Journal of the Korean Institute
of Industrial Engineers
Vol.21, No.2, June. 1995

트립에 기초한 물자취급 시스템에서 자재의 평균 체류시간에대한 추정*

Estimation of the Expected Time in System of Trip-Based Material Handling Systems

조 면 식**

Myeon-Sig Cho**

Abstract

We develop an analytical model to estimate the time a workpiece spends in both input and output queues in trip-based material handling systems. The waiting times in the input queues are approximated by M/G/1 queueing system and the waiting times in the output queues are estimated using the method discussed in Bozer, Cho, and Srinivasan [2]. The analytical results are tested via simulation experiment. The result indicates that the analytical model estimates the expected waiting times in both the input and output queues fairly accurately. Furthermore, we observe that a workpiece spends more time waiting for a processor than waiting for a device even if the processors and the devices are equally utilized. It is also noted that the expected waiting time in the output queue with fewer faster devices is shorter than that obtained with multiple slower devices.

1. INTRODUCTION

A manufacturing system consists of three principal components: machines or stations, a material handling system, and a control system. Among manufacturing systems, one of the most challenging system in terms of analysis is the job shop and its variations. In particular,

---

* 이 논문은 1994년도 경기대학교 학술연구 조성비에 의하여 수행되었음.

** Dept. of Industrial Engineering, Kyonggi University.

batch-type metalworking manufacturing operations represent approximately 40% of total manufacturing employment, as noted by Merchant [10]. The author also reports that the average workpiece in a batch-type metal cutting shop spends only about 5% of its time on machine tools. Furthermore, of the time the material is loaded on a machine tool, less than 30% is actual processing time. For the remaining 95% of the time, the workpiece is either moving or waiting.

The time a workpiece spends in manufacturing system consists of four components: waiting time to be processed, processing time, waiting time to be picked up by a handling device, and traveling time. Suppose, early in the design phase, a designer develops a layout for the stations, determines the flow volume and patterns (or device paths) between these stations, and selects the type of material handling system to be used and the appropriate equipment parameters (such as travel speed and load pick-up/deposit times). If such parameters are known, both processing time and traveling time are straightforward to compute. However, waiting times to be processed and to be picked up by a handling device are not straightforward to obtain. Furthermore, it is not known which waiting time, that in the input queue or that in the output queue, contributes significantly toward the time in system. In addition, these waiting times depend on the type of material handling system employed.

Among many different types of material handling systems, we will focus on a trip-based material handling system which consists of one or more self-powered "devices" that are capable of operating independently in an asynchronous fashion. There are many examples of trip-based handling systems. Unit load automated guided vehicles(AGVs), storage/retrieval(S/R) machines in microload automated storage/retrieval systems(AS/RS), lift trucks, numerous manual handling systems, industrial robots, and cranes are good examples of such systems. In general, trip-based handling systems provide flexibility in the movement path.

In trip-based handling systems, each move request in the system is served by one of the devices. In order to serve a move request, a device has to perform a trip. Each trip is composed of empty travel followed by loaded travel. The former accounts for the time it takes for the empty device to travel from its current location to the station which contains the move request while the latter represents the time it takes the loaded device to deliver the loads to its destination. The load travel time also includes the load handling time which consists of a pick-up and deposit operation. On each trip the device typically handles only one load (which represents one move request).

Consider a trip-based material handling system in which devices are used for moving material between stations. (For convenience, we will use a pick and drop AGV system as an example, although the model we develop can be applied to any trip-based handling system.) Such a system is depicted in Figure 1, where input/output(I/O) stations are represented by stations 1 and 2, and processor stations are denoted by stations 3 and 4. There is an input queue and an output queue for each station including the I/O stations. We assume that all the input and output queues have infinite capacity. That is, devices and processors never get blocked. We also assume that there is no significant interaction among the devices. That is, there are no major delays due to congestion at intersections and/or at processor and I/O stations.

Jobs from outside the system enter through one of the I/O stations and-when all the operations have been completed-they exit through one of the I/O stations. Incoming jobs arrive at the output queue of an I/O station while outgoing jobs are deposited at the input queue of an I/O station where they are instantly removed from the system. That is, no processing takes place at the I/O stations.

The oldest job in the output queue of the I/O station is picked by a device and delivered to the input queue of its destination. A device traveling with a job is termed a loaded device. When a loaded device arrives at the input queue of a station, it unloads that job and becomes empty to serve another job in the system. If none of the jobs is ready to be

