Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

DECENTRALIZED DISPATCHING FOR BLOCKING AVOIDANCE IN AUTOMATED MATERIAL HANDLING SYSTEMS

Yen-Shao Chen

Department of Electrical Engineering
National Taiwan University
No. 1, Sec. 4, Roosevelt Rd., Da'an Dist.
Taipei City 106, TAIWAN

Cheng-Hung Wu
Shi-Chung Chang

Institute of Industrial Engineering
National Taiwan University
No. 1, Sec. 4, Roosevelt Rd., Da'an Dist.
Taipei City 106, TAIWAN

ABSTRACT

Advancements in communication and computing technologies have made promising the decentralized control of automated material handling systems (AMHS) to alleviate blocking and congestion of production flows and raise productivity in an automated large-scale factory. With the growing availability of edge computing and low-cost mobile communications, either among vehicles (V2V) or between vehicles and machines (V2M), decentralized vehicle control may exploit frequent and low latency exchanges of neighborhood information and local control computation to increase AMHS operation efficiency. In this study, a decentralized control algorithm design, BALI (blocking avoidance by exploiting location information) algorithm, exploits V2X exchanges of local information for transport job matching, blocking inference, and job exchange for vehicle dispatching in AMHS. Performance evaluation of the BALI algorithm by discrete-event simulation shows that the BALI algorithm can significantly reduce blocking and congestion in production flows as compared to commonly used Nearest Job First rule-based heuristics.

1 INTRODUCTION

In semiconductor manufacturing, since one fully loaded front-opening unified pod (FOUP) containing 300-mm wafers weighs 9 kgs (Wang, Chung and Wu 2004) discouraging manual handling, almost all existing 300-mm fabs apply automated material handling systems (AMHS) for material handling. The AMHS consists of annular and unidirectional rails and transport vehicles called overhead hoist transport (OHT), as shown in Figure 1. The annular rails surrounded by machines are termed intra-bays while the rails connecting intra-bays are referred to as inter-bays. Among intra-bays and inter-bays, OHTs pick up and deliver FOUPs from and to machines.

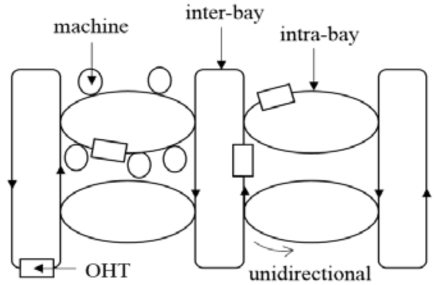

Figure 1: Typical AMHS configuration.

OHT dispatching in 300-mm fabs is a complex and challenging problem. (Yan 2003) indicates that the dispatching problem of one vehicle transporting goods in a loop is equivalent to the travelling salesman problem (TSP) which is NP-hard. The computational complexity of applying enumeration to achieve the optimal solution of the 1-OHT problem is already $\Theta(n!)$, and the complexity climbs to $\Theta(2^n n_1! n_2!)$ in the 2-OHT problem where n_1 represents the number of transport jobs of the first OHT, n_2 represents the number of transport jobs of the second OHT, and n equals $n_1 + n_2$. Modeling of the TSP can be found in (Applegate et al. 2011), (Bellman 1962), and (Miller, Tucker, and Zemlin 1960). While there are hundreds of OHTs running simultaneously in a fab, the computational complexity is extremely high.

In most fabs at semiconductor industry, OHT dispatching is executed by one central controller based on centralized information. In consideration of computation time limitation and fab-wide dispatching problem complexity, most central controllers adopt the Nearest Job First (NJF) method and its variants to match demand and supply of transport jobs. Machines generate transport requests and represent the demand side, while OHTs execute transport jobs and represent the supply side. The NJF method allows empty OHTs to find the nearest machine with transport requests. The NJF method is easy to be implemented in the central controller and requires computation using local information, which is essentially distributed decision by centralized computation. The NJF-based dispatching has led to reasonable AMHS performance when little uncertainty exists in the factory. However, once uncertainty such as nondeterministic machine process time rises, the time delay and the non-optimal solution of OHT dispatching incurred by the NJF method would cause congestion, deteriorate AMHS efficiency, and increase FOUP waiting time and production cycle time.

