Proceedings of the 2003 IEEE
International Conference on Robotics & Automation
Taipei, Taiwan, September 14-19, 2003

Platform-based AS/RS for Container Storage

Chuanyu CHEN Shell-Ying HUANG
Wen-Jing HSU
Center for Advanced Information Systems,
School of Computer Engineering, Nanyang
Technological University
Nanyang Avenue, Singapore 639798

Ah Cheong TOH Chee Kit LOH
Technology Division, Maritime and Port
Authority of Singapore
460 Alexandra Road, PSA Building #18-00,
Singapore 119963

Abstract—**The ever-increasing demand on container port throughput has mounted the necessity of a more efficient container storage structure, the requirements of which apparently can be met by AS/RS because of its abilities of fully automating material handling and the potentials for delivering high performance. In this study, we present an AS/RS for container storage, develop the travel time models for the proposed system, and confirm its feasibility in terms of throughput performance. A pre-processing scheme for operation scheduling that significantly improves the AS/RS throughput is proposed and evaluated. Additional constraints arising from the container size and weight on AS/RS design and scheduling are taken into considerations, which in effect leads to a platform-based AS/RS.**

I. INTRODUCTION

In a conventional storage yard for containerized cargos, containers are stacked side by side, one on top of another. The major disadvantage of this stacking scheme is that the reshuffling procedure, which incurs additional unproductive moves, has to be performed in order to retrieve a container that is not at the top of the stack. Moreover, due to the strength constraint of containers, the stacking height is limited, which implies a low floor-space utilization. An Automated Storage/Retrieval System (AS/RS) is capable of providing random access to individual storage locations; it is also structurally possible to construct a high AS/RS structure. Existing studies and applications have assured that AS/RS can deliver great improvement in material flow and inventory control with carefully designed scheduling [2], [7], [9], [5], [8], [6], [1]. Based on these observations, we will assess the idea of storing containers in AS/RSs to achieve the significant throughput in a hub container terminal.

Compared with other types of cargos handled by AS/RS in industry, sea containers are specially large in size and weight, which imposes additional mechanical and safety requirements on their handling systems. Stacker cranes used by the conventional AS/RS are not adequate for handling heavy loads at a high turnover rate in the container application. Therefore, a split-platform AS/RS is proposed, where transports of the containers within individual storage aisles are separated into vertical and horizontal movements and handled by different devices, namely the vertical platform (VP) and the horizontal platform (HP), respectively [3]. Fig. 1 gives a schematic view of a standard aisle in this proposed AS/RS. The I/O stations are the interface with the external system that carries loads to/from the AS/RS, and the hand-over stations are the locations where a loaded VP delivers the container to an empty HP or vice versa. The VPs transfer loads in between the I/O stations and the hand-over stations at any tier of the storage racks, whereas the HPs provide the horizontal connection from the hand-over stations to the individual storage cells. Such a system is capable of concurrent operations, that is, the VPs and the HPs can move independently and in parallel. The split HPs and VPs bring the potential benefits of higher handling rate, easier maintenance, and reduced down-time.

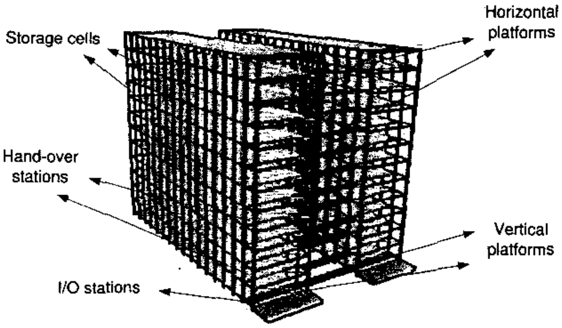


Fig. 1. Structure of the platform-based AS/RS

The storage operations are performed in the platform-based AS/RS as follows; the retrieval operations are just the reversal of the sequence:

1) The VP moves from its dwell position to the I/O station to pick up the load, and lifts it to the hand-over station on the destination tier; meanwhile, the HP at the corresponding tier moves from its dwell position to the hand-over station.
2) The load is transferred from the VP to the HP at the hand-over station.
3) The HP carries the load to the destination storage cell and returns to its dwell position afterwards; at

0-7803-7736-2/03/$17.00 ©2003 IEEE

the same time the VP travels to its dwell position if no new job arrives.