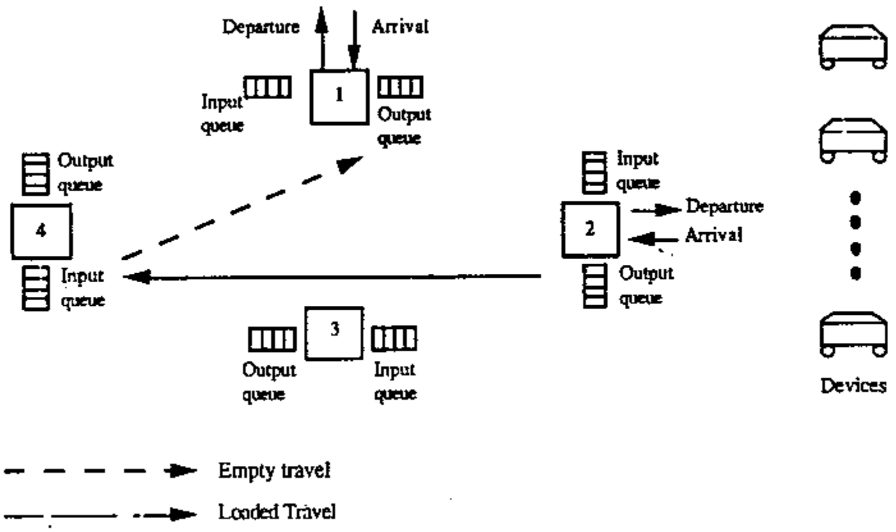


Figure 1. A typical trip-based handling system.

moved, the device becomes and remains idle at the current station. Otherwise, if the next job to be moved is located at, say, station i, the device must first travel without a job, which is termed empty device travel. If there are several jobs ready to be moved, the device must select one of them according to the Modified First-Come-First-Served(MOD FCFS) rule proposed by Srinivasan, Bozer, and Cho [13]. When the empty device arrives at station i, it picks up the job from output queue i and delivers it to its destination, and becomes empty again.

Note that, at certain times, there may be one or more idle devices in the system (which implies that there are no jobs awaiting to be picked up). When a job eventually arrives and finds two or more devices idle, it will select (or "call") one of the devices which has been idle for the longest time period.

A processor station represents either one machine, or group of machines (a cell), or a department.· Jobs to be processed are removed from the corresponding input queue in FCFS order: later, when processing is complete, they are placed in the corresponding output queue without delay.

2. LITERATURE REVIEW

Cho [3] presents an approximate analytical model to estimate throughput capacity and expected waiting times for move requests for an automated guided vehicle systems with single vehicle. To develop a model, they assume that the empty vehicle is dispatched according to the First-Come-First-Served rule, which is simple. However, the performance of this dispatching rule is not as efficient as other known dispatching rules.

Srinivasan, Bozer, and Cho [13] present a general-purpose analytical model to estimate the expected device utilization and the expected station cycle times of a trip-based material handling system. This model is the frist analytical model to explicitly consider an empty device dispatching rule, namely, the MOD FCFS rule which is also employed in this paper. In the MOD FCFS rule, an arriving loaded device first inspects the output queue of the current station. If one or more jobs are waiting, the device is allowed to pick up the one in the front of the queue regardless of its "ages" compared to move requests at other stations. If the output queue of the current station is empty, however, the device serves the oldest move request as the FCFS rule. Therefore, the performance of the MOD FCFS is better than that of the FCFS.

Bozer, Cho, and Srinivasan [2] present an approximate analytical model to estimate the expected waiting times for move requests in the output queues that occur in single-device, trip-based handling systems. They assume that an empty device is dispatched according to the MOD FCFS rule, which performs nearly as good as the Shortest-Travel-Time-First(STTF) rule which is known as one of the best empty device dispatching rules in literature. They show that the single device waiting time model works well as long as the processing time is exponential.

Egbelu and Tanchoco [6] compare a wide set of empty vehicle dispatching rules. In doing so, they define two cases which arise due to the nature of multiple vehicle system: a work center initiated task assignment problem(that is, a job arriving at the output queue selects an idle vehicle) and a vehicle initiated task assignment problem(that is, an empty vehicle selects a waiting job). With infinite output queue capacity, they observe that there are virtually no performance differences among work center initiated task assignment rules and that, among the tested vehicle initiated task assignment rules, a simplified version of FCFS rules maximizes throughput.

In a similar empirical study, Russell and Tanchoco [12] simulate a four machine job shop served by a single vehicle. They compare four vehicle dispatching rules: output queue with largest number of move requests served first, service in random order, FCFS, and STTF. They observe that while the maximum output queue lengths are affected by the vehicle dispatching rules, the throughput capacity (i.e., vehicle utilization) and the mean time a job spends in the system are not affected by the dispatching rules. Although the first observation seems reasonable, the second observation contradicts the results presented by Egbelu [5] and Egbelu and Tanchoco [6], which empirically show that the throughput capacity depends on the vehicle dispatching rule.

Hodgson, King, and Monteith [7] develop a heuristic empty vehicle dispatching rule for a single vehicle system. This dispatching rule, labelled "rule", is based on certain characteristics they observed in an analytical model that was developed for very simple systems (where the maximum number of stations is equal to four and the buffer space for each output queue is limited to one). The "rule" developed by Hodgson et al. is extended by King, Hodgson, and Monteith [8] to a system with two vehicles. For the two vehicle case, it is assumed that there is no interaction between the vehicles. Although both "rule"s are truly dynamic in the sense that the destination of the empty vehicle is reevaluated at every station it passes, three scaling factors are required for reevaluation. (Each scaling factor is determined subjectively.) In both studies, the performance of "rule" is tested against the Vehicle-Looks-For-Work (VLFW) rule, which is equivalent to the STTF rule. The authors empirically observe that "rule" provides shorter mean output queue lengths.