This paper is amongst the first papers demonstrating the potential of decentralized vehicle control methods in AMHS. Most papers propose centralized dispatching rules to solve the vehicle dispatching and blocking problems. (Huang, Lu, and Fu 2007) utilize the real-time traffic information and apply Markov decision model to estimate the cost of traversing a congested edge. (Bartlett et al. 2014) indicate that vehicular congestion causes transport delay and reduces production efficiency. A congestion-aware dynamic routing strategy is demonstrated to reroute vehicles as congestion status changes. (Nishi and Tanaka 2012) propose a Petri net decomposition approach to accomplish simultaneous dispatching and conflict-free routing for bidirectional automated guided vehicle (AGV) systems. Yet few papers have discussed decentralized OHT dispatching rules addressing blocking issues.

The rest of this paper is organized as follows. Section 2 describes the problem and a mathematical model. The proposed decentralized dispatching algorithm is given in Section 3. Section 4 illustrates the simulation results and managerial insights for fab managers. Section 5 concludes this paper.

2 PROBLEM DESCRIPTION AND MODEL FORMULATION

For the convenience of simulation, the manufacturing and transportation problem in practical fabs is reasonably simplified in the current study. A practical 300-mm fab has 40 to 50 intra-bays with each intra-bay containing 15 to 20 machines (Tung et al. 2013) and on average 10 OHTs per intra-bay. The current study refers to the fab layout of spine configuration in (Peters and Yang 1997) and simplifies the practical problem as follows. Consider 1 intra-bay, 4 machines, and 3 OHTs in the intra-bay. The machines send out transport requests stochastically and the OHTs execute the transport jobs. A decentralized dispatching algorithm is proposed to generate better performance than that of the NJF method. The performance indices considered in this study include throughput, cycle time, and FOUP waiting time.

A mathematical representation of the AMHS is constructed to model the characteristics of the manufacturing and transportation systems and the advantages and disadvantages of the NJF method. As shown in Figure 2, assume the $\mathcal{v}^{\text{th}}$ OHT arrives at the load port of its target machine at time $q_{\mathcal{v}}$, picks up the FOUP and then delivers it to the unload port of the destination machine at time $s_{\mathcal{v}}$. The FOUP loading and unloading time are both assumed to be $\mathcal{L}$. $q_{\mathcal{v}\mathcal{M}}$ denotes the time when the $\mathcal{v}^{\text{th}}$ OHT arrives at machine $\mathcal{M}$. As shown in Figure 3, on a unidirectional rail, OHT $\mathcal{b}$ is behind OHT a and OHT a is loading a FOUP at the load port of machine $\mathcal{A}$. The sufficient condition of blocking is

$$q_a + \mathcal{L} > q_{b\mathcal{A}}. \tag{1}$$

If blocking happens, then the blocking time will be

$$q_a + \mathcal{L} - q_{b\mathcal{A}}. \tag{2}$$

Similarly, assume OHT a is unloading a FOUP at the unload port of machine $\mathcal{A}$. Inequality (1) and formula (2) still hold after q_a is replaced with s_a. Thus, the possible blocking location will be load/unload ports and the sufficient condition of blocking is relevant to the relative position of OHTs and machines and the loading/unloading time.

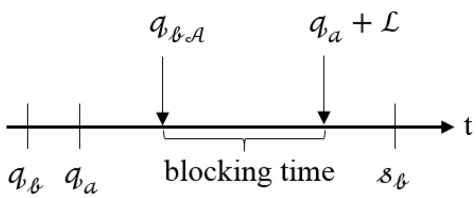

Figure 2: Time sequence of transport jobs.

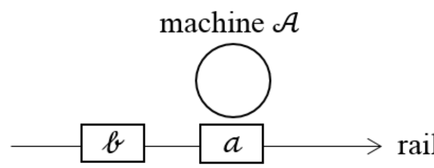

Figure 3: The front OHT is loading/unloading a FOUP.

3 DECENTRALIZED DISPATCHING ALGORITHM

BALI (blocking avoidance by exploiting location information) algorithm, a decentralized dispatching algorithm, is proposed to effectively reduce blocking and increase AMHS efficiency. The core ideas of the BALI algorithm are transport job matching, blocking inference, and job exchange. The BALI algorithm is executed by individual OHTs in empty state and assigned state, as shown in Figure 4. Assume v represents the OHT executing the BALI algorithm. ℓ_v denotes the location of the OHT v. O_v denotes the transport job of v and ℓ_{O_v} denotes the load port location of that transport job. A negative value of ℓ_{O_v} means v has no transport job and thus is available for receiving transport requests. The units of ℓ_v and ℓ_{O_v} are both meter. f denotes the OHT in front of v. The loading/unloading time is $\mathcal{L}$ and the OHT speed is $\mathcal{S}$. The unit of $\mathcal{L}$ is second while that of $\mathcal{S}$ is meter per second. $\mathcal{R}$ denotes the total number of transport requests in the entire manufacturing system. A positive value of $\mathcal{R}$ means at least one transport request exists. Detailed steps of the BALI algorithm are presented as follows:

