flow asymmetry, dimensional asymmetry, and the number of transporters. Zhang et al.[7][8] demonstrated that the shortest path is efficient in the low-traffic situation but causes congestion in the moderately high-traffic condition. They explained that it may be better to use the longer path as an alternative route to mitigate congestion. Moreover, Cheng[23] showed that the Largest Rectilinear Distance (LRD) dispatching rule can alleviate congestion. Zhang et al.[7] and Chiang et al.[24] also mentioned that congestion usually occurs at intersections and pick-up and drop-off points (P/D points) on the layout due to crossing and overlaying of flows.

There is some research on the alleviation of congestion in the general factory layout design. Benjaafar[6] proposed a layout design methodology to minimize work in progress rather than distance, and Chiang[24] developed a layout design to reduce the workflow interferences caused by crossing at an intersection or overlaying of flows. Zhang et al.[7][8] considered an optimized model for traffic balance on the layout to minimize the total expected travel time.

In summary, most previous research on FAB layout and AMHS has focused on minimization of the flow-weighted distance between tools. **Most research dealing with congestion has focused on deadlock detection and vehicle control[25][26][27], but the effects of the configuration of the FAB layout and AMHS design on congestion were not considered.** Moreover, there has been little research to verify the min-distance FAB layout using actual FAB data to assess congestion in the FAB. With this research, we verify the effectiveness and limitations of the min-distance FAB layout through simulation experiments.

III. Problem Definition

This section describes the type of FAB and AMHS on which this research focuses. Also, scenarios for the simulation analysis are presented.

A. Description of semiconductor FAB layout and AMHS

A typical configuration used in a semiconductor FAB is the spine type[4][18][19][20][28]. With a spine configuration of the AMHS, a FAB is divided into several bays that are arranged on both side of a central aisle. One bay usually contains similar types of processing tools or a group of tool sets[28]. In a FAB layout, tools performing similar processes are generally clustered together to support utility configuration and maintenance. Therefore, the bays with similar processing tools are generally allocated adjacent to each other. The bays are connected by the tracks and stockers as shown in Figure 2. Two types of AMHS use the spine configuration: the inter-bay system and the intra-bay system[29]. A stocker is installed in each bay to serve as a storage system, and it also works as a connection point for movement of lots between the inter-bay and intra-bay system. The OHT track of the inter-bay system on the central aisle transports the lots between stockers located in front of each bay. The intra-bay material handling system transports the lots within the bay. Since similar processing sequences are repeated during all of the steps involved in semiconductor FAB processing, there are many re-entrances of flow of the wafers to each bay. In the inter-/intra-bay type of FAB, the performance of the inter-bay system is highly important to the overall material flow[28]. Much research on the FAB layout has focused on the inter-bay system[3][4]. In this study, we also consider the inter-bay system only to observe the flow behavior of lots between the bays.

In Figure 2, the main loop is a single monorail track that connects all bays in the FAB. We also consider an additional loop within the main loop in this study. Many practical FABs use custom track configurations with shortcuts to enhance their performance[29]. Shortcuts on the track allow OHT vehicles to travel shorter distances. A lot is loaded or unloaded at the P/D point at a stocker.

B. Flow and volumes

In this research, we perform a simulation analysis using actual data from a 300-mm DRAM memory chip FAB. The data include information on the processing steps for each product and corresponding tools for each step. There are four product groups in the data and the product mix ratio is 4.5:3.5:1:1. The first two products are major products, and the last two are minor products. **The number of steps for each product are 289, 302, 249, and 322, respectively. Note that although the rest of the analysis is based on this particular product mix scenario, we show later that the results from the analyses are less sensitive to the variation in product mix. We discuss this issue in section V. Practical Implication and Suggestions.** In our analysis, we consider multiple production volume scenarios. With the product mix ratio, three different volume cases are analyzed: 10,000 wafers/week,

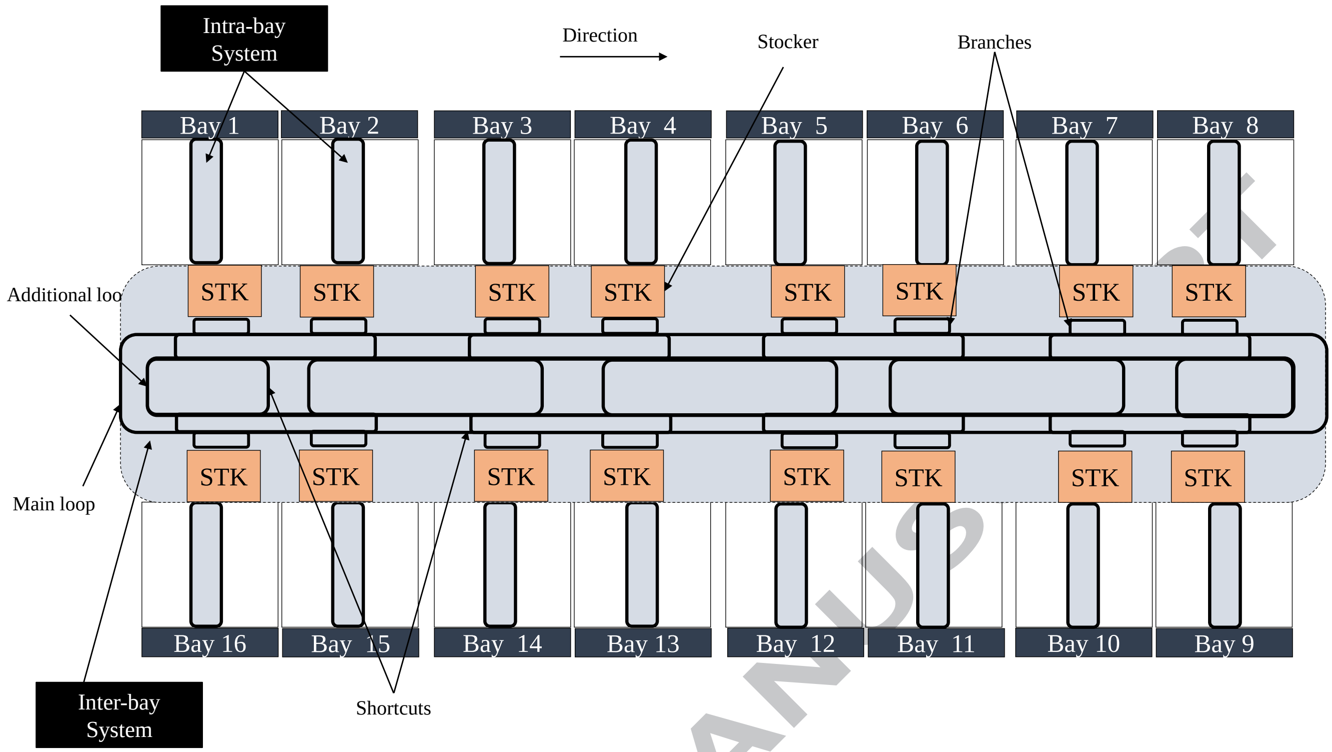


Fig. 2. The inter- and intra-bay FAB configuration

20,000 wafers/week, and 30,000 wafers/week. This volume scenario is similar to the production ramp-up scenario for the FAB as shown in Figure 3. When a new FAB starts operation, the production volume is gradually increased to the target production level. In our volume scenario, the target volume is 30,000 wafers/week, and there are two milestones, 10,000 wafers/week and 20,000 wafers/week. The number of tools and vehicles are gradually increased to meet the demand at each milestone. That is, more tools are added to the bays to meet the increased production volume. Also, as the tools are added, the intra-bay systems are expanded.

C. Tool groups

Our analysis includes eight groups of semiconductor wafer fabrication processes: chemical mechanical planarization (CMP), chemical vapor decomposition (CVD), diffusion (DIF), dry etching (DRY), implant (IMP), physical vapor decomposition (PVD), photo lithography (PHOTO), and wet etching (WET). We do not consider the metrology process in this study. Each process group has a number of tool groups that process similar steps. For example, the PVD process group includes tool groups such as PVD_END1, PVD_AMAT2, PVD_AMAT3, and others. Tools in each tool group are assumed to have the same

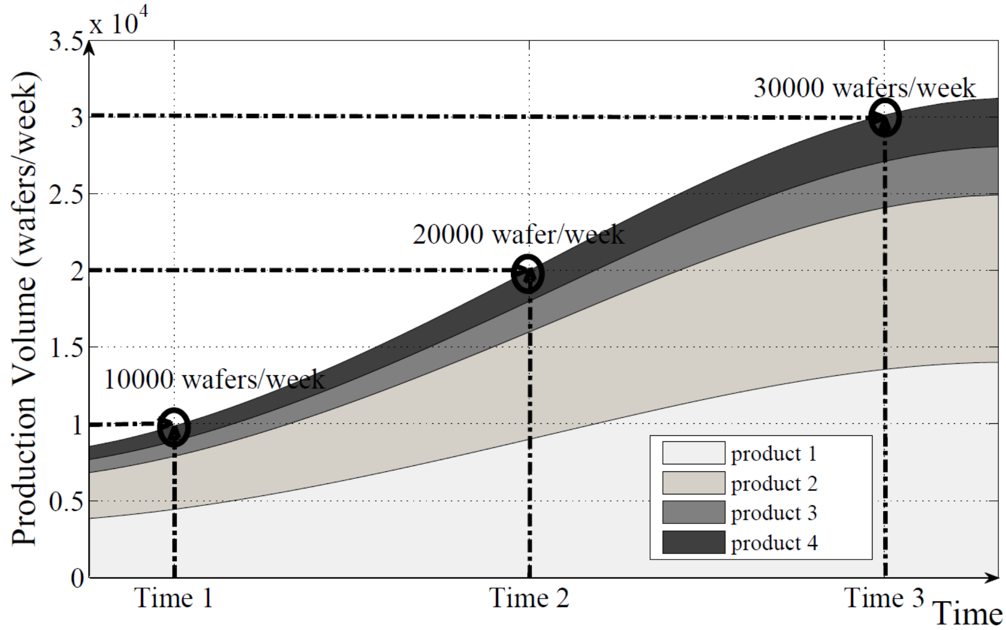


Fig. 3. Production ramp-up scenario

processing time, utilization, processing lot quantity, and other tool characteristics. The number of tools for each tool group is determined for each production volume scenario using the optimization model presented in the **Appendix I.A.** Table I shows the number of tool groups and tool numbers for each process group for three different volume scenarios.

TABLE I

TOOL COUNT

process group	Number of tool groups in each process group	Total number of tools for 10000 wafers/week	Total number of tools for 20000 wafers/week	Total number of tools for 30000 wafers/week
CMP	9	28	56	84
CVD	19	77	154	231
DIF	17	70	140	210
DRY	19	135	270	405
IMP	11	23	46	69
PVD	6	10	20	30
PHOTO	6	18	36	54
WET	18	79	158	237

Tables II presents the hourly flow-interaction data among the process groups for 10000 wafers/week. The data for 20000 and 30000 wafers/week can be obtained by multiplying by two and three times. These flow rates are used to create the bay layouts.