3. Time in System

In this section, we develop an expression to find TS, the expected time a job spends in the system from its arrival at the output queue of an I/O station to its departure from the system. (Recall that once a job is delivered to the input queue of an I/O station, it is assumed to leave the system instantly.) Let TS_i denote the expected time in system for a job entering through I/O station i, $i=1,\cdots,N$, where N is the number of I/O stations. Then TS is represented as follows:

$$TS=\sum_{i=1}^{N} TS_i\,P\{\text{a job enters system through I/O station } i\}$$
$$=\frac{\sum_{i=1}^{N}\lambda_i TS_i}{\sum_{j=1}^{N}\lambda_j} \tag{1}$$

where λ_i is the arrival rate of jobs at the output queue of I/O station i.

Time in system consists of four components: waiting time in the input queue of a processor, processing time on a processor, waiting time in the output queue of either an I/O or a processor station, and traveling time while being transported by the device. Therefore, TS_i is given by:

$$TS_i = \sum_{k=N+1}^{M} NV_{ik}(WI_k+PR_k+WO_k+TR_k) + WO_i+TR_i, \tag{2}$$

where NV_{ik} = expected number of times that a job entering through I/O station i visits processor station k,

WI_k = expected waiting time in the input queue of processor station k,

PR_k = expected processing time on processor k,

WO_k = expected waiting time in the output queue of station k,

TR_k = expected loaded travel time out of station k,

M = number of stations, including the I/O stations.

While the first term on the right hand side represents the time associated with processors, the last two terms represent the time associated with the I/O station through which a job enters the system.

In order to find the value of WI_k, $k=N+1,\cdots,M$, we assume that the arrival process of loaded devices to processor station k is Poisson. Hence, WI_k is obtained from an M/G/1 queueing system. The expected processing time is directly obtained from the input data. The value of WO_k is obtained as discussed in Bozer et al. [2]. Lastly, TR_k is obtained by $\sum_l p_{kl}\,\tau_{kl}$, where p_{kl} is the probability that a job picked up by the device at station k needs to be delivered to station l and τ_{kl} is the mean loaded travel time from station *k* to station *l*. Note that while PR_k and TR_k are given by input data, WI_k and WO_k should be estimated. Therefore, possible sources of error in estimating the expected time in system are the first and third components.

To evaluate TS_l, we next find the value of NV_{lk}, $k=N+1,\cdots,M$, as follows:

$$NV_{ik} = \sum_{j \in v \text{ and } i} P_{jk} NV_{ij} \quad k \in v \qquad (3)$$

where $NV_{ll}=1$ and $NV_{lj} = 0$, $j \neq i$ and $j \in \Omega$, where Ω is a set of I/O stations and v is a set of processor stations. Therefore, substituting the four components and NV_{lk}, $k=N+1,\cdots,M$, into equation (1), we obtain the expected time in system.

In order to obtain WO_k, in Bozer et al. [2], they assumed that there is a single device which serves all the move requests. One possible way to extend the single device model to multiple devices is to use a "single faster device." That is, in order to model a system with K devices, we assume that there is a single device which travels K times faster. Using a proportionally faster single server assumption to represent multiple servers has been used earlier by Dukhovnyy [4] who studied urban transportation systems. Brumelle [1], Newell [11], and Köllerström [9] also used the same approach to approximate the behavior of the G/G/c queueing system.

The single faster device is modeled simply by dividing all the mean loaded travel time from station i to station j, (including pick up time from output queue i and deposit time at input queue j), and mean empty travel time from station i to station j by k. We realize that this approach is not likely to yield satisfactory results for expected waiting times. However, as discussed in Srinivasan et al. [13], such an approximation yields reasonably accurate results for steady state device statistics and cycle times.

4. NUMERICAL RESULTS

4.1 Experimental Framework

There different layouts, labeled L1, L2, and L3, are considered for the experiment. For all three layouts, it is assumed that the device travels at a speed of 15 distance units per minute (except for L3 where it travels at 30 distance units per minute) and that the pickup or deposit time is equal to 1/3 minutes. While devices are allowed to move in only one direction in L1 and L3, devices in L2 are allowed to move in both directions. Both loaded and empty travel times are computed by assuming that the device always follows the shortest path. (The travel time from the input queue to the output queue of a station is assumed to be negligible.)