It is obvious that the proposed system works in a completely different fashion than the conventional AS/RS with combined single S/R mechanism. As such, conceptualizing tools are needed to support the physical design and performance feasibility decisions for the platform-based AS/RS. This paper addresses the main issues in AS/RS design and control, i.e., rack configuration design, storage assignment, and interleaving rules. In particular, it develops for the platform-based AS/RS the travel time models that are essential to preliminary rack configuration design, studies and compares two dwell point policies for the platforms, proposes a *pre-processing* scheme for operation scheduling that can significantly improve the throughput performance, and builds a queuing network model for the evaluation of various storage allocation schemes.

The rest of this paper is organized as follows: In Section 2, we will develop discrete travel time models for the system under randomized storage, which take as parameters the basic system attributes similar to those associated with conventional AS/RS. Based on these an optimal layout of the system is proposed, and a dwell point policy is chosen from among several alternatives. Other scheduling rules, i.e., storage interleaving and assignment, will also be briefly discussed in Section 3 and 4, respectively. Concluding remarks are given in Section 5.

II. TRAVEL TIME MODELS FOR PERFORMANCE EVALUATION

Under randomized storage assignment, the probability of accessing any storage cell is identical. Therefore, a straight-forward way of deriving the expected travel time for an AS/RS is to calculate the travel times to individual storage cells and take the average of the sum.

Consider a single rack with B bays and T tiers for an aggregated capacity of $N = T \times B$ cells. As illustrated in Fig. 2, the I/O station is located at the lower left-hand corner, and the hand-over stations are at the 0^{th} bay of each tier. Let (b,t) denote the cell location at bay b and the tier t. We would calculate the travel time for an access to position (x,y) assuming that the location of the immediately preceding operation is (x',y').

By the above definitions, we have the following observations:

Property 1: If d_h denotes the dwell position of the corresponding HP, and d_v denotes that of the VP, then we have $d_v = y'$ and $d_h = x'$ upon completing a storage, and $d_v = 1$, $d_h = 0$ upon a retrieval.

The main assumptions for the analysis are as follows:

1) Randomized storage is used, which means that any empty cell within the rack is equally likely to be selected for storage and any occupied cell is equally likely to be the location for retrieval.

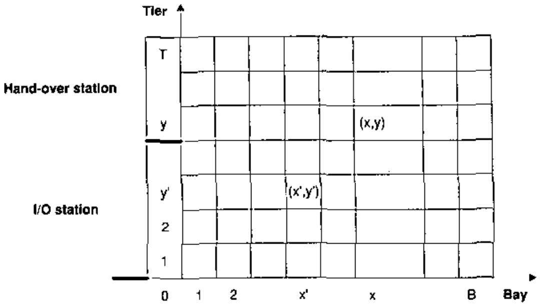


Fig. 2. Definitions of open locations

2) The system operates on *single command* basis for unit loads, i.e., in each operation cycle, either a storage or a retrieval is performed.
3) The dwell point policy for platforms is *stay policy*, i.e., they will stay where they are at the end of each operation until they are required for another operation later.
4) Each request in the infinite sequence of requests is independent of the previous operations, provided that a storage is to an empty cell and a retrieval is from an occupied cell.
5) The VP takes one time unit to travel one tier, and the HP takes one time unit to travel one bay.
6) Pick-up, deposit, and load transfer times associated with load handling are ignored.
7) The acceleration and deceleration of the platforms are ignored. The difference in speed between empty travel and loaded travel is ignored.
8) There is no concurrent movement of platforms for different requests.