TABLE II

HOURLY FLOW-INTERACTION DATA BETWEEN PROCESS GROUPS (WAFERS/HOUR)

From-To	CMP	CVD	DIF	DRY	IMP	PVD	PHO	WET
CMP	0	220.24	101.19	0	0	0	160.71	**244.049**
CVD	125	752.98	220.24	101.19	71.43	0	395.83	**220.24**
DIF	333.33	300.6	113.1	95.24	151.79	0	184.52	**854.175**
DRY	0	5.95	17.86	0	0	0	142.86	**1773.81**
IMP	5.95	11.9	125	53.57	494.05	0	59.52	**732.14**
PVD	0	5.95	53.57	0	0	53.57	5.95	**252.9762**
PHO	53.57	0	53.57	1113.1	452.38	0	71.43	**178.57**
WET	208.33	589.29	1348.21	577.38	312.5	318.45	901.79	**2633.929**

D. Layouts

For the min-distance FAB layout, the bays with high flow interactions are located next to each other. As discussed earlier, to logically analyze congestion, we also consider the max-distance FAB layout. which is the counter case to the min-distance layout. Using the optimal bay allocation method presented in the **Appendix I.B**, the two layouts are generated as shown in Figure 4. This optimization allocates the tool groups to bays by considering the tool footprints and lot moves between and within tools. For the min-distance layout, the optimization model allocates the tools to bays to accomplish the total flow distance, whereas the max-distance layout maximizes the total flow distance. Details of the optimization model are presented in the **Appendix I.C**. The first layout is the result generated from the min-distance layout, and the second layout is that generated from the max-distance layout. **The flow-weighted distances for the layouts are 25902.82 and 48504.12 for the min-distance layout and max-distance layout, respectively.** Note that the flow-weighted distance for the min-distance layout is shorter than that of the max-distance layout.

According to the flow-interaction data between the process groups in **Tables II,** the process groups WET and DIF have high flow interactions. Therefore, in the min-distance layout, DIF is adjacent to WET. However, as shown in Figure 4, WET and DIF are far from each other in the max-distance layout, as expected. In this study, we identify the differences between the two layouts by comparing delivery times of the OHT vehicles(hereafter, vehicles).

E. Track configurations

Many practical FABs use a customized track configuration to provide optimal paths for the vehicles[29]. The shortcut refers to an inner loop linked to the main loop to provide a shorter travel path. Two or more vehicles compete to cross an intersection when vehicles meet each other there. If flows are concentrated on or many vehicles pass an intersection, a large number of vehicles cluster at the intersection. In this

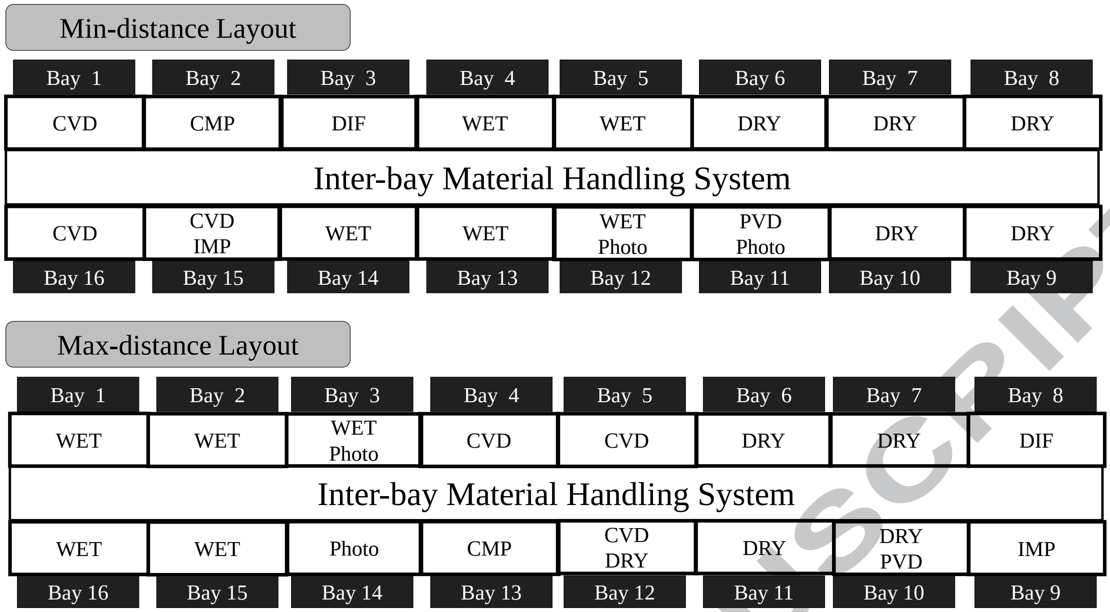


Fig. 4. Two FAB layouts, the min- and max-distance layouts

study, we investigate how the number of shortcuts affects vehicle flows. In this analysis, we consider three types of track configurations for each FAB layout as illustrated in Figure 5. Each configuration is distinguished by the number of shortcuts. The first type(track type 1) has the highest number of shortcuts, whereas the third type (track type 3) has the least number of shortcuts. The number of shortcuts in the second type (track type 2) lies somewhere between types 1 and 3. In this analysis, we try to understand how the number of shortcuts affects material flows.

As discussed earlier, the numbers of tools and vehicles are different depending on the production volume scenario. That is, during the ramp-up period, tools are added, and the intra-bay system is also expanded. However, the OHT track in the inter-bay system is normally installed to support the target production level even as the FAB starts its operation because the inter-bay OHT track cannot be easily expanded once a FAB starts operation. That is, the OHT track has enough capacity to handle 30,000 wafers/week, even if the experiments here only assess production of 10,000 or 20,000 wafers/week. However, the number of vehicles may vary depending on the volume scenario. Table III summarizes the results of the estimation of vehicles for each scenario.

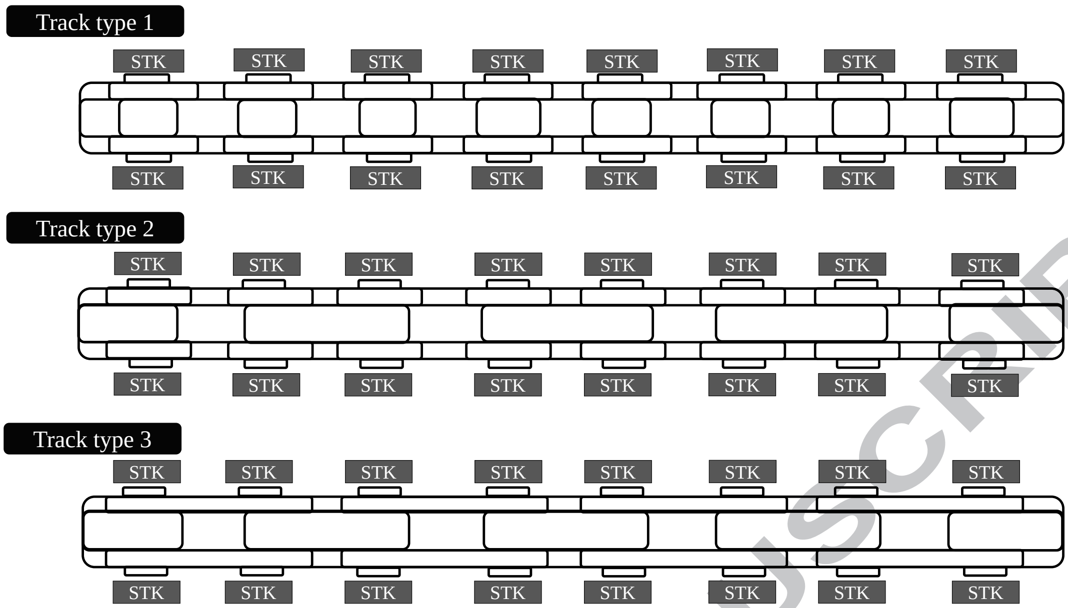


Fig. 5. The three track configurations

TABLE III

THE REQUIRED NUMBER OF VEHICLES FOR EACH SCENARIO

	Track configuration	Track type 1		Track type 2		Track type 3	
	FAB layout design	Min-distance layout	Max-distance layout	Min-distance layout	Max-distance layout	Min-distance layout	Max-distance layout
Flow Volume	10000 wafers/week	15	19	16	18	15	17
	20000 wafers/week	29	32	30	33	31	31
	30000 wafers/week	52	57	51	58	53	57

IV. VERIFICATION OF FAB LAYOUT THROUGH DOE

A. DOE settings

We analyze the simulation results of the AMHS performance of the two layouts with different scenarios through DOE. DOE is a statistical method used to examine the effect of more than two input factors on output results. DOE provides statistical reliability through the proper interpretation of the results of simulation. There are several types of DOE, including the factorial design, Taguchi method, and the Response Surface Method (RSM). The factorial design analyzes the factors, which are k-fixed input parameters, to examine the interaction between them using a response variable, which is the results of an experiment. We use factorial design because the objective in this study is to analyze factors or interactions between them that we are interested in.

Factors and levels for the factorial design are summarized in Table IV. We consider three factors and each factor has its own levels. The levels of the first factor 'A' are the min- and max-distance FAB layouts. The second factor 'B' has three levels: track type 1, 2, and 3, as described in Figure 5. The third factor 'C', flow volume, also has three levels: low, mid, and high flow volumes.

TABLE IV

FACTORS AND LEVELS

Factors	characteristics	Level	Number of levels
A	FAB layout design	Min- and max-distance layout	2
B	OHT track configuration	Track type 1,2,3	3
C	Flow volume	Low, Mid, High	3

In the factorial design, the average delivery time of the vehicles is used as the response variable. Delivery time is one of the widely used performance measures of the AMHS in a FAB. Numerous papers including Lin et al.[1] and Mackulak and Savory[30] also support delivery time as the key performance factor of the AMHS. We define the average delivery time as the sum of the average transport time of the lots by vehicles and the average waiting time of the lots.

B. Simulation settings

Simulation experiments are performed by using AutoMod 12.3. The input data for the simulation is the lot move request described in the From/To data. **The From/To data inputs for each experiment are described in Tables V and VI. The data for 20000 and 30000 wafers/week can be obtained by multiplying by two or three times.** The method of evaluation of these tables are explained in the **Appendix I.D**. Five replications are performed in each experiment, and 90 experiments are performed overall (=$2\times3\times3\times5$). The simulation settings are similar to those of simulation analysis using AutoMod including Lin et al.[31], Wang and Lin[29], and Kong[32]. The simulation model is set as follows:

- A lot contains 25 wafers.
- The vehicle moves along the shortest path on the OHT track.
- The closest rule is used for the vehicle dispatching policy – the nearest idle vehicle to the job request is dispatched.
- The lot priority is based on the FIFO rule.
- Inter-arrival times of the lots at the stockers are exponentially distributed.