The first layout, namely, L1, consists of seven stations where stations 1 and 2 are the

I/O stations. Note that no jobs are received through station 2. The mean interarrival time of jobs received through stations 1 is 30 minutes. In L2, the mean interarrival times of jobs received through stations 1, 2, and 4 are 4.9, 9.8, and 14.7 minutes. In L3, the mean interarrival times of jobs received through stations 1, 2, 3, 4, and 5 are 5, 10, 15, 20, 25 minutes. Layouts for L1, L2, and L3 are presented in Cho [3]. Tables 1-6 in Cho [3] also represent the routing matrix and the distance matrix for L1, L2, and L3, respectively.

For each problem, we examined three travel time distributions and two processing time distributions. The device travel time distributions considered are deterministic, uniform with a coefficient of variation equal to 0.4, and exponential. The processing time distributions considered are exponential and uniform with a CV equal to 0.4. For all three problems, the mean processing time at each processor station (shown in Table 1) is set equal to that value which yields an expected processor station utilization of 0.75. Also, the mean processing time that yields an expected station utilization of 0.9 can be obtained from $\rho = \lambda / \mu$.

Table 1. Mean processing time at the processor stations.

Layout and throughput level Station NO.	L1	L2	L3
1	0.0	0.0	0.0
2	0.0	0.0	0.0
3	24.010	0.0	0.0
4	35.934	0.0	0.0
5	34.278	4.339	0.0
6	32.103	4.101	8.556
7	107.010	7.796	9.035
8		4.549	7.787
9		3.924	8.318
10		5.592	5.917
11		5.425	4.463
12			6.695
13			5.866
14			4.434
15			6.015
16			9.829
17			6.011
18			7.043
19			5.393
20			6.737

In order to obtain steady state statistics, we first make a single simulation run starting with an empty system and devices idling at the I/O stations. For "warm-up" purposes, appropriate statistics are cleared after 5,000 loaded trips per device are performed. After this warm-up period, ten observations (i.e., replications) on each measure of performance are recorded. Each observation is based on 5,000 loaded trips per device.

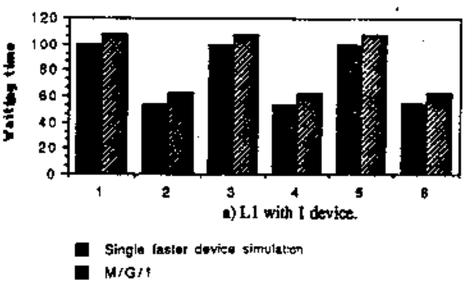
4.2 Experimental Results

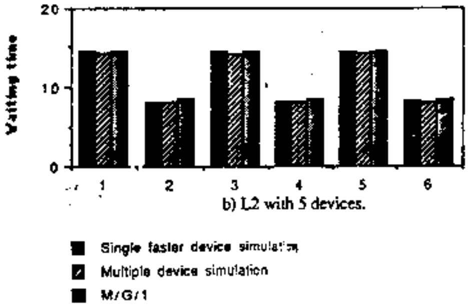
Due to the direct relationship between the expected queue length (level of WIP) and the expected waiting time, we can discuss the WIP in terms of expected waiting times.

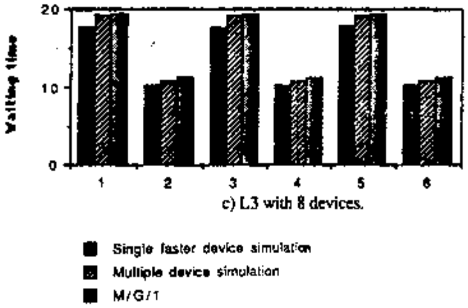
To estimate the expected waiting time in an input queue, we assume, as mentioned earlier, that loaded devices (i.e., jobs) arrive according to a Poisson process. That is, we use an M/G/1 queueing system. Figure 2 depicts the simulated and estimated weighted expected waiting times in all the input queues with various combinations of distributions. This Figure shows that the M/G/1 approximation works quite well although it slightly overestimates the weighted expected waiting time in all the input queues. It also shows that the weighted expected waiting time in the input queue does not depend on the number of devices or the travel time distributions. That is, the weighted expected waiting time depends significantly on the type of the processing distribution. (Similar observations are made for problems where the expected processor utilization is set equal to 0.9)

To estimate the expected waiting time in an output queue with multiple devices, we assume that the expected waiting time is the same regardless of the number of devices. That is, we use the single (faster) device model, discussed in Srinivasan et al. [13], to obtain the expected waiting time in multiple device system.

Table 2 shows the time in system under different processor utilizations obtained from the simulation experiment and the analytical model. (Note that only difference in estimating the expected time in system for a single (faster) device and multiple (slower) devices is the time spent on the devices.) The results show that the difference between the estimated and simulated time in system is less than 10% in most problems.


a) L1 with 1 device.


b) L2 with 5 devices.


c) L3 with 8 devices.

1) Deterministic travel and exponential processing time.
2) Deterministic travel and uniform processing time.
3) Uniform travel and exponential processing time.
4) Uniform travel and uniform processing time.
5) Exponential travel and exponential processing time.
6) Exponential travel and uniform processing time.