A. Storage Time Model

Consider a storage operation in a stable system, its cycle time is affected by the dwell point of the VP and that of the related HP . In an infinite request sequence consisting of both storage and retrieval, any operation except the very first one is preceded by either a storage or a retrieval with equal probability, i.e., 0.5. A storage to cell (x,y) could be decomposed into two phases:

1) The VP moves from tier d_v to tier 1 and after picking up the load, moves to tier y; meanwhile, the HP at tier y moves from d_h to the handover station at $(0,y)$ for the transfer of the load.
2) After the transfer of the load from the VP to the HP, the HP at tier y moves to cell (x,y) while the VP stays at tier y.

Property 2: For each phase, the completion time is decided by the maximum travel time of VP and HP according to the ability of concurrent movement of platforms for a particular operation.

Note that the probability of accessing any cell is equally likely and recall that N denotes the total number of cells in the rack. There are two cases to be considered:

(a) $(x,y) = (x',y')$, with a probability of $\frac{1}{N}$;

(b) $(x,y) \neq (x,y)$, with a probability of $\frac{N-1}{N}$.

The storage time $t_s(x,y)$ to (x,y) should therefore be:

$$t_s(x,y) = \frac{1}{N} \times t_s^a + \frac{N-1}{N} \times t_s^b \qquad (1)$$

where t_s^a is the operation time for case (a) and t_s^b for case (b).

Claim 1: The operation time for case *(a) is:*

$$t_s^a(x,y) = x + (y-1) \qquad (2)$$

For case (b), we should further decompose it into two sub-cases as follows:

(b1) $y = y'$, with a probability of $\frac{B-1}{N-1}$: the preceding operation is to the same tier but a different bay, which can be a storage or a retrieval with equal probability, i.e., 0.5.

(b2) $y \neq y'$, with a probability of $\frac{N-B}{N-1}$: the preceding operation is to a different tier, which can be a storage or a retrieval with equal probability, i.e., 0.5.

The storage time $t_s^b(x,y)$ is therefore represented as:

$$t_s^b(x,y) = \frac{B-1}{N-1} \times t_s^{b1} + \frac{N-B}{N-1} \times t_s^{b2} \qquad (3)$$

where t_s^{b1} is the operation time for case (b1) and t_s^{b2} for case (b2).

Claim 2: The operation time for case *(b1) is:*

$$t_s^{b1}(x,y) = \begin{cases} \frac{1}{B-1}y^2 - \frac{B+6}{2(B-1)}y \\ \quad + \frac{B^2-B+8}{4(B-1)} \\ \quad + x - \frac{max(2(y-1),x)}{2(B-1)} & \textit{if } 2(y-1) < B, \\ \\ x + \frac{3}{2}y - \frac{3}{2} & \textit{otherwise.} \end{cases} \qquad (4)$$

Claim 3: The operation time for case *(b2) is:*

$$t_s^{b2} = \frac{(B+1)(2T-3)}{4B(T-1)}y + x + \frac{(B+1)(T^2-5T+6)}{4B(T-1)} + \frac{1}{4B(T-1)} \sum_{y' \neq y} \sum_{x'' \neq x} (max(y+y'-2,x'') + max(y-1,x'')) \qquad (5)$$

where x'' is the bay location of the previous operation on tier y.

The proofs of these claims involve the application of Properties 1 and 2 (refer to [4] for details). Now we have the expressions of the operation times for all considered cases. In summary, we combine all the equations, with their probabilities as weights, to yield the equation for the storage time to cell (x,y) as follows:

$$t_s(x,y) = \frac{1}{N} \times t_s^a + \frac{B-1}{N} \times t_s^{b1} + \frac{N-B}{N} \times t_s^{b2} \qquad (6)$$

Substituting (2), (4), and (5) into (6), the final expression of the storage time could be obtained [4]. In application, it is more convenient to calculate (2), (4), and (5) individually, and substitute the results into (6), instead of working with the long expression.