- There are no failures or unexpected events for the stockers and vehicles.
- The number of vehicles used in the simulation for each scenario is adequate to operate the AMHS.
- The OHT track has adequate capacity to meet the lot move demand.
- The size of the buffer in the stockers is set to infinity.

TABLE V

FROM/TO DATA BETWEEN BAYS FOR THE MIN-DISTANCE LAYOUT (LOTS/DAY)

	Bay 1	Bay 2	Bay 3	Bay 4	Bay 5	Bay 6	Bay 7	Bay 8	Bay 9	Bay 10	Bay 11	Bay 12	Bay 13	Bay 14	Bay 15	Bay 16
Bay 1	0	63.4	172.6	0.0	24.8	36.6	0.6	0	2.9	0	167.1	17.3	38.6	9.6	52.9	20.7
Bay 2	99.4	0	97.1	31.3	38.4	0	0	0	0	0	81.7	102.7	56.4	78.1	0.0	112.1
Bay 3	204.6	320.0	0	123.0	147.3	63.6	23.1	0	0	4.8	75.3	249.2	173.6	228.7	157.1	72.6
Bay 4	26.6	19.0	240.6	0	23.2	43.8	5.0	1.3	0	12.6	44.9	175.9	48.3	121.0	127.1	50.2
Bay 5	114.7	40.3	196.8	40.4	0	13.4	44.1	5.9	22.2	17.9	249.1	39.9	149.6	42.5	25.0	4.5
Bay 6	0.1	0	2.4	117.6	25.0	0	0	0	0	0	91.4	167.3	16.7	139.0	0	0.6
Bay 7	1.0	0	2.4	27.8	220.7	0	0	0	0	0	0	20.1	186.6	34.5	0	0
Bay 8	0.3	0	0	24.6	8.5	0	0	0	0	0	0	36.6	7.7	25.4	0	0
Bay 9	1.3	0	0	0.0	147.8	0	0	0	0	0	0	25.4	117.2	11.9	0	0
Bay 10	0.6	0	12.2	68.8	44.5	0	0	0	0	0	0	113.4	43.4	118.3	0	1.8
Bay 11	1.9	19.3	72.0	63.5	69.1	24.6	241.5	17.3	190.3	132.2	0	60.6	109.5	47.0	204.7	1.9
Bay 12	24.3	72.5	144.3	172.2	16.2	226.1	59.2	64.2	31.7	178.2	118.1	0	1.5	36.9	251.8	65.2
Bay 13	78.2	40.3	406.2	76.8	96.9	12.7	85.3	5.5	56.5	8.4	228.2	23.7	0	50.7	62.7	20.2
Bay 14	20.7	60.1	337.2	18.4	11.9	57.3	10.1	8.0	0	47.5	94.0	227.3	12.4	0	104.6	121.3
Bay 15	6.4	40.0	130.3	155.5	122.9	26.8	24.1	0	0	0.5	42.9	100.2	224.9	130.8	0	1.8
Bay 16	26.9	22.3	28.6	19.5	9.1	55.3	0.1	0.9	0	0.9	62.8	159.8	9.1	56.6	21.2	0

TABLE VI

FROM/TO DATA BETWEEN BAYS FOR THE MAX-DISTANCE LAYOUT (LOTS/DAY)

	Bay 1	Bay 2	Bay 3	Bay 4	Bay 5	Bay 6	Bay 7	Bay 8	Bay 9	Bay 10	Bay 11	Bay 12	Bay 13	Bay 14	Bay 15	Bay 16
Bay 1	0	58.4	191.6	0	0	0	0	0	0	38.6	0	9.4	0	39.5	53.4	27.5
Bay 2	20.0	0	306.8	0.2	14.0	0	8.7	118.3	51.1	23.4	31.5	24.0	19.7	87.4	108.1	38.4
Bay 3	19.2	59.3	0	36.6	170.7	53.3	133.8	418.8	216.1	112.4	166.5	24.7	79.8	108.0	52.8	52.4
Bay 4	0	0	4.7	0	82.0	2.6	0	16.9	21.5	0.6	0	55.2	61.7	214.3	68.3	27.4
Bay 5	0	19.7	147.0	11.0	0	0.5	0.7	126.0	47.1	1.1	0	29.0	58.3	78.5	10.8	0.8
Bay 6	9.6	23.5	43.0	1.0	0	0	0	12.2	0	0	0	0	0	0	140.5	127.4
Bay 7	72.1	95.6	104.7	0.1	1.4	0	0	0	0	0.0	0	0	0	0	12.6	18.9
Bay 8	25.7	68.4	294.9	84.0	127.7	0	42.0	0	145.7	15.2	14.4	96.7	320.0	144.4	340.1	123.7
Bay 9	75.2	205.1	163.7	5.6	5.8	0	9.0	120.0	0	0.3	10.0	32.1	5.7	57.1	128.7	130.1
Bay 10	79.7	51.5	68.5	3.9	1.9	0	0	53.9	0	0	0	0	0	4.6	172.4	102.3
Bay 11	86.4	165.2	120.8	1.4	1.1	0	0	2.4	0	0	0	0	0	0	7.7	15.1
Bay 12	12.2	48.7	62.4	0.7	0.0	0	0	68.6	0	0	0	0	0	140.4	137.5	185.5
Bay 13	14.4	29.6	92.3	135.1	69.2	0	0	97.1	0	0	0	7.1	0	128.8	71.7	51.7
Bay 14	2.9	7.8	13.7	0	0	247.1	93.3	41.1	308.6	165.6	147.5	239.4	51.4	0	122.8	0.7
Bay 15	0.3	8.0	74.6	183.1	42.1	25.0	4.7	475.6	128.2	135.8	27.8	84.6	79.8	317.2	0	96.5
Bay 16	0.9	10.6	15.5	92.5	14.5	28.8	13.2	292.0	30.3	45.7	2.3	53.6	20.6	121.8	256.1	0

Simulation experiments are run for 5 days after a warm-up period of 3 days. We continuously measure the average delivery time of the vehicle and that it does not change by more than 10%

as it reaches the steady-state. We found that the simulation experiments showed a small variation after 3 days. Once the simulation reaches to the steady-state, it is run for an additional 5 days for the analysis. Finally, we measure the average delivery time for one hour(to assess the performance of the AMHS) for the final 8th day.

TABLE VII

OHT VEHICLE SPECIFICATIONS

Category	Specification
Straight speed	2 m $/sec$
Curve speed	0.5 m $/sec$
Acceleration	0.98 $/sec^2$
Deceleration	1.47 $/sec^2$
Load/unload time	10 sec
Capacity (FOUP)	1
Rotation speed	90 degrees $/sec$

The specifications of the vehicles used in the simulation are summarized in Table VII. Note that these specifications are not very different from those used in other simulation studies such as that of Kong[32].

C. Simulation results

We first identify the significance of the factors A, B, and C and the interactions between them with ANOVA. The results are summarized in Table VIII. The p-values of the main factors are close to zero; therefore, we identify FAB layout, OHT track configuration, and the flow volumes as being significant. The p-values of the interactions between them are also close to zero, and therefore, they are also identified as being significant. **The value of adjust–R^2 is 74.89%. The 74.89% of total variation for the response variable is explained by significant factors A, B, C, AB, AC, BC, and ABC. All VIF values are less than 10, and therefore, there is no multicollinearity between factors.**

Next, we obtain three interaction graphs between the significant factors as shown in Figure 6, 7, and 8. First, Figure 6 shows the interaction between track configuration and flow volumes. Note that track type 1 has the higher number of shortcuts and type 3 has the least number, with type 2 somewhere in between. The average delivery times for the low flow volume are 231.90, 228.49, and 234.40 seconds for track types 1, 2, and 3, respectively, and those for the medium flow volume are 214.26, 202.24, and 227.270 seconds, respectively. The number of shortcuts are insignificant when the volume is low

TABLE VIII

RESULTS OF ANOVA

Effect	Degrees of freedom	Sum of square	Mean square	F ratio	P-value	**VIF**
A	1	10658996	10658996	18.15	0.000	**1**
B	2	14682987	7341494	12.50	0.000	**1.33**
C	2	39803822	19901911	33.89	0.000	**1.33**
AB	2	16234071	8117036	13.82	0.000	**1.33**
AC	2	23146553	11573277	19.71	0.000	**1.33**
BC	4	29459139	7364785	12.54	0.000	**1.78**
ABC	4	31841248	7960312	13.56	0.000	**1.78**

or medium. That is, the shortcuts are not effective in these cases because having more shortcuts does not improve the travel time. Furthermore, the shortcuts make the situation worse when the flow volume is high. More interesting insights are found in Figure 7.

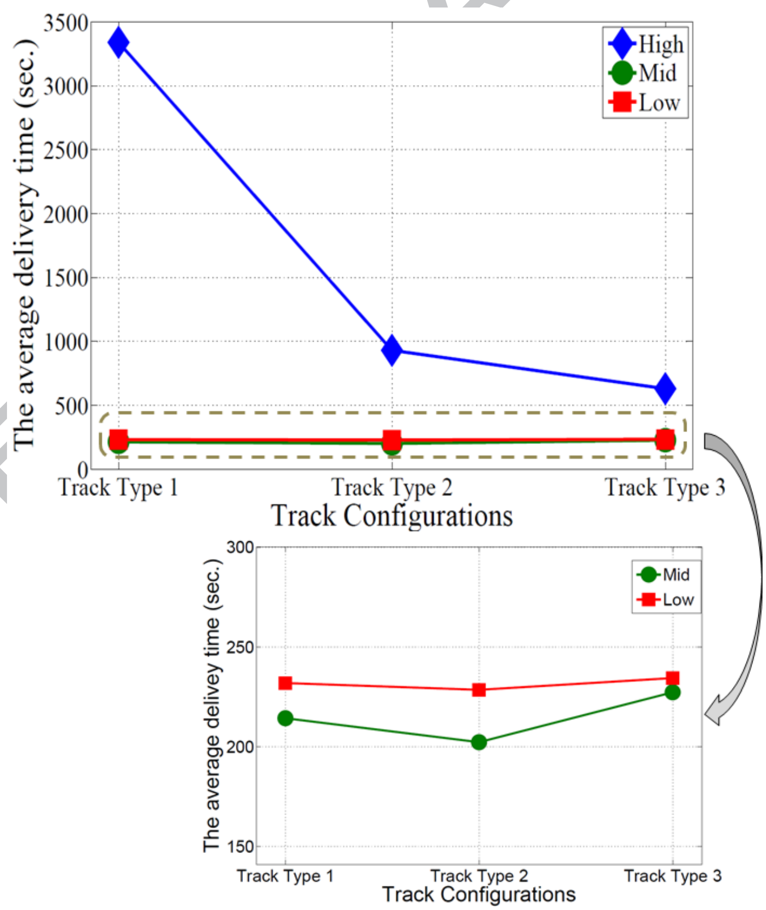


Fig. 6. Interactions between track configurations and flow volume

Figure 7 shows the interaction between the FAB layout and the flow volumes. The average delivery times for the min-distance layout are 217.42, 199.87, and 2695.10 seconds for the low, mid, and high flow volumes, respectively, but those for the max-distance layout are 245.79, 229.32, and 572.43 seconds. Note that these are slight increases in the delivery times for the low-volume case (from 217.42 to 245.79 sec.) and the mid-volume case (from 199.87 to 229.32 sec.), when the layout is changed from the minimum to maximum distance. That is, the min-distance approach is more effective because the travel time is smaller. However, the changes are insignificant compared to those of the high-volume case. Interestingly, the delivery time is significantly higher in the min-distance layout when the volume is high. In contrast to the low- and medium-volume cases, the min-distance layout performs worse than the max-distance layout when the volume is high. This result indicates that the min-distance layout design is effective only when the flow volume is limited. If the volume increases, vehicle congestion is created in concentrated areas in the min-distance layout, causing inefficient delivery in the overall FAB.