Figure 2. Weighted expected waiting time in all the input queues with various combinations of distributions (Processor utilization=0.75)

Table 2. Expected time in system under MOD FCFS obtained from the simulation model and the analytical model(Processor util.=0.75).

	Simulation		Analytical model	
	1 device	single faster	1 device	single faster
Deter/Expon(1)	471.45±35.83(2)	Same as left	494.91 (5.0%)	Same as left
Deter/Unifo	316.99±26.43		353.16 (11.4%)	
Unifo/Expon	472.89±35.68		495.82 (4.8%)	
Unifo/Unifo	318.03±26.69		354.07 (11.3%)	
Expon/Expon	480.13±36.53		500.61 (4.3%)	
Expon/Unifo	325.88±27.09		358.87 (10.1%)	

(1) A/B : A is the travel time distribution and B is the processing time distribution.

(2) 95% confidence interval.

a) L1 with 1 device.

	Simulation		Analytical model	
	5 devices	single faster	5 devices	single faster
Deter/Expon	72.86±2.44	65.70±5.87	71.07 (-2.5%)	64.89 (-1.2%)
Deter/Unifo	53.25±2.03	45.34±3.53	53.39 (0.3%)	47.21 (4.1%)
Unifo/Expon	72.84±2.44	66.02±5.93	71.18 (-2.3%)	65.00 (-1.5%)
Unifo/Unifo	53.40±2.01	45.68±3.50	53.50 (0.2%)	47.32 (3.6%)
Expon/Expon	73.48±2.48	67.32±6.06	71.73 (-2.4%)	65.55 (-2.6%)
Expon/Unifo	54.06±2.23	47.08±3.70	54.05 (-0.0%)	47.87 (1.7%)

b) L2 with 5 devices.

	Simulation		Analytical model	
	8 devices	single faster	8 devices	single faster
Deter/Expon	114.13±5.99	97.79±7.85	110.11 (-3.5%)	103.10 (5.4%)
Deter/Unifo	81.18±1.99	68.15±4.83	79.07 (-2.6%)	72.06 (5.7%)
Unifo/Expon	114.03±6.01	97.98±7.88	110.17 (-3.4%)	103.15 (5.3%)
Unifo/Unifo	81.26±2.01	68.40±4.81	79.13 (-2.6%)	72.12 (5.4%)
Expon/Expon	114.11±6.10	98.81±7.88	110.48 (-3.2%)	103.46 (4.7%)
Expon/Unifo	81.25±1.99	69.16±4.92	79.44 (-2.2%)	72.42 (4.7%)

c) L3 with 8 devices.

Table 2. (Contd.) Expected time in system under MOD FCFS obtained from the simulation model and the analytical model(Processor util.=0.9).

	Simulation		Analytical model	
	1 device	single faster	1 device	single faster
Deter/Expon(1)	1254.36±208.93(2)	Same as left	1394.90 (11.2%)	Same as left
Deter/Unifo	764.45±136.68		884.61 (15.7%)	
Unifo/Expon	1256.23±208.95		1395.82 (11.1%)	
Unifo/Unifo	765.60±135.86		885.52 (15.6%)	
Expon/Expon	1263.54±210.06		1400.60 (10.8%)	
Expon/Unifo	774.57±138.98		890.31 (14.9%)	

(1) A/B : A is the travel time distribution and B is the processing time distribution.

(2) 95% confidence interval.

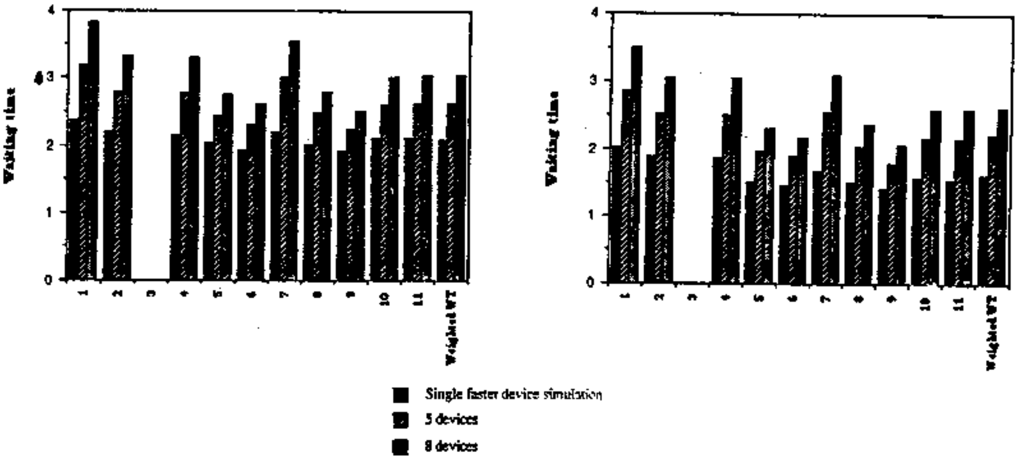
a) L1 with 1 device.