Moreover, the overall average storage time of the rack can be derived in a straightforward manner:

$$t_s = \frac{1}{T \cdot B} \sum_{1 \leqslant x \leqslant B} \sum_{1 \leqslant y \leqslant T} t_s(x,y) \qquad (7)$$

This is useful for estimating the storage performance for the preliminary evaluation of AS/RS design configurations.

B. Retrieval Time Model

The algorithm for calculating storage time is applicable to developing retrieval time model. The retrieval operation could be decomposed into two phases:

1) The VP moves from tier d_v to tier y; at the same time, the HP at tier y moves from d_h to bay x to retrieve the load, and then moves to the handover station for the transfer of the load.
2) After the transfer, the VP at tier y moves to the I/O station to deliver the load.

Property 1 and 2 hold, and the possible cases are similar to those of a storage operation. Therefore the processes for deriving the retrieval time model are simply the same as the previous section. In particular, the retrieval time to cell (x,y) can be represented as the combination of various cases with different probabilities:

$$t_r(x,y) = \frac{1}{N} \times t_r^a + \frac{B-1}{N} \times t_r^{b1} + \frac{N-B}{N} \times t_r^{b2} \qquad (8)$$

The corresponding expressions and their proofs are given in [4], which take similar processes as those for deriving the storage time model. The overall average retrieval time of the rack can be obtained by summing up the retrieval times from all the cells and taking average over the number of cells.

TABLE I

DETAILS OF DIFFERENT RACK CONFIGURATIONS

Config -uration	Layout 1		Shape factor	Layout 2		Shape factor
	Tiers	Bays		Tiers	Bays	
1	3	48	0.125	4	72	0.111
2	6	24	0.500	8	36	0.444
3	9	16	1.125	12	24	1.000
4	12	12	2.000	16	18	1.778
5	NA	NA	NA	18	16	2.250
6	16	9	3.556	24	12	4.000
7	24	6	8.000	36	8	9.000
8	48	3	32.000	72	4	36.000

TABLE II

COMPARISONS OF DISCRETE MODELS WITH SIMULATION RESULTS (LAYOUT 1)

Config -uration	Simulation results	Model results	Errors (% Dev)
1	95.023	94.980	0.046
2	56.845	56.966	0.212
3	50.944	51.050	0.208
4	54.447	54.482	0.064
5	NA	NA	NA
6	64.524	64.752	0.354
7	91.341	91.246	0.105
8	178.929	178.639	0.162

TABLE III

COMPARISONS OF DISCRETE MODELS WITH SIMULATION RESULTS (LAYOUT 2)

Config -uration	Simulation results	Model results	Errors (% Dev)
1	141.912	141.975	0.044
2	83.483	83.723	0.288
3	73.001	73.085	0.116
4	75.994	76.202	0.273
5	80.174	80.203	0.036
6	96.935	96.865	0.072
7	137.421	137.527	0.078
8	269.564	269.303	0.097

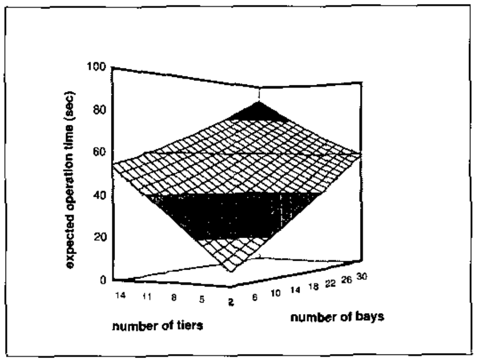


Fig. 3. Variation of operation time

C. Verification and Discussions:

In this section, we examine the derived operation time models and discuss their applications. Note that although the travel time model is done on a single rack with one VP, the result is also valid for a dual-rack structure with two VPs and two I/O stations as long as the sequence of access operations are carefully planned such that the two VPs do not compete for the shared HPs at the same time.