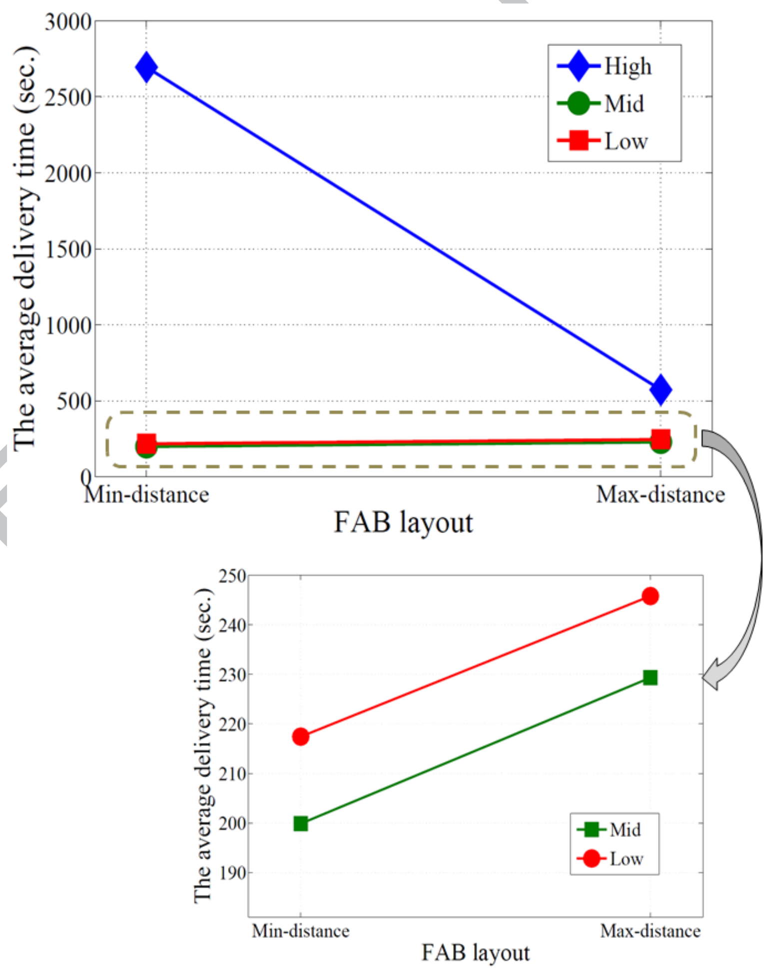


Fig. 7. Interaction between the FAB layout and flow volume

Figure 8 represents the interaction of FAB layouts and track configurations. The average delivery times for the min-distance layout are 2205.40, 535.78, and 371.21 seconds for track types 1, 2, and 3, respectively, whereas those for the max-distance layout are 318.96, 371.79, and 356.79 seconds. From the graphs, it can be observed that the number of shortcuts is insensitive in the max-distance layout. Interestingly, the delivery time is significantly higher with more shortcuts for the min-distance case because the shortcuts create more concentrated congestions. Fewer shortcuts increase the travel distance of the vehicles and spread the vehicle movers in the min-distance layout.

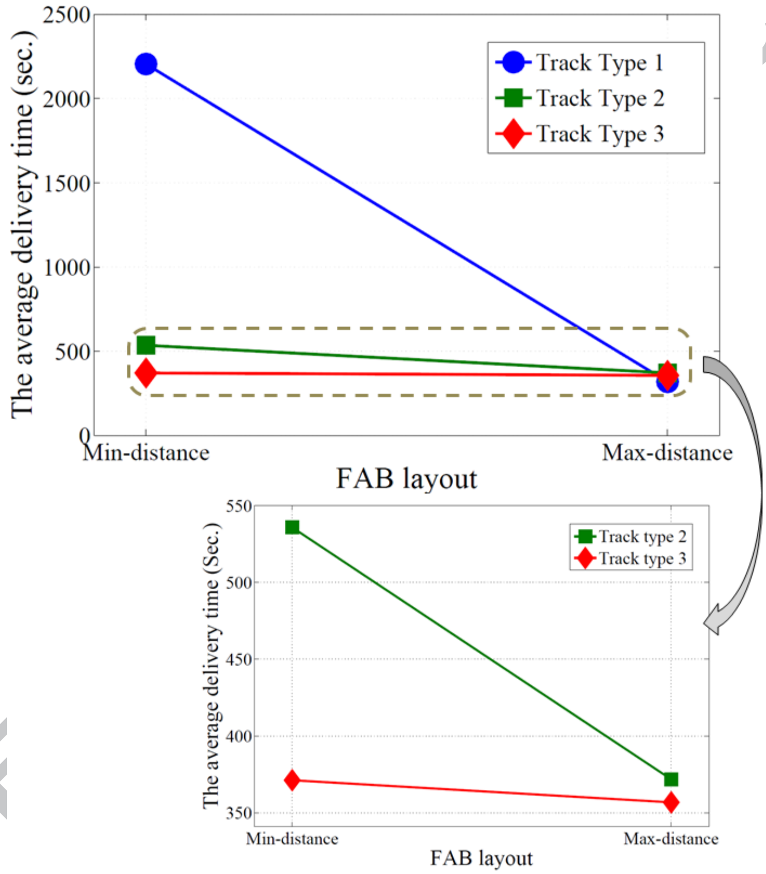


Fig. 8. Interaction between the FAB layouts and track configurations

D. *Analysis of results*

Considerable insight is provided from the results presented in the interactions graphs. With the actual FAB data, the high flow volume largely effects the performance of the AMHS. From the simulation results, we conclude the following. First, the min-distance layout is not efficient when the flow volume increases. We do not conclude that the max-distance layout is better than the min-distance layout because the max-distance layout is analyzed only to understand the opposite extreme of the min-distance layout. This is why we conclude that the min-distance layout is not effective rather than that the max-distance layout is effective. Another interesting finding is that the shortcuts in the OHT track can improve delivery times when the flow rate is moderate. However, if the volume increases, the shortcuts cause the vehicle traffic to worsen in the min-distance layout. Finally, the performance of lot delivery is sensitive to the flow volume and OHT track configuration in the min-distance layout. To obtain robust performance, the min-distance layout is not recommended.

V. Practical Implication and Suggestions

In this analysis, most of the congestions in the min-distance layout occur in the areas around Bays 2, 3, 14, and 15 as shown in Figure 9. This congestion is caused by relatively heavy flows within and between the DIF and WET processes. Since the bays for these processing tools are located next to each other, heavy traffic is created in this area. When the flow volume increases, the congestion becomes worse, and the traffic in the area eventually blocks the flow on the entire track.

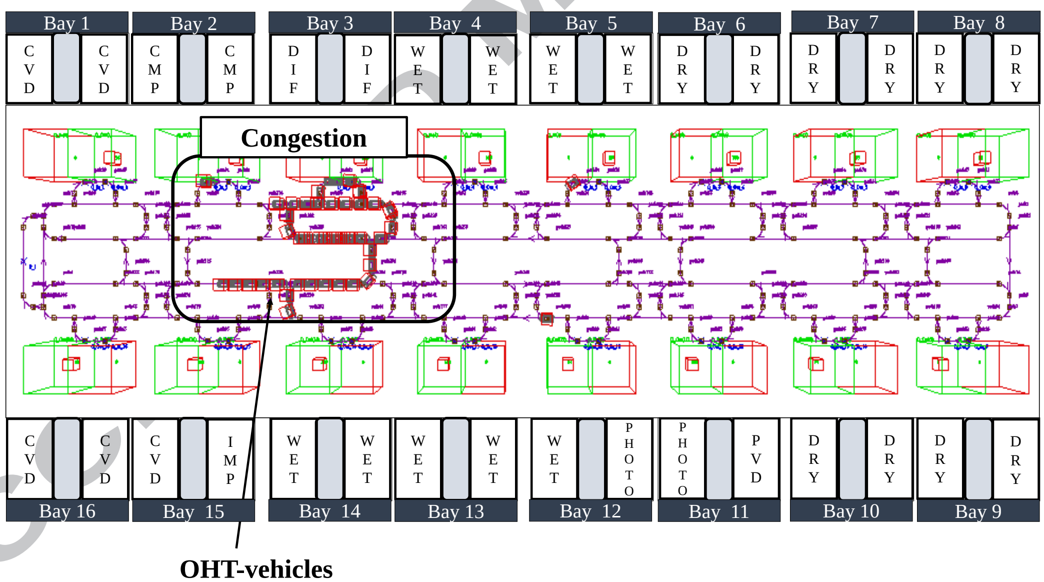


Fig. 9. **Congestion on the track(Track type 2) around Bays 2, 3, 14, and 15 (between the DIF and WET process groups)**

With the understanding of the cause of the congestion, we provide practical design suggestions with the following additional experiments. Again, note that the max-distance layout is not the alternative solution. The max-distance layout is analyzed simply as the extreme case to validate the effectiveness of the min-distance layout. Therefore, we do not suggest that practitioners adopt the max-distance approach. Let us consider the following layouts.