	Simulation		Analytical model	
	5 devices	single faster	5 devices	single faster
Deter/Expon	179.08±13.44	169.10±13.72	183.33 (2.4%)	177.15 (4.8%)
Deter/Unifo	115.72±9.95	105.66±9.88	119.68 (3.4%)	113.50 (7.2%)
Unifo/Expon	179.26±13.39	169.83±13.61	183.43 (2.3%)	177.25 (4.4%)
Unifo/Unifo	115.72± 9.93	106.30±9.83	119.78 (3.5%)	113.61 (6.9%)
Expon/Expon	180.03±13.64	171.14±13.74	183.98 (2.2%)	177.81 (3.9%)
Expon/Unifo	116.42±10.02	107.92±10.06	120.39 (3.4%)	114.16 (5.8%)

b) L2 with 5 devices.

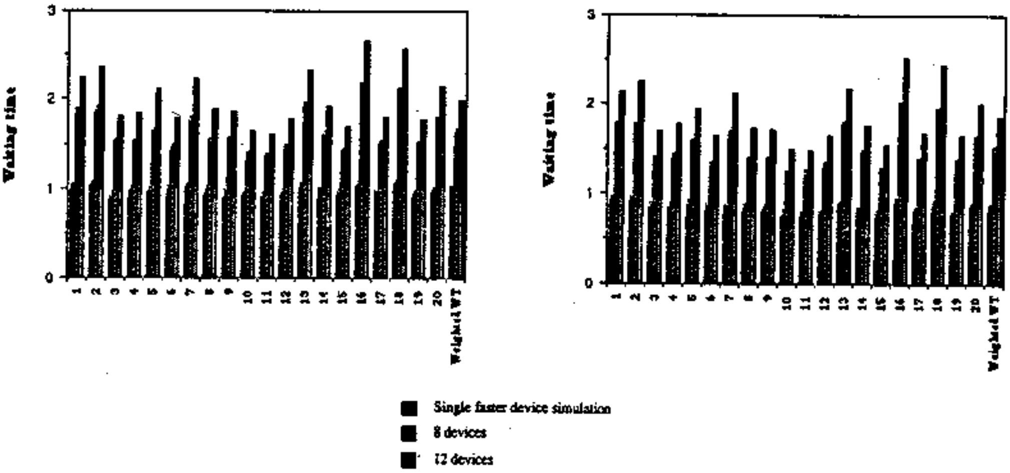
	Simulation		Analytical model	
	8 devices	single faster	8 devices	single faster
Deter/Expon	317.74±37.32	272.16±42.62	307.19 (-3.3%)	300.18 (10.3%)
Deter/Unifo	193.43±18.61	168.03±31.42	195.45 (1.0%)	188.43 (12.1%)
Unifo/Expon	317.67±37.34	272.14±42.67	307.25 (-3.3%)	300.24 (10.3%)
Unifo/Unifo	193.58±18.78	168.20±31.61	195.51 (1.0%)	188.49 (12.1%)
Expon/Expon	317.66±37.14	272.97±42.47	307.56 (-3.2%)	300.54 (10.1%)
Expon/Unifo	193.54±18.68	169.04±31.36	195.81 (1.2%)	188.80 (11.7%)

c) L3 with 8 devices.



a) Uniform travel time and exponential processing time in L2.

b) Uniform travel time and uniform processing time in L2.



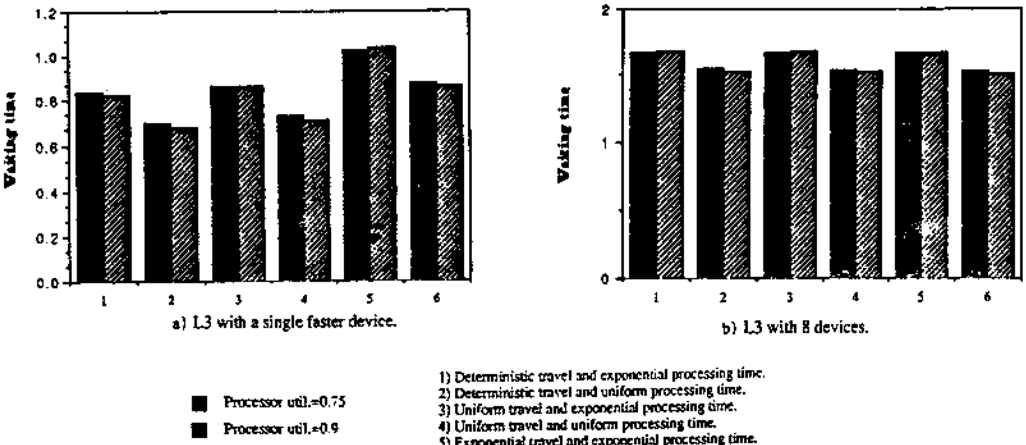
c) Exponential travel time and exponential processing time in L3.

d) Exponential travel time and uniform processing time L3.