The evaluations of the models are done by comparing them with computer simulation results. 100,000 operations (which is considerably large compared with the number of cells in an AS/RS rack) were executed in each experiment to simulate the infinite sequence of operations. For each operation, the probability that the preceding operation is a storage is set to be the same as that of a retrieval, i.e., 0.5. The initial status of the rack is unknown, which means that if a cell is selected for the first time as the destination of an operation, it can be executed regardless of its operation type. In order to test the sensitivity of the models to the rack configurations, the experiments are done based on two different racks, one with 144 cells and the other with 288 cells. Table I gives the details of the two rack layouts. The outputs of the models and the simulation results are summarized in Table II and III.

Note: % Dev refers to % deviation which was used to indicate the accuracy of the models and calculated as:

$$\%Dev = \frac{modelResults - simulationResults}{simulationResults} \times 100\%$$

The results show that for sufficiently long sequences of accesses, the discrete models can accurately describe the operation times for various rack configurations. The average error is at an acceptable level of 0.14% with the maximum error not exceeding 0.4%, insensitive to the discreteness of the rack.

The travel time models are useful for the AS/RS rack configuration design. For a desired throughput requirement, we can obtain a set of possible rack configurations that yield a corresponding expected operation time from the models. This is reduced to looking for a line on a surface as shown in Fig. 3. Together with other constraints, such as shape factor, we can preliminarily decide an optimal rack configuration. On the other hand, for a given rack configuration, the result suggests a class-based storage assignment policy. For instance, Fig. 4 illustrates possible class boundaries of the zones based on the operation times for a 12-by-28 rack.

We can also use the model to estimate the optimal shape factor *b*, which represents the shape of storage rack in

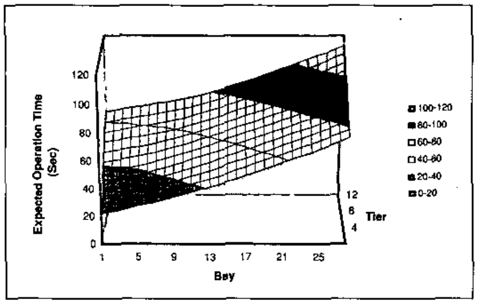


Fig. 4. Zoning of rack by operation time

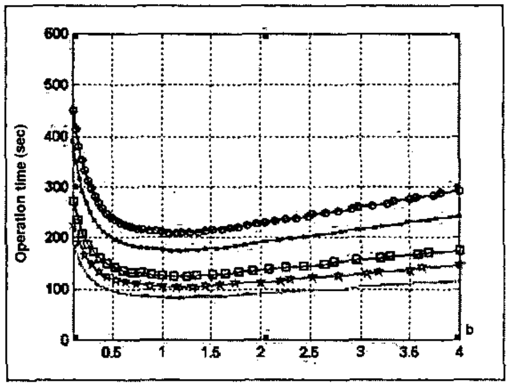


Fig. 5. Influence of shape factor on operation time

terms of time, i.e.,

$$b=\frac{\text{travel time to the highest tier}}{\text{travel time to the farthest bay}}$$

. An experiment was done which varies b from $[0.1,4]$ in five rack configurations to investigate its impact on the operation time and to find out an optimal b. The results are illustrated in Fig. 5.

D. An Alternative Dwell Point Policy

In this section we consider another dwell point policy for comparisons. Under this *return* policy, the VP returns to the I/O station and the HP returns to the handover station upon finishing a job. We will build the discrete time models and compare its efficiency with the previous policy in terms of expected operation time. In this case, the dwell point of VP is $d_v=1$, and that of HP is $d_v=0$. Therefore, from Property 2, both of the storage time and retrieval time can be represented by:

$$t_s=t_r=y+max(y,2x)$$

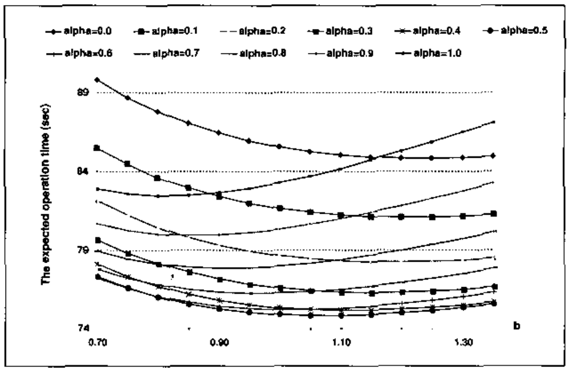


Fig. 6. Impacts of b and α on expected operation times

where (x,y) denotes the destination of current operation.