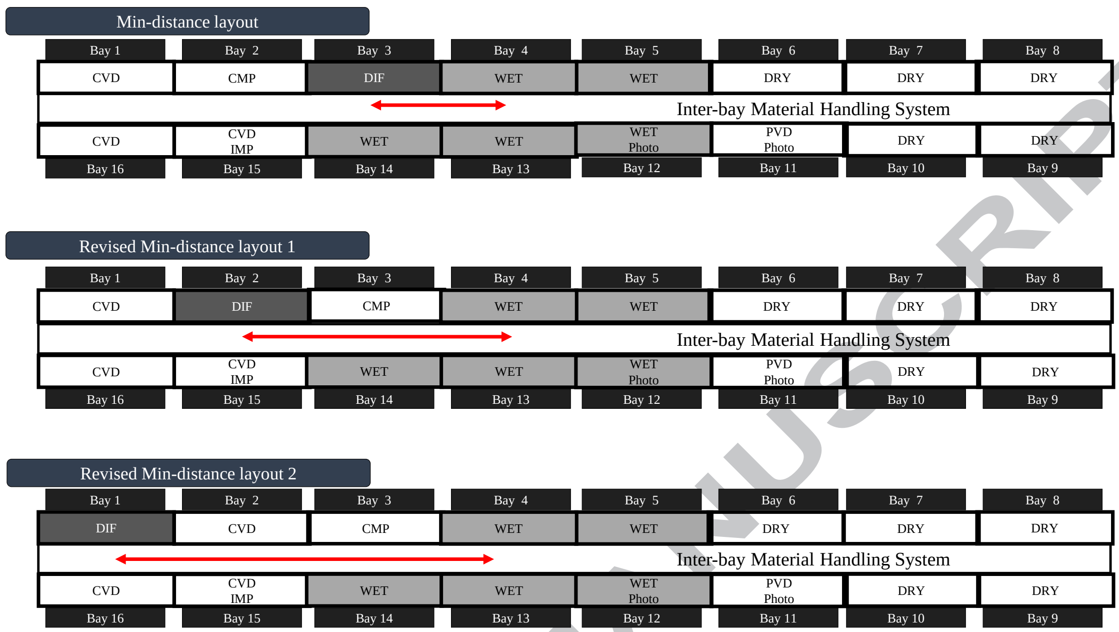


Fig. 10. Min-distance and revised min-distance FAB layouts

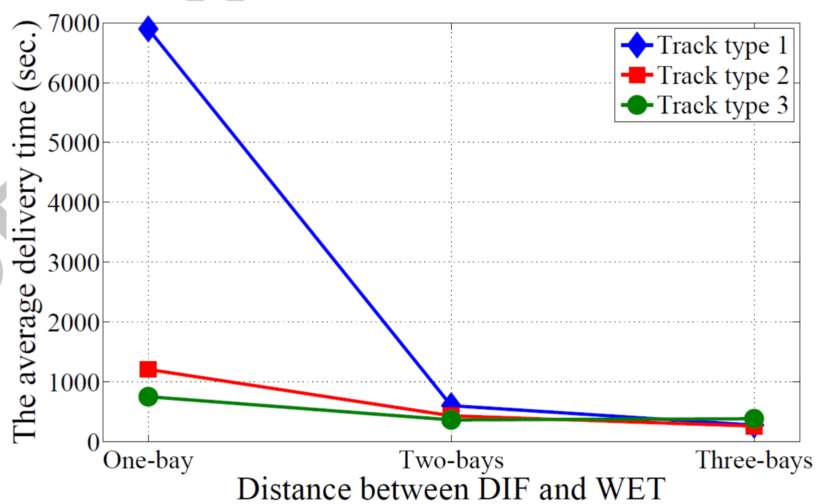


Fig. 11. Results of sensitivity analysis for distance between DIF and WET

First, allocate the bays based on the min-distance. Let us focus on the two tool groups with the highest flows between them. Since the layout design is based on the min-distance approach, the groups must be located next each other, and, in our case, they are DIF and WET. We call this layout the one-bay-away configuration, as depicted in the first panel in Figure 10, because the tool groups are located next to each other. Then, we re-allocate these tool groups (i.e., DIF and WET) two bays and three bays away from each other from their locations in the original min-distance layout. These layouts are illustrated in the second and third panels in Figure 10, respectively. Figure 11 shows the simulation experiments for the one-, two-, and three-bay-away layouts. It can be observed that the delivery time is radically reduced even when DIF and WET are separated in the two-bay-away configuration. Also note that for the two-bay- and three-bay-away solutions, the delivery time is less sensitive to the track configurations. These revised min-distance layouts (two-bay and three-bay cases) also outperform the max-distance layout. **To quantify and compare the levels of congestion for the min-, max-, and revised min-distance cases, track utilization for the top 10 most highly utilized tracks are indicated in Figure 12, and their utilization rates are provided in Table IX. For the min-distance cases, some of the tracks have greater than 50% utilization. In contrast, for the max-distance and revised min-distance cases, the tracks have significantly less utilization. From this figure, we can also observe that the delivery service request meets the demand for both configurations. However, due to the heavy congestion in the min-distance case, the delivery is delayed.**

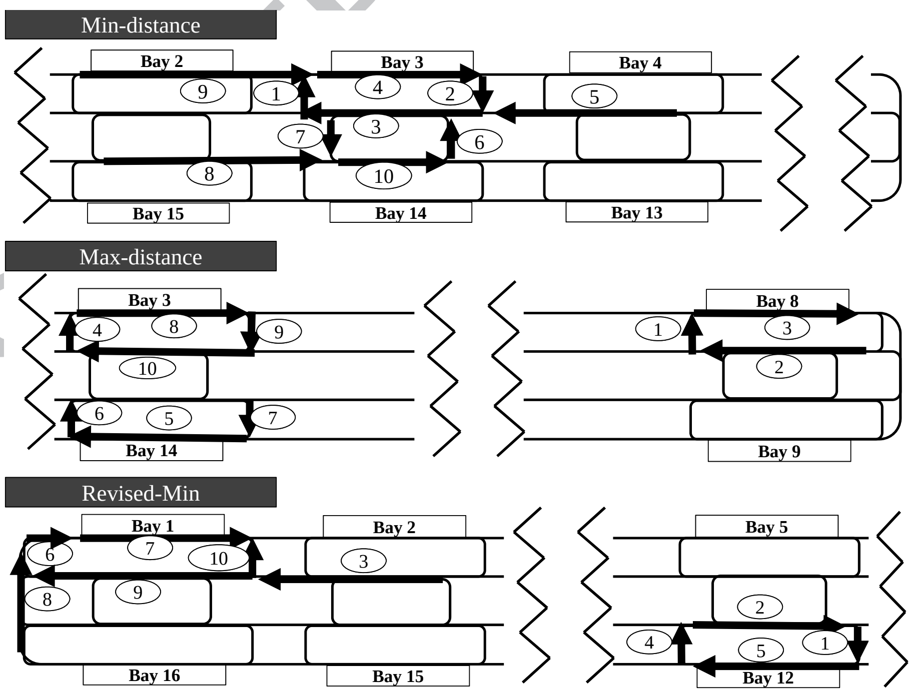


Fig. 12. Key track segments of track type 1 for each FAB layout

TABLE IX

UTILIZATION OF KEY SEGMENTS FOR HIGH FLOW VOLUME

Segment Idex	1	2	3	4	5	6	7	8	9	10
Min-distance	60.50%	55.50%	51.19%	33.88%	32.13%	21.13%	5.63%	4.83%	4.08%	1.81%
Max-distnace	32.67%	15.00%	9.69%	9.50%	8.83%	7.00%	5.50%	5.44%	3.44%	2.17%
Revised-Min2	17.50%	10.44%	10.42%	9.17%	8.38%	5.50%	2.88%	1.31%	0.75%	0.33%

Figure 13 shows the average utilization vs. flow volume (top panel) and the average delivery time vs. flow volume (bottom panel) for the min-, max-, and revised min-distance cases. The average utilization value is the average utilization for the first five highly utilized segments in the corresponding tracks configurations. For the min-distance configuration, the utilization radically increases at 28,000 wafers/week, and this is the point at which the delivery time starts to increase. However, the utilization and delivery times for the max- and revised min-distance configurations are both relatively low even though they reach 30,000 wafers/week.

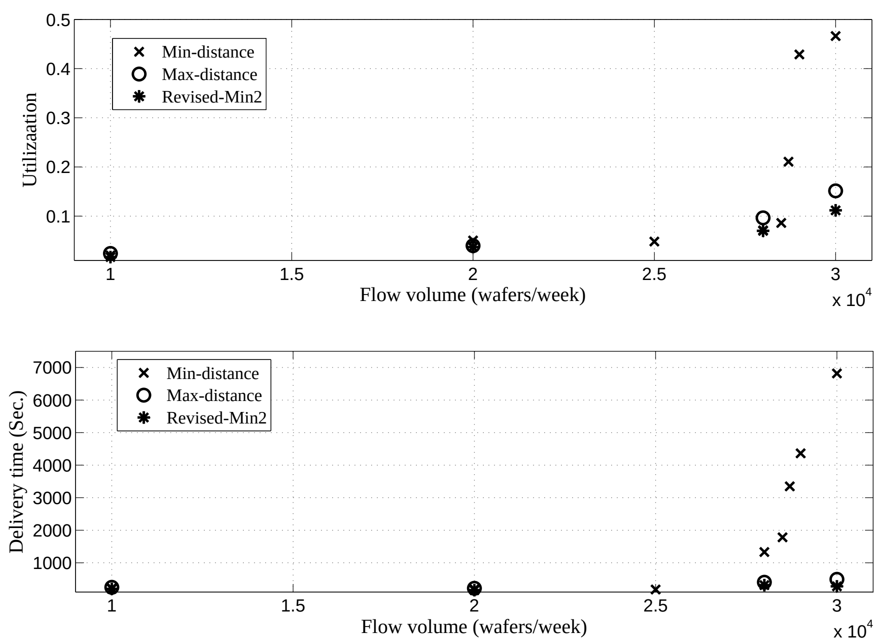


Fig. 13. Relationship between utilization and flow volume (top panel) and average delivery time and flow volume (bottom panel)

The analysis shown in Figure 13 demonstrates that the delivery and utilization trends follow the concepts of the queueing theory. As the utilization of the service reaches a certain level, the

delivery time starts to increase radically. This nonlinear behavior is also observed in our track-vehicle system. It might be interesting to investigate queueing modeling to show the behavior of the system analytically. We reserve this investigation for a future study.

Therefore, from a practical layout design, we suggest the following:

1) Allocate the bays based on the min-distance layout.
2) Try to increase the flow volume until a significant delivery delay is identified.
3) Identify the process groups with the highest traffic or flows.
4) Re-allocate the process groups by separating them by one bay.
5) If the delivery time is substantially reduced and it is insensitive to the track configuration, stop. Otherwise repeat step 4.

The idea behind this approach is that the bays with high flow are separated to avoid congestion while other bays are allocated to minimize the flow-weighted distance. This approach has been tested with other product mix scenarios, and the layouts using this approach outperform both the min- and max-distance layouts.