Step 1. *Transport Job Matching*
    if $\ell_{O_v} < 0$
        if $\mathcal{R} > 0$
            find the nearest machine with transport job requests and update ℓ_{O_v}
        else if $\ell_{O_f} > 0$
            go to Step 2.
        else return to Step 1.
    else if $\ell_{O_f} > 0$
        go to Step 2.
    else return to Step 1.

Step 2. *Blocking Inference*
    if $\frac{\ell_f - \ell_v}{\mathcal{S}} < \mathcal{L}$
        if $\ell_{O_v} > \ell_{O_f}$ or $\ell_{O_v} < 0$
            go to Step 3.
        else return to Step 1.

else return to Step 1.

Step 3. *Job Exchange*
exchange the value of ℓ_{O_v} and ℓ_{O_f}
return to Step 1.

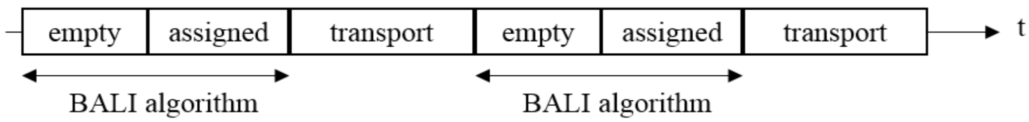

Figure 4: Time periods when the BALI algorithm is executed.

4 NUMERICAL RESULTS AND INSIGHTS

The BALI algorithm and the NJF method have been implemented in Flexsim, a 3D discrete event simulation software, and run on a desktop with 3.4 GHz CPU and 24 GB RAM. The above two dispatching rules are tested and analyzed under four scenarios: high factory loading, medium-high factory loading, medium factory loading, and low factory loading. Each scenario is repeated for 35 times with each repetition possessing different random seeds. The simulation horizon is set to 7 days with warm-up time lasting 24 hours. The numerical values of system parameters are shown in Table 1. The values of the OHT speed, acceleration/deceleration, and loading/unloading time are extracted from (Campbell and Ammenheuser 2000).

Table 1: Numerical value and unit of parameters.

Parameter	Value and Unit	Parameter	Value and Unit
OHT maximum speed	2 m/sec	input buffers of machines	10
OHT acceleration/deceleration	1 m/sec^2	output buffers of machines	1
OHT loading/unloading time	20 sec	inter-arrival time of FOUP entering the manufacturing system	high/medium-high/medium/low factory loading: exponentially distributed with mean 55/60/80/ 100 sec
machine process time	exponentially distributed with mean 50 sec		

The numerical results with statistical significance are shown in Table 2. MCT and SDCT stand for the mean and the standard deviation of cycle time respectively. MWT and SDWT stand for the mean and the standard deviation of FOUP waiting time respectively. MBT stands for mean blocking time and NB represents the number of total blocking. THP and SDTHP stand for the throughput and the standard deviation of the throughput respectively. The results illustrate the value of the BALI algorithm. Confronted with process time uncertainty and high or medium-high factory loading, the BALI algorithm reduces mean cycle time by more than 43% and increases the throughput by more than 13%, as shown through Figures 5 to 8. The significant reduction in mean waiting time, mean blocking time, and the number of total blocking also demonstrates the potential of the BALI algorithm.

Table 2: Numerical results of both dispatching rules in each scenario.