Hence, regardless of the fraction of storage operations, the operation time to cell (x,y) is:

$$t_o(x,y)=y+max(y,2x)$$

The expected operation time of a discrete rack can be henceforth calculated as:

$$E(t_o)_{dis}=\frac{1}{T\cdot B}\sum_{y=2}^{T}\sum_{x=1}^{B}(y+max(y,2x)) \tag{9}$$

Note that the VP is needed only for tier 2 and tiers above, therefore in the summation y value starts from 2. The discrete expression is the accurate calculation of the expected operation time.

In order to make the two dwell point policies comparable, we have purposely used the same rack layout with 288 storage cells for experiments. In dwell policy 1, we found that the expected operation time has a minimum value of $E(t_o)^1_{min}=74.781$seconds when $b=1.05$. In dwell point policy 2, the minimum value of expected operation time was found to be $E(t_o)^2_{min}=84.773$seconds when $b=1.236$. It can be concluded that policy 1 outperforms policy 2 for a stable system where the fractions of storage and retrieval operations are identical. However, the expected operation time varies with the fraction of storages under policy 1, in some instances it exceeds 84.773 seconds (Fig. 6). Therefore, the optimal rack design will largely depend on the characteristics of the demand in our application. From a long-term point of view, the container yard has balanced work-flow which suggests that policy 1 be more preferable.

III. STORAGE SCHEDULING

On scheduling of incoming requests, due to the constraint that the re-ordering of job sequences will complicate the traffic system in a container terminal, we proposed a *preprocessing scheme* to pre-fetch containers for incoming retrieval requests (or pre-positioning the HPs for storages) to the handover stations and thus reduces the cycle time. In the current design, on each tier of an AS/RS aisle, there is one horizontal platform, whereas there is only one vertical platform serving the entire rack. Therefore, most of the time the horizontal platforms would be idle. If the horizontal platforms can pre-process jobs according to the incoming job sequence such that the VP waiting time for HPs is minimized, the handling time for each operation would be significantly reduced. A dispatching algorithm that enables both HPs and VP to detect their jobs and pre-respond accordingly before the jobs arrive is presented as follows.

```
For $j = 1$ to $k$ ($k$ is the number of jobs in the sequence)
    Step 1. Scan the incoming jobs using a time window to collect jobs to be pre-processed.
    Step 2. For all jobs in a same batch, execute the following:
        (a) For retrieval job, the HP pre-fetches the requested container and waits at the handover stations whenever it is free after finishing the previous job; the container will then be transferred to the VP, lowered down and roll into the I/O station whenever the VP is free after finishing its previous job.
        (b) For storage job, the HP is reserved at the transfer point and the VP at ground tier whenever they are available after finishing their previous jobs.
    Step 3. Advance time window and go to Step 1.
Endfor
```

To examine how the scheduling scheme improves the throughput performance, we ran an example of 10,000 randomly generated jobs with and without pre-processing and collected statistics data on job handling time for analysis. The inter-arrival time of jobs followed Poisson distribution with different mean values varying from 30 seconds to 240 seconds.

As shown in Fig. 7, the improvement on job handling time is significant when the schedule of jobs is not too tight. The results confirm that the pre-processing scheme for AS/RS optimization is effective. Note that the improvement in throughput might change with different job sequences. If the destinations of incoming requests are distributed more evenly, we can expect to see even higher improvement. Particularly, for a given job inter-arrival distribution, the best case for the proposed scheme is that individual jobs in the time window are located at different tiers, where the horizontal platforms could concurrently work to the maximum extent. In this case, if the mean value of the inter-arrival distribution is larger than the time it takes for the platforms to pre-process the jobs, the AS/RS response time to each job would be the cycle time of the yard crane that transfers load between the I/O station and the prime movers.