Note that all of the analyses are based on a particular product mix scenario. However, even though the product mix is varied, the results will not be changed. To show this, we investigated 10 different product cases in Appendix II. Table XIV shows the different product mix scenarios. The number of required tools is evaluated, and the tools are allocated with the same approach used in the previous analysis for each product mix case. The tool counts are shown in Table XV. Interestingly, even though the product mix is varied, the tool counts and allocations to the bays are not very different because DRAM process flows are not radically different from one product to another as long as they are in the same product family using the same technology node within the same allocation rule. Because the tool allocations are similar, material moves between bays will be the same. In our analysis, we used the particular product mix ratio of 4.5:3.5:1:1 because this product mix figure is provided from the manufacturer. However, as our analysis in Appendix II shows, even though the product mix is changed, the results remain the same.

VI. Conclusion

In this paper, we validate the min-distance layout design through simulation. Although some research in the past has reported the limitations of the min-distance layout approach in the context of general factory layout design, little research exists in the context of FAB design. The actual wafer flow data from a 300-mm DRAM memory chip manufacturing FAB is used for the simulation analysis. The layouts and AMHS are constructed based on the standard optimization procedures widely used in the semiconductor industry. To logically understand the effectiveness of the min-distance layout, we generate the max-distance layout, which is the opposite extreme to the min-distance approach, and compare the two layouts. We hypothesize that if the performance of the max-distance layout is better than that of the min-distance layout, we can conclude that the min-distance layout may not be the best approach to use. We use the DOE method to analyze the simulation results generated from 18 different cases: min- and max-distance layouts and each layout with three different flow-volume scenarios and with three different OHT track configurations.

We discover that the min-distance layout is not effective when the flow volume is heavy. The delivery time for the min-distance layout is significantly larger compared to that of the max-distance layout when the flow volume increases. This performance degradation remains even when additional shortcuts on the OHT tracks are provided. In fact, more shortcuts even increase the travel time on the min-distance layout when the flow volume is high. The delay in delivery time is caused by the congestion of vehicles on the OHT track. The heavy flows within and between the DIF and WET bays create significant vehicle traffic, and traffic back-up propagates to the entire track when the flow volume is increased. In the max-distance layout, however, the flow congestion is not concentrated to a limited area and flows are spread out.

Although this finding is somewhat intuitive and expected, it is significant because we numerically show that there is a certain limitation with the min-distance layout design in the context of FAB design. Our finding aligns exactly with that of the actual FAB-K case mentioned in the Introduction. While ramping up the FAB, the increased flow volume created serious congestion in a limited area that caused serious delay in lot delivery by the OHT vehicles. The layout of the FAB-K was based on the min-distance approach.

For practitioners, we provide tips for layout design from the insights gained from this analysis. Note that the goal of this paper is not to provide the best layout design method or approach. The tips are not analytically or mathematically proven (if, indeed, there is any way to prove that there is an optimal method of layout design). However, the approach is insightful and proven with the numerical data. We propose more rigorous research on the development of layout design methods from the insight gained with this research. We believe that there are great opportunities in this area.

There are certainly limitations in this study. First, the OHT vehicle movement logic is simple: lot delivery is assigned to the closest vehicle, and vehicles move according to the shortest path algorithm. Vehicle zone-control or more advanced OHT vehicle logic may need to be applied. However, we conjecture that advanced logic may improve the performance if there is enough track capacity and its effectiveness is limited when too much flow occurs within a concentrated area. However, we leave the further discussion of this issue for future research. In our research, the track configurations varied by the number of shortcuts. However, there are countless different alternatives for track configuration. The analysis of more complicated OHT track configurations with different layout options for FAB design would be an interesting study. In this research, we only consider the FAB using an inter- and intra-bay separated layout concept. However, more FABs are adopting the direct delivery approach without inter-/intra-bay separation. Analyzing the effectiveness of the min-distance approach for the direct-delivery AMHS will be very practical and interesting research. We leave this topic to future research.

APPENDIX I: DESIGN OF THE FAB LAYOUT

A. *Number of tool evaluations*

The first step in designing the layout is to evaluate the number of tools necessary to meet a certain production volume. Layout design is basically a process of allocating the necessary tools to the specific areas while meeting certain constraints. The number of tools directly affects the configuration of the layout design.

There are various methods already published to evaluate the number of tools in the FAB environment. Most modern semiconductor processing tools can process multiple recipes for multiple part types, and evaluation of the right number of tools for multiple product mixes is not a trivial job. In this research, we used the optimization method of the tool counts method presented in [33]**[34]**.

The tools counts for each of the tool groups in our study are summarized in Table XIII.

B. *Layout design*

The term *Layout Design* in this paper indicates the process of allocating the process groups to the specific bay on the FAB floor. The goal of this process is to determine which bay should be located in which area of the FAB. We use the optimization method for allocation of the process tool groups to the bays. The input values for the optimization are the tools counts for each tool group, the footprint of each tool, and the processing steps. The optimization method we use in this study is distinguished from the conventional QAP or its variant in that we consider the material flow of each of the process steps. By considering step-flow, more detailed allocation can be achieved. The OPT-Layout model is the basic structure of the layout optimization model. The definition of decision variables and parameters are summarized in Table X.

$$
\begin{aligned}
\text{(OPT-Layout)} \quad \text{Min/Max} \quad & \sum_{ijkl} C_{jl} \cdot f_{ijkl} & & (1)\\
\text{subject to} \quad & \sum_{j} t_{ij} = N_i & \forall i \quad & (2)\\
& \sum_{i} A_i^t \cdot t_{ij} \leq A_j^b & \forall j \quad & (3)\\
& \sum_{jl} f_{ijkl} = F_{ik} & \forall i,\ k \quad & (4)\\
& \sum_{l} f_{ijkl} = \tfrac{F_{ik}}{N_i} \cdot t_{ij} & \forall i,\ j,\ k \quad & (5)\\
& \sum_{j} f_{ijkl} = \tfrac{F_{ik}}{N_k} \cdot t_{kl} & \forall i,\ k,\ l \quad & (6)\\
& \sum_{ijk} f_{ijkl} = \sum_{ijk} f_{klij} & \forall l \quad & (7)\\
& t_{ij} \leq M \cdot \delta_{ij} & \forall i,\ j \quad & (8)\\
& f_{ijkl} \leq M \cdot (\delta_{ij} + \delta_{kl} - 1) & \forall i,\ j,\ k,\ l \quad & (9)\\
& \sum_{j} \delta_{ij} = B_i & \forall i \quad & (10)\\
& \sum_{j} s_{ij} = 1 & \forall i \quad & (11)\\
& \delta_{i1} = s_{i1} & \forall i \quad & (12)\\
& \delta_{ij} \leq s_{ij} + \delta_{i,j-1} & \forall i,\ j = 2, ..., J \quad & (13)
\end{aligned}
$$

The main decision variables are the number of tools in a tool group to be allocated in a bay. The objective in this OPT-Layout, which is expressed in (1), is the total flow distance. Note that the objective function is either minimum or maximum depending on which layout approach is used as described in Section III-D. The constraint (2) is the tool counts equality condition. The equation (3) indicates the bay space constraint. The equations from (4) to (7) are the flow constraints. The conservation of flow between the bays and tools must be maintained. The equations (8) and (9) are the constraints enforcing tools with the same tool types to be allocated to close-by bays. For example, in Figure 4, which shows the layouts generated from the optimization algorithm, the WET process requires 5 bays. These WET bays are located closer together in both the min-distance and max-distance layouts. The remaining constraints cover auxiliary conditions. **This model is basically the extension of the model proposed in [34].** A more detailed analysis and an algorithm for optimization are found in [35]. Note that in practice, the size of the bay is continuously modified during this design stage. This issue is also discussed in [35].

TABLE X

NOTATIONS FOR THE LAYOUT DESIGN MODEL

	Notation	Definition
Decision Variables	f_{ijkl}	Flow amount from process group *i* of bay *j* to process group *k* of bay *l*.
	t_{ij}	Integer, number of tools for process group *i* assigned to bay *j*.
	δ_{ij}	Binary, 1 if any tool of process group *i* is assigned to bay *j*, 0 otherwise.
	s_{ij}	Binary, 1 if *j* is the first index of the bays assigned for process group *i*, 0 otherwise.
Parameters	C_{ij}	**Distance** from bay j to bay l.
	P_i	Penalty cost of assigning additional tools to process group *i*.
	N_i	Number of tools planned for process group *i*.
	B_i	Number of bays planned for process group *i*.
	A_i^t	Area space required for a tool of the process group *i*.
	A_j^b	Area space available for the bay *j*.
	F_{ik}	Material flow amount from process group *i* to *k*.
	M	Big number M, the upper bound for the variables t_{ij} and f_{ijkl}.

C. Tool allocation to the bay

The previous step concerns allocation of the bay in the FAB, whereas this step concerns which specific tool has to be allocated to which bay. In Figure 4, for example, there are five WET bays, and we need to figure out which tools in the WET process group should be allocated to which WET bay. We also use optimization modeling for this allocation problem and call it OPT-Tool. The primary goal of the OPT-Tool model is to improve the flow efficiency of the material flow within the bay while it offloads the material moving on the inter-bay system. Therefore, it is also called the intra-bay optimization model. The decision variables and parameters are listed in Table XI.

$$
\begin{aligned}
&\text{(OPT-Tool)} \quad \text{Minimize} && \sum_{si^s kn} M_{kn} \cdot f_{si^s kn} && && (14)\\
&\qquad\qquad\quad \text{subject to} && \sum_{kn} f_{si^s kn} \geq D_s && \forall s,\ i^s && (15)\\
& && \sum_{k} f_{si^s kn} = \sum_{k} x_{s,i^s+1,n,k} && \forall s,\ n,\ and\ i^s = 1, ..., I^s - 1 && (16)\\
& && \sum_{n} f_{si^s kn} \leq \sum_{j} C_{si^s j} \cdot y_{si^s jk} && \forall s,\ i^s,\ k && (17)\\
& && \sum_{si^s} y_{si^s jk} \leq t_{jk} && \forall j,\ k && (18)\\
& && \sum_{k} t_{jk} = N_j && \forall j && (19)\\
& && \sum_{si^s j} A^t_j \cdot t_{jk} \leq A^b_k && \forall k && (20)\\
& && t_{jk} \leq M \cdot S_{jk} && \forall j,\ k && (21)
\end{aligned}
$$

The objective function, which is depicted in (14), is to minimize the material flow, and constraints from (15) to (17) are concerned with flow conservation and flow demand. The constraints (18) and (19) relate to tool counts. The constraint (20) is the restriction on bay space, and the last equation, (21), is an auxiliary constraint relating to allowable tools in the bay. **This model is the extension of the model proposed in [34], and more specific details on the optimization model can be found in [35].**

TABLE XI

NOTATIONS FOR THE TOOL ALLOCATION MODEL

	Notation	Definition
Decision Variables	$f_{si^s kn}$	Wafer flow rate per minute at step *i* of product *s* from bay *k* to *n*.
	$y_{si^s jk}$	Number of tools *j* allocated to bay *k* for step *i* of product *s*
	t_{jk}	Integer, Number of tools *j* allocated to bay *k*.
Parameters	M_{kn}	**Distance** between bay *k* to bay *n*.
	D_s	Required production rate of the product *s*.
	$C_{si^s j}$	Production rate or capacity of the tool *j* for step *i* of product *s*.
	N_j	Number of tools *j* available for allocation, output result from the design step 1.
	S_{jk}	0 or 1 setup indicator for the tool *j* at bay *k*, output result from the design step 2.
	A^t_j	Relative space size of the tool *j*.
	A^b_k	Relative space size of the bay *k*.