Scenario	Dispatching Rule	MCT	SDCT	MWT	SDWT	MBT	NB	THP	SDTHP
High factory loading	NJF	48648.98	3133.95	37.41	0.12	15.84	21016.86	6228.20	49.03
	BALI	27586.32	2505.68	28.70	0.08	13.87	15238.23	7083.37	37.60
	ratio[1]	-43.30	-20.05	-23.28	-33.33	-12.44	-27.50	13.73	-23.31
Medium-high factory loading	NJF	46641.47	2669.90	37.41	0.12	15.84	21048.09	6236.29	38.30
	BALI	24017.98	3095.89	28.70	0.10	13.88	15321.54	7099.86	35.87
	ratio	-48.51	15.96	-23.28	-16.67	-12.37	-27.21	13.85	-6.34
Medium factory loading	NJF	8449.51	3922.12	38.06	0.15	15.90	22225.66	6267.46	44.60
	BALI	1133.97	79.97	28.08	0.13	14.38	18172.23	6423.37	66.21
	ratio	-86.58	-97.96	-26.22	-13.33	-9.56	-18.24	2.49	48.45
Low factory loading	NJF	938.85	41.07	38.11	0.15	16.07	25991.54	5181.46	53.39
	BALI	711.86	22.78	27.17	0.14	15.08	21201.57	5198.46	72.87
	ratio	-24.18	-44.53	-28.71	-6.67	-6.16	-18.43	0.33	36.49

[1]: $\text{ratio} = \frac{(\text{variable}_{\text{BALI}} - \text{variable}_{\text{NJF}})}{\text{variable}_{\text{NJF}}} \times 100\%$

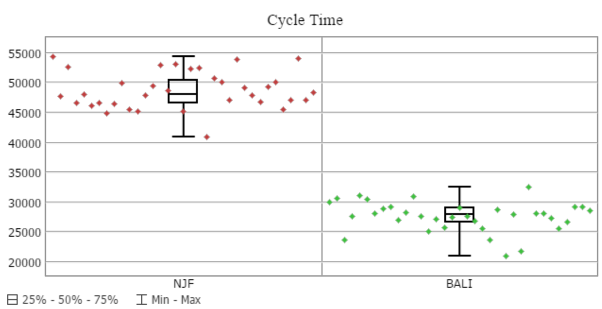

Figure 5: MCT under high factory loading.

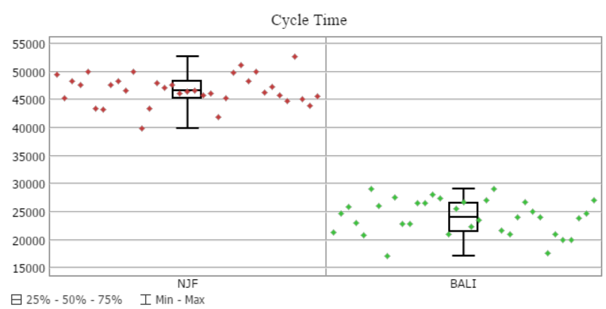

Figure 6: MCT under medium-high factory loading.

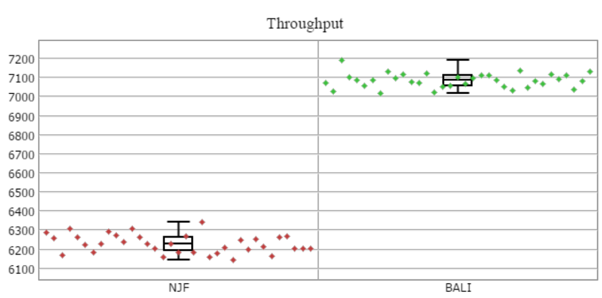

Figure 7: THP under high factory loading.

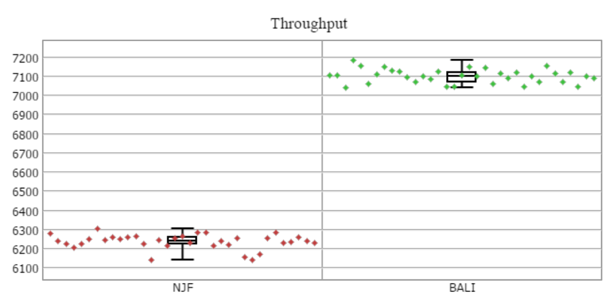

Figure 8: THP under medium-high factory loading.

The above result analysis may give managerial insights to fab managers. Facing process time uncertainty, managers can reduce mean cycle time, increase manufacturing throughput, and meet the increasingly harsher queue time constraints by applying the BALI algorithm. Fab managers can also dynamically alter the dispatching rule of the AMHS, either the NJF method or the BALI algorithm, to dynamically adjust the manufacturing throughput, instead of altering the machines which consumes much time and capital.

5 CONCLUSION

The vehicle dispatching problem in AMHS is NP-hard. Compared with the centralized control method widely applied at modern industry, the proposed decentralized dispatching algorithm exploiting the notion of edge computing and local exchange of neighborhood information through V2V and V2M communications significantly alleviates blocking, reduces cycle time, and increases throughput in production flows. Further improvement is needed to assign the routes of each OHT to prevent possible blocking and enable the decentralized algorithm to identify the optimal dispatching solution.