Figure 3. Simulated expected waiting times in the output queues under MOD FCFS with different number of devices.

In Table 2, one may observe that the model with multiple (slower) devices appears to work better than the model with a single (faster) device. However, this can be a misleading observation, since the time in system is the sum of four components described earlier. As discussed above, the M/G/1 approximation slightly overestimates the expected waiting time in the input queue. On the other hand, as shown in Figure 3, the single (faster) device model underestimates the expected waiting time in the output queue. Therefore, summing up these components, the model with multiple (slower) devices appears to perform better than the model with a single (faster) device.

Consider next the fraction of time a job spends in each one of the four components once it is released to the system. The results obtained from the analytical model and the simulation experiment are shown in Table 3. Recall that the expected processor utilizations are set equal to 0.75 and the (simulated) expected device utilizations obtained from these problems vary between 0.7 and 0.9 as shown in Srinivasan et al. [13]. From Table 3 we observe that a job spends most of its time in the input queues waiting for the processors. For example, a job spends more than 80% of its time in the system waiting for the processors or being processed (except in L2 with 5 devices). That is, although the expected utilizations of the processors and the devices are comparable as shown above, the processors are responsible for a significant portion of the time a job spends in the system. This observation implies that, as long as the handling system satisfies the throughput requirement, in order to reduce the expected time in system (and WIP), we have to improve only



a) L3 with a single faster device.

b) L3 with 8 devices.

1) Deterministic travel and exponential processing time.
2) Deterministic travel and uniform processing time.
3) Uniform travel and exponential processing time.
4) Uniform travel and uniform processing time.
5) Exponential travel and exponential processing time.
6) Exponential travel and uniform processing time.

Figure 4. Simulated weighted expected waiting times in the output queues under MOD FCFS with different processor utilizations.

Table 3. Breakdown of time in system under MOD FCFS.

Processor util. = 0.75

		L1				L2				L3			
		Simulation		Analytical model		Simulation		Analytical model		Simulation		Analytical model	
	Components	1 device	single faster	1 device	single faster	5 devices	single faster	5 devices	single faster	8 devices	single faster	8 devices	single faster
Deter/Expon(1)	Input queue	66.5%	Same as left	68.2%	Same as left	56.6%	64.3%	59.2%	64.9%	64.7%	69.7%	67.1%	71.1%
	Processor	23.9%		22.7%		19.2%	21.5%	19.7%	21.6%	21.5%	25.2%	22.4%	23.9%
	Output queue	6.8%		6.4%		13.9%	11.9%	10.2%	11.1%	7.1%	4.1%	3.2%	3.5%
	Device	2.8%		2.7%		10.3%	2.3%	10.9%	2.4%	6.8%	1.0%	7.3%	1.0%
Deter/Unifo	Input queue	52.9	,	55.4	,	43.7	52.2	45.7	51.7	51.1	57.8	54.2	59.5
	Processor	35.5		31.9		26.3	31.1	26.3	29.7	30.2	35.9	31.2	34.2
	Output queue	7.6		9.0		15.9	13.4	13.5	15.3	9.2	4.9	4.5	4.9
	Device	4.1		3.8		14.1	3.3	14.5	3.3	9.5	1.4	10.1	1.4
Unifo/Expon	Input queue	66.3	,	68.1	,	56.6	63.9	59.1	64.8	64.6	69.6	67.1	71.6
	Processor	23.9		22.7		19.2	21.4	19.7	21.6	21.5	25.1	22.4	23.9
	Output queue	7.1		6.6		13.9	12.4	10.3	11.3	7.1	4.2	3.3	3.5
	Device	2.8		2.7		10.3	2.3	10.9	2.4	6.8	1.0	7.3	1.0
Unifo/Unifo	Input queue	52.6	,	55.3	,	43.7	51.9	45.6	51.6	51.2	57.7	54.2	59.4
	Processor	35.3		31.8		26.2	30.9	26.2	29.7	30.2	35.8	31.1	34.2
	Output queue	7.9		9.2		16.1	13.9	13.7	15.5	9.1	5.1	4.6	5.0
	Device	4.1		3.8		14.1	3.3	14.4	3.3	9.5	1.4	10.1	1.4
Expon/Expon	Input queue	65.5	,	67.4	,	56.3	63.1	58.7	64.2	64.7	69.1	66.9	71.4
	Processor	23.5		22.5		19.0	21.0	19.6	21.4	21.5	24.9	22.3	23.8
	Output queue	8.3		7.5		14.5	13.7	11.0	12.0	7.1	5.0	3.6	3.8
	Device	2.7		2.7		10.2	2.3	10.8	2.4	6.8	1.0	7.3	1.0
Expon/Unifo	Input queue	51.9	,	54.6	,	43.5	51.0	45.2	51.0	51.2	57.1	54.0	59.2
	Processor	34.5		31.4		25.9	30.0	26.0	29.3	30.2	35.4	31.0	34.0
	Output queue	9.6		10.4		16.7	15.8	14.6	16.5	9.1	6.1	4.9	5.4
	Device	4.0		3.7		13.9	3.2	14.3	3.2	9.5	1.4	10.1	1.4

(1) A/B: A is the travel time distribution and B is the processing time distribution.

the processor stations, i.e., reduce the processing times, not the handling times.