D. From/To matrix

Once the previous step is completed, tools are assigned to specific bays. However, to simulate inter-bay material flow, the lot move rate between bays needs to be evaluated. This rate, i.e., the rate of lot moves from one bay to another bay, is often described in a matrix form and is called a *From/To* table or matrix. Each element in the table indicates the lot move per unit time from one bay to another bay. This From/To table is used as input to the inter-bay simulation.

Note that due to the complexity of the modern FAB, a simulation program such as AutoMod, which is the most widely used software package for FAB simulation, cannot simulate the lot moves based on the flow of the processing steps. Instead, lot moves between bays in the From/To table are used as an input and simulate the vehicle deliveries.

The generation of the From/To table is not intuitive because tools are processing multiple steps, and the same tools may be allocated to different bays. We also use the optimization model for the generation of the From/To table as shown in OPT-From/To. The decision variables and parameters are listed in Table XII.

$$\text{(OPT-From/To)} \quad \text{Minimize} \quad \sum_{ijklmn} x_{ijklmn} \cdot D_{kn} \tag{22}$$

$$\text{subject to} \quad \sum_{kmn} x_{ijklmn} = F_{il} \cdot P_{ij} \quad for \quad all \quad i, j, l \tag{23}$$

$$\sum_{ijk} x_{ijklmn} = \sum_{pqr} x_{lmnpqr} \quad for \quad all \quad l, m, n \tag{24}$$

$$\sum_{mn} x_{ijklmn} = F_{il} \cdot P_{ij} \cdot A_{jk} \quad for \quad all \quad j, k \tag{25}$$

TABLE XII

NOTATIONS FOR THE LAYOUT PLANNING MODEL

	Notations	Definition
Decision variable	x_{ijklmn}	Number of lots from step *i*, tool *j*, and bay *k* to step *l*, tool *m*, and bay *n*.
Parameters	F_{il}	Number of lots from step *i* to step *l*.
	D_{kn}	Rectilinear distance between the bays *k* and *n*.
	P_{ij}	Step-to-tool ratio for step *i* and tool *j*.
	A_{jk}	Tool-to-bay ratio for tool *j* and bay *k*.

Again, more details on this optimization model are given in [36].

E. Number of vehicles evaluation

The last step in running the simulation for the inter-bay moves is to evaluate the appropriate number of OHT vehicles needed. Too few vehicles will cause an inter-bay AMHS bottleneck, whereas too many will result in unnecessary vehicle blockage by idling vehicles because the idle vehicles may block the working vehicles on the track. **We evaluate the number of vehicles with a standard simulation method used in industry as described in [2][36].** The basic idea of the method is that it first analytically evaluates a rough estimate of the required number of vehicles described in Equation (26).

Then number of vehicle is increased for each simulation run and evaluate it average delivery time. As shown in Figure 14, the average delivery time is decreased when more vehicle is added. Then it starts increases more vehicles. Then we select the lowest point as the vehicle counts for the case. This simulation based approach is employed to evaluate the vehicle counts for the analysis in the study. We understand that there are more rigorous method available proposed in the academic literature. However, we found that this method is good enough for our purpose of the study.

$$Number\ of\ vehicles = \frac{(Total\ flow\ in\ inter\text{-}bay\ AMHS/unit\ time) * (Average\ travel\ time/one\ vehicle)}{(Utilization\ rate/one\ vehicle)} \tag{26}$$

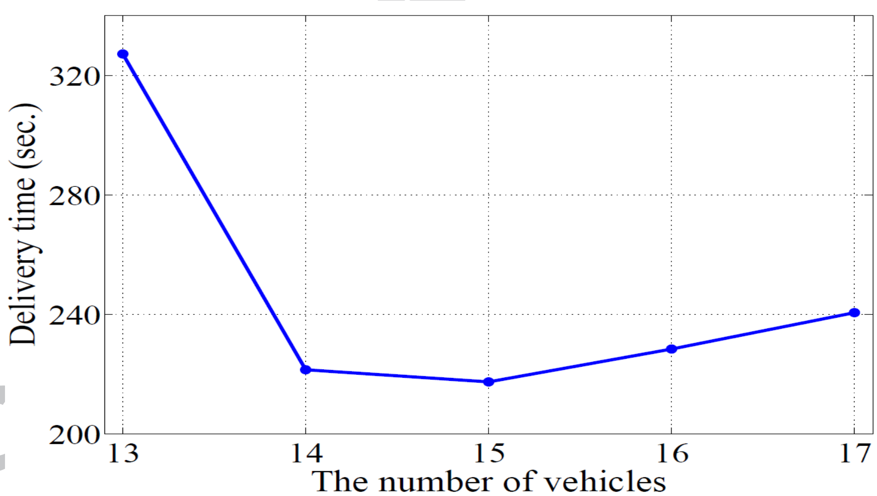


Fig. 14. Evaluation of the number of vehicles

TABLE XIII

RESULTS OF THE CAPACITY PLANNING STEP

Tool Group ID	Process group	Total number of tools for 10000 wafers/week	Total number of tools for 20000 wafers/week	Total number of tools for 30000 wafers/week	Tool Group ID	Process group	Total number of tools for 10000 wafers/week	Total number of tools for 20000 wafers/week	Total number of tools for 30000 wafers/week
1	CMP	7	14	21	100	DRY	2	4	6
3	CMP	4	8	12	101	DRY	4	8	12
4	CMP	3	6	9	104	DRY	4	8	12
5	CMP	2	4	6	105	DRY	7	14	21
7	CMP	2	4	6	106	DRY	4	8	12
14	CMP	3	6	9	107	DRY	4	8	12
15	CMP	1	2	3	108	DRY	1	2	3
16	CMP	4	8	12	110	DRY	7	14	21
17	CMP	2	4	6	111	DRY	1	2	3
19	CVD	2	4	6	112	DRY	22	44	66
20	CVD	3	6	9	115	DRY	12	24	36
22	CVD	2	4	6	118	IMP	1	2	3
23	CVD	2	4	6	120	IMP	2	4	6
24	CVD	1	2	3	121	IMP	2	4	6
25	CVD	10	20	30	122	IMP	2	4	6
27	CVD	8	16	24	123	IMP	1	2	3
28	CVD	7	14	21	124	IMP	1	2	3
29	CVD	4	8	12	127	IMP	6	12	18
35	CVD	1	2	3	128	IMP	2	4	6
36	CVD	3	6	9	130	IMP	3	6	9
37	CVD	6	12	18	131	IMP	1	2	3
39	CVD	1	2	3	133	IMP	2	4	6
41	CVD	2	4	6	134	PVD	1	2	3
42	CVD	9	18	27	135	PVD	1	2	3
43	CVD	3	6	9	136	PVD	1	2	3
44	CVD	7	14	21	143	PVD	2	4	6
45	CVD	4	8	12	145	PVD	3	6	9
47	CVD	2	4	6	147	PVD	2	4	6
48	DIF	7	14	21	150	PHO	6	12	18
50	DIF	2	4	6	154	PHO	4	8	12
51	DIF	2	4	6	158	PHO	1	2	3
53	DIF	5	10	15	159	PHO	1	2	3
56	DIF	2	4	6	162	PHO	1	2	3
59	DIF	1	2	3	170	PHO	5	10	15
60	DIF	1	2	3	175	WET	1	2	3
61	DIF	4	8	12	177	WET	1	2	3
62	DIF	5	10	15	178	WET	5	10	15
64	DIF	1	2	3	180	WET	4	8	12
72	DIF	11	22	33	181	WET	3	6	9
73	DIF	9	18	27	183	WET	1	2	3
76	DIF	7	14	21	184	WET	1	2	3
77	DIF	3	6	9	192	WET	1	2	3
79	DIF	2	4	6	196	WET	4	8	12
80	DIF	3	6	9	198	WET	3	6	9
84	DIF	5	10	15	199	WET	1	2	3
86	DRY	15	30	45	200	WET	8	16	24
88	DRY	1	2	3	201	WET	1	2	3
90	DRY	1	2	3	203	WET	2	4	6
91	DRY	10	20	30	204	WET	34	68	102
92	DRY	2	4	6	206	WET	3	6	9
94	DRY	8	16	24	211	WET	3	6	9
95	DRY	26	52	78	212	WET	3	6	9
99	DRY	4	8	12					

APPENDIX II: DIFFERENT PRODUCT MIX CASE

In this study, we focus on one product mix for the simulation experiments. The product mix can be changed according to the production ramp-up phases or the production plans. A different product mix can cause the required number of tools for each tool group to change, and thus, it may affect the FAB layout. However, our analysis found that as long as the products are in the same DRAM family, FAB layout is less likely to change even if the product mix is varied.

Table XIV shows 10 different product mix scenarios for the four products in the DRAM product family, and Table XV lists the tool counts for each scenario. Note that although the product mix is varied, the number of tool counts are not radically different because as long as the products are in the same DRAM family, the process steps are similar and the required tools are not very different.

Figure 15 shows the from-to flow between the tool groups. The x-axis for each subplot indicates the 10 different product mix scenarios, and the y-axis indicates the from-to amount of flow. Note that for each subplot, the flow amount is almost similar even though the product mix scenarios are different. The tools are allocated based on the tool counts and flow volume between tool groups. According to the tool counts in Table XV and the from-to flow volumes in Figure 15, no significant difference was observed among the different product mix scenarios. As a result, the tool allocations should also be similar for the different product mix scenarios.