ACKNOWLEDGMENT

The authors would like to thank Global Logistics & Supply Chain Management Research Group, National Taiwan University of Science and Technology, for offering software and hardware helpful for conducting the simulation experiments in the current study. The authors also thank the anonymous reviewers for their invaluable suggestions greatly improving this paper. This research is supported in part by the Ministry of Science and Technology, Taiwan under grant numbers: 105-2923-E-002 -009 -MY3, 104-2815-C-002-117-E, and 103-2221-E-002 -229 -MY2.

REFERENCES

Applegate, D. L., R. E Bixby, V. Chvatal, and W. J. Cook. 2011. *The Traveling Salesman Problem: A Computational Study.* Princeton University Press.

Bartlett, K., J. Lee, S. Ahmed, G. Nemhauser, J. Sokol, and B. Na. 2014. "Congestion-Aware Dynamic Routing in Automated Material Handling Systems." *Computers & Industrial Engineering* 70: 176-182.

Bellman, R.. 1962. "Dynamic Programming Treatment of the Travelling Salesman Problem." *Journal of the ACM (JACM)* 9: 61-63.

Campbell, E., and J. Ammenheuser. 2000. "300 mm Factory Layout and Material Handling Modeling: Phase II Report." *Tech Transfer Document*.

Huang, H.-W., C.-H. Lu, and L.-C. Fu. 2007. "Lot Dispatching and Scheduling Integrating OHT Traffic Information in the 300mm Wafer Fab." *IEEE International Conference on Automation Science and Engineering 2007:* 495-500.

Miller, C. E., A. W. Tucker, and R. A. Zemlin. 1960. "Integer Programming Formulation of Traveling Salesman Problems." *Journal of the ACM (JACM)* 7: 326-329.

Nishi, T., and Y. Tanaka. 2012. "Petri Net Decomposition Approach for Dispatching and Conflict-Free Routing of Bidirectional Automated Guided Vehicle Systems." *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans* 42: 1230-1243.

Peters, B. A, and T. Yang. 1997. "Integrated Facility Layout and Material Handling System Design in Semiconductor Fabrication Facilities." *IEEE Transactions on Semiconductor Manufacturing* 10: 360-369.

Tung, J., T. Sheen, M. Kao, and C. H. Chen. 2013. "Optimization of AMHS Design for a Semiconductor Foundry Fab by Using Simulation Modeling." In *Proceedings of the 2013 Winter Simulation Conference: Simulation*, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, 3829-3839. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Wang, M.-J. J., H.-C. Chung, and H.-C. Wu. 2004. "Evaluating the 300mm Wafer-Handling Task in Semiconductor Industry." *International Journal of Industrial Ergonomics* 34: 459-66.

Yan, C. -Y. (in Chinese) 2003. "Dynamic OHT Dispatching in a 300mm Intrabay Loop." *Master's thesis*. Graduate Institute of Electrical Engineering, National Taiwan University.

AUTHOR BIOGRAPHIES

YEN-SHAO CHEN is a student at the Department of Electrical Engineering, National Taiwan University, Taipei, TAIWAN. His research interests include optimization theory and algorithms, transportation systems, and security analysis. His email address is ericchenys@gmail.com.

CHENG-HUNG WU is an Associate Professor at the Institute of Industrial Engineering, National Taiwan University, Taipei, TAIWAN. He received his Ph.D. and M.S. degree in industrial and operations engineering from the University of Michigan – Ann Arbor in 2005 and 2006. His research interests include: (1) decisions under uncertainties, particularly in applications such as production/supply-chain system analysis and dynamic control; (2) information and decision support systems (especially operations management and supply chain management); and (3) theoretical work in Markov decision processes and stochastic programming. His email address is wuchn@ntu.edu.tw.

SHI-CHUNG CHANG (S'82-M'83) received a B.S.E.E. from National Taiwan University (NTU) and his M.S. and Ph.D. from the University of Connecticut. Dr. Chang is now a Professor, jointly appointed by the EE Dept., IIE, and GICE, NTU. He was a commissioner of the National Communications Commission, Taiwan, ROC, from 2010–2012. His research interests include system optimization and control, network management, Internet economics, and distributed decision making. His email address is scchangee@ntu.edu.tw.