The simulated weighted expected waiting times in the output queues of L3 with different processor utilizations are shown Figure 4 which indicates that the expected waiting times in the output queues are not sensitive to the mean processor utilization. Similar observations are made for L1 and L2. That is, the only components affected by the reduced mean processor utilizations are the first and second components which correspond to the contents of the input queues and the jobs on the processors, respectively.

Form Table 3 we make another observation: the expected time in system depends largely on the processing time distribution. This is due to the fact that the expected waiting time in an input queue depends on the processing time distribution, and the fraction of time a job spends in the input queues is larger than the time it spends elsewhere in the system. Recall that, while the expected times being processed or moved is determined directly form the input data, the expected times in both the input and output queues are estimated analytically or via simulation. The results in Table 3 indicate that the analytical model estimates the expected waiting time levels in both the input and output queues relatively accurately.

5. CONCLUSIONS

We develop an analytical model to estimate the time a workpiece spends in both input and output queues. The analytical results are tested via simulation experiment. The result indicates that the analytical model estimates the expected WIP levels in both the input and output queues relatively accurately.

We observed two interesting phenomena in our study: 1) a job spends more time waiting for a processor than waiting for a device even if the processors and the devices are equally utilized and 2) the expected waiting time in an output queue with fewer (faster) devices is shorter than that obtained with multiple (slower) devices. The first observation implies that, as long as the handling system satisfies the throughput requirement, in order to reduce the expected time in system (and the WIP), we have to improve only the processor system, i.e., reduce the processing times, not the handling times.

The second observation implies that, if possible, we would purchase faster but fewer devices to reduce time spent in the output queues as well as time spent for transportation. Furthermore, a fewer number of devices would generally result in less blocking and congestion delays. Obviously, the final decision must be made on the basis of cost and safety considerations.

BIBLIOGRAPHY

[1] Brumelle, S. L., "Some Inequalities for Parallel Server Queues," *Operations Research*, Vol. 19, pp. 402-413, 1971.

[2] Bozer, Y.A, Cho, M.S., Srinivasan, M.M., "Expected Waiting Times in Single-Device Trip-Based Material Handling System," *European Journal of Operational Research*, Vol. 75, No. 1. pp. 200-216, 1994.

[3] Cho, M.S., "A Performance Evaluation Model of AGV Systems under First-Come-First-Served Rule," Journal of the Korean Institute of Industrial Engineers, Vol. 18, No. 2, pp. 65-87, 1992.

[4] Dukhovnyy, I.M., "An Approximate Model of Motion of Urban Passenger Transportation over Annular Routes," *Engineering Cybernetics*, Vol. 17, No.1, pp. 161-162, 1979.

[5] Egbelu, P.J., "The Use of Non-simulation Approaches in Estimating Vehicle Requirements in an Automated Guided Vehicle Based Transport System," *Material Flow*, Vol. 4, pp. 17-32, 1987.

[6] Egbelu, P.J. and Tanchoco, J.M.A., "Characterization of Automatic Guided Vehicle Dispatching Rules," *International Journal of Production Research*, Vol.22, No.3, pp. 359-374, 1984.

[7] Hodgson, T.J., King, R.E., and Monteith, S.K., "Developing Control Rules for an AGVS using Markov Decision Processes," *Material Flow*, Vol. 4, pp. 85-96, 1987.

[8] King R.E., Hodgson, T.J., and Monteith, S.K., "Extracting Heuristic Control Rules for AGV's using Markov Decision Processes," *Belgian Jour. of OR, Stat. and Comp. Sci.*, Vol. 27, No. 2, pp. 3-17, 1987.

[9] Köllerström, J., "Heavy Traffic Theory for Queues with Several Servers : I," *Journal of Applied Probability*, Vol. 11, pp. 544-552, 1974.

[10] Merchant, M.E., "The Factory of the Future," paper presented to the American Association of Engineering Societies Annual Meeting, Port St. Lucie, Fla, 1982.

[11] Newell, G.F., "Approximate Stochastic Behavior of n-Server Service System with Large n," Lecture Notes in Economics and Mathematical Systems, No. 87, Springer-Verlag, New York, 1973.

[12] Russell, R.S. and Tanchoco, J.M.A., "An Evaluation of Vehicle Dispatching Rules and Their Effect on Shop Performance," Material Flow, Vol. 1, pp. 271-280, 1984.

[13] Srinivasan, .M., Bozer, Y.A., and Cho, M. S., "Trip-Based Material Handling Systems: Throughput Capacity Analysis," *IIE transactions*, Vol. 26, No. 1, pp. 70-89, 1994.