TABLE XIV
PRODUCT MIX SCENARIOS

Product mix scenario	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
Product 1	1.5	2	2.5	2.5	3.5	4	4.5	5	3	2
Product 2	1.5	2	2.5	3.5	3.5	4	3.5	3	3	2
Product 3	3	2.5	2	2	1.5	1	1	1	1	4
Product 4	4	3.5	3	2	1.5	1	1	1	3	2

TABLE XV

TOTAL NUMBER OF TOOLS IN EACH PROCESS GROUP FOR EACH SCENARIO

Process group	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
CMP	27	28	29	29	29	29	28	28	28	27
CVD	69	70	72	73	77	77	77	77	77	67
DIF	67	68	69	69	68	70	70	70	70	64
DRY	135	136	136	135	134	135	135	135	139	128
IMP	24	25	23	23	24	23	23	23	23	24
PVD	11	10	10	10	10	10	10	10	10	9
Photo	18	18	17	18	17	18	18	18	18	18
WET	80	79	80	79	80	79	79	79	81	78

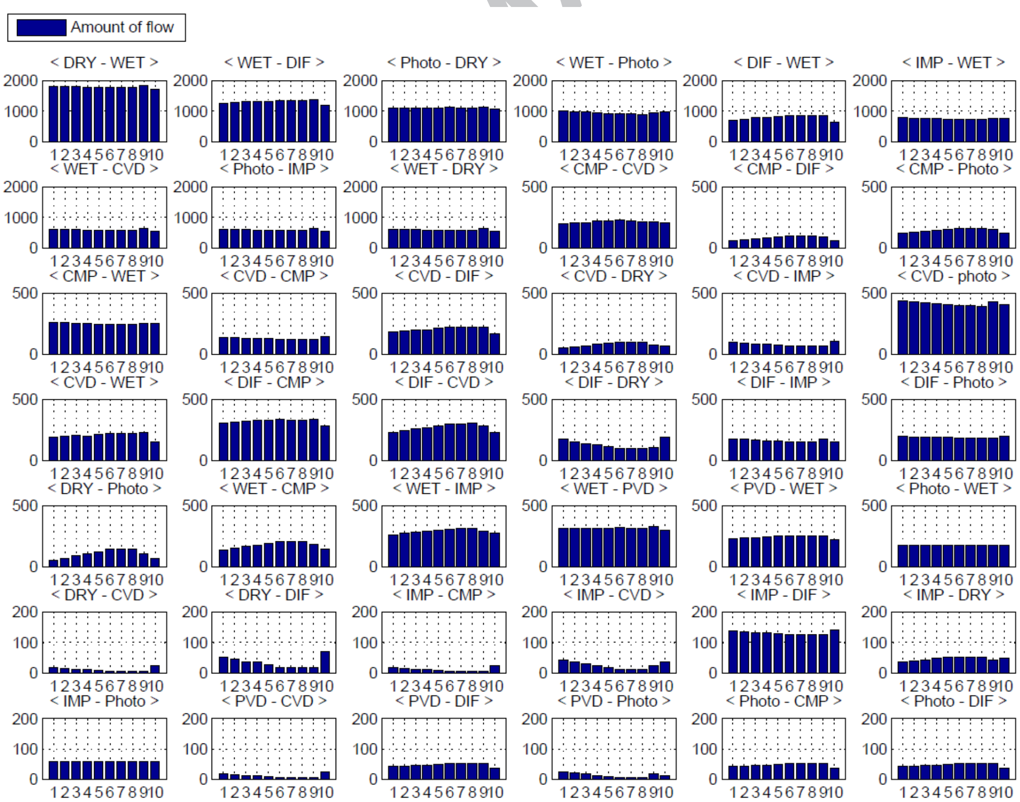


Fig. 15. The from-to flow between the tool groups for all product mix scenarios

References

[1] J. T. Lin, F. Wang, and C. Yang, "The performance of the number of vehicles in a dynamic connecting transport amhs," *International Journal of Production Research*, vol. 43, no. 11, pp. 2263–2276, 2005.

[2] Y. J. Jang, "The fab operation analysis for the major semiconductor fabs in south korea (korean) kaist-internal report," Department of Industrial and Systems Engineering, KAIST, Daejeon, South Korea, Tech. Rep., 2014.

[3] R. Meller, "The multi-bay manufacturing facility layout problem," *International Journal of Production Research*, vol. 35, no. 5, pp. 1229–1237, 1997.

[4] I. Castillo and B. A. Peters, "Integrating design and production planning considerations in multi-bay manufacturing facility layout," *European Journal of Operational Research*, vol. 157, no. 3, pp. 671–687, 2004.

[5] Y.-C. Ho and T.-W. Liao, "The layout design of semiconductor bays with spine and perimeter inter-bay guide path loops," *International Journal of Production Research*, vol. 50, no. 3, pp. 719–741, 2012.

[6] S. Benjaafar, "Modeling and analysis of congestion in the design of facility layouts," *Management Science*, vol. 48, no. 5, pp. 679–704, 2002.

[7] M. Zhang, R. Batta, and R. Nagi, "Modeling of workflow congestion and optimization of flow routing in a manufacturing/warehouse facility," *Management Science*, vol. 55, no. 2, pp. 267–280, 2009.

[8] ——, "Designing manufacturing facility layouts to mitigate congestion," *IIE Transactions*, vol. 43, no. 10, pp. 689–702, 2011.

[9] S. Singh and R. Sharma, "A review of different approaches to the facility layout problems," *The International Journal of Advanced Manufacturing Technology*, vol. 30, no. 5-6, pp. 425–433, 2006.

[10] A. Drira, H. Pierreval, and S. Hajri-Gabouj, "Facility layout problems : A survey," *Annual Reviews in Control*, vol. 31, no. 2, pp. 255–267, 2007.

[11] Y. A. Bozer and R. D. Meller, "A reexamination of the distance-based facility layout problem," *IIE transactions*, vol. 29, no. 7, pp. 549–560, 1997.

[12] I. Castillo and B. A. Peters, "An extended distance-based facility layout problem," *International Journal of Production Research*, vol. 41, no. 11, pp. 2451–2479, 2003.

[13] P. M. Hahn, B.-J. Kim, M. Guignard, J. M. Smith, and Y.-R. Zhu, "An algorithm for the generalized quadratic assignment problem," *Computational Optimization and Applications*, vol. 40, no. 3, pp. 351–372, 2008.

[14] Y. Bukchin, R. D. Meller, and Q. Liu, "Assembly system facility design," *IIE Transactions*, vol. 38, no. 1, pp. 53–65, 2006.

[15] J.-G. Kim and Y.-D. Kim, "Layout planning for facilities with fixed shapes and input and output points," *International Journal of Production Research*, vol. 38, no. 18, pp. 4635–4653, 2000.

[16] R. Muther, *Systematic layout planning*. Cahners books, 1973.

[17] J. A. Tompkins, J. A. White, Y. A. Bozer, and J. M. A. Tanchoco, *Facilities planning*. John Wiley & Sons, 2010.

[18] J. C. Chen, R.-D. Dai, and C.-W. Chen, "A practical fab design procedure for wafer fabrication plants," *International Journal of Production Research*, vol. 46, no. 10, pp. 2565–2588, 2008.

[19] B. A. Peters and T. Yang, "Integrated facility layout and material handling system design in semiconductor fabrication facilities," *IEEE Transactions on Semiconductor Manufacturing*, vol. 10, no. 3, pp. 360–369, 1997.

[20] T. Yang, M. Rajasekharan, and B. A. Peters, "Semiconductor fabrication facility design using a hybrid search methodology," *Computers & Industrial Engineering*, vol. 36, no. 3, pp. 565–583, 1999.

[21] T. Yang, C.-T. Su, and Y.-R. Hsu, "Systematic layout planning : A study on semiconductor wafer fabrication facilities," *International Journal of Operations & Production Management*, vol. 20, no. 11, pp. 1359–1371, 2000.

[22] J.-H. Ting and J. Tanchoco, "Unidirectional circular layout for overhead material handling systems," *International Journal of Production Research*, vol. 38, no. 16, pp. 3913–3935, 2000.

[23] T. Cheng, "A simulation study of automated guided vehicle dispatching," *Robotics and Computer-Integrated Manufacturing*, vol. 3, no. 3, pp. 335–338, 1987.

[24] W.-C. Chiang, P. Kouvelis, and T. L. Urban, "Incorporating workflow interference in facility layout design : The quartic assignment problem," *Management Science*, vol. 48, no. 4, pp. 584–590, 2002.

[25] J. Jang, J. Suh, and P. M. Ferreira, "An agv routing policy reflecting the current and future state of semiconductor and lcd production lines," *International Journal of Production Research*, vol. 39, no. 17, pp. 3901–3921, 2001.

[26] K. Im, K. Kim, Y. Moon, T. Park, and S. Lee, "The deadlock detection and resolution method for a unified transport system," *International Journal of Production Research*, vol. 48, no. 15, pp. 4423–4435, 2010.

[27] K. Bartlett, J. Lee, S. Ahmed, G. Nemhauser, J. Sokol, and B. Na, "Congestion-aware dynamic routing in automated material handling systems," *Computers & Industrial Engineering*, vol. 70, pp. 176–182, 2014.

[28] T. Yang and B. A. Peters, "A spine layout design method for semiconductor fabrication facilities containing automated material-handling systems," *International Journal of Operations & Production Management*, vol. 17, no. 5, pp. 490–501, 1997.

[29] F.-K. Wang and J. T. Lin, "Performance evaluation of an automated material handling system for a wafer fab," *Robotics and Computer-Integrated Manufacturing*, vol. 20, no. 2, pp. 91–100, 2004.

[30] G. T. Mackulak and P. Savory, "A simulation-based experiment for comparing amhs performance in a semiconductor fabrication facility," *IEEE Transactions on Semiconductor Manufacturing*, vol. 14, no. 3, pp. 273–280, 2001.

[31] J. T. Lin, F.-K. Wang, and C.-K. Wu, "Simulation analysis of the connecting transport amhs in a wafer fab," *IEEE Transactions on Semiconductor Manufacturing*, vol. 16, no. 3, pp. 555–564, 2003.

[32] S. H. Kong, "Two-step simulation method for automatic material handling system of semiconductor fab," *Robotics and Computer-Integrated Manufacturing*, vol. 23, no. 4, pp. 409–420, 2007.

[33] S. J. M. Lars Mnch, John W. Fowler, *Production planning and control for semiconductor wafer fabrication facilities: Modeling, analysis, and systems.* Springer Science & Business Media, 2012, vol. 52.

[34] G. Yu, "Analytical method for the layout design of a semiconductor fab to maximize its operation efficiency," Master's thesis, KAIST, Daejeon, Korea, 2013. [Online]. Available: http://hdl.handle.net/10203/182470

[35] G. Yu and Y. J. Jang, "Analytical approach for the semiconductor fab design procedure," in-Progress.

[36] J. Kim, Y. J. Jang, and B. Kurtz, "Simulation verification for layout design - Is shortest distance always good?" in *Proceedings of the International Symposiumon Semiconductor Manufacturing*, 2014.

Paper Title: Semiconductor FAB Layout Design Analysis with 300-mm FAB Data: "Is minimum distance-based layout design best for semiconductor FAB design?"

Highlights:

- The minimum distance based layout is verified.
- The min-distance based layout is not effective when the product volume is high.
- The process groups between which the highest flow traffic should be separated.
- Tips for robust layout design are provided.