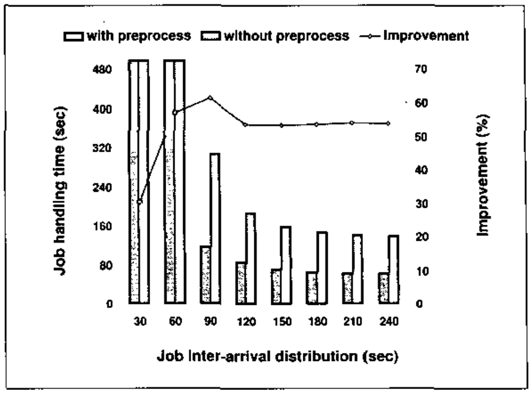
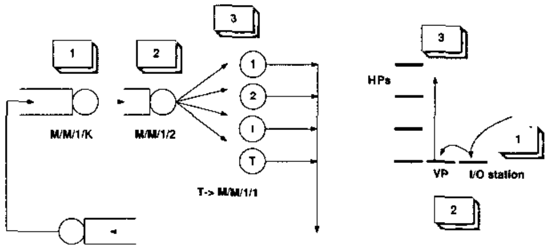


Fig. 7. Improvement on job handling time



Proposed queuing model for the system

Schematic diagram of the system

Fig. 8. The queuing network model for the system

IV. STORAGE ALLOCATION: ON-GOING WORK

The algorithms for estimating the operation time are also essential to the evaluation of storage schemes that allocate open locations in the rack to incoming items. This is because the behaviors of the platforms, which are decided by the storage assignment schemes, determine the service time of any incoming requests. We can represent the AS/RS rack in terms of a queuing network (Fig. 8) to highlight the impacts of storage schemes on the system performance as follows:

In the network, the system is decomposed into three queuing models. The arrivals and departures of the requests at the I/O station are represented by an $M/M/1/K$ queue where K indicates the available queuing spaces for the vehicles, the service rate is decided by the crane cycle time which can be derived from crane specifications. The departure of this queue is in effect the arrival of the second queue representing the VP behaviors. Note that the I/O station can be considered as a queuing space for the VP, therefore an $M/M/1/2$ queue is used. In the third model, the HPs are represented as T independent servers whose arrival rates are decided by the storage scheme and the service time of VP. Solving the queuing network with classical queuing theories, we could find an optimal storage scheme to serve a predefined arrival rate of vehicles.

V. SUMMARY

In this study we have proposed a platform-based AS/RS for container storage. The proposed system deploys platforms to transfer containers for safety and efficiency considerations. The mechanical and economical feasibilities of such a structure have been proven by similar systems applied in air-cargo handling. In order to evaluate the throughput feasibility of the system, we have developed discrete travel time models to preliminarily estimate the expected operation time. Two dwell point policies were evaluated and discussed, namely the *stay* policy and the *return* policy.

For the stay policy, we first gave the expression of the expected operation time, which was validated by simulation results. The discrete models can accurately describe the platform behaviors and they are computationally feasible. However, the expressions are rather complicated and to some extent, not practical for quick estimation of the operation time. To avoid this drawback, we also approximated the discrete rack as a continuous pick face to simplify the expression for the operation time. It was observed that the global minimum of the expected operation time is obtained when the fraction of storages in the operation sequence is 0.5 and the shape factor is around 1.05.

For the return policy, the discrete expression is straightforward and simple, due to the fact that both storage and retrieval operations will start from a known position. It appears that the rack has the minimum expected operation time when the shape factor is about 1.236.

Comparing the two policies, we found that for a balanced system, i.e., the inbound work-flow is equal to the outbound work-flow, the stay policy outperforms the return policy. However, during certain time slots when the system works under an unbalanced situation, the latter policy yields better throughput performance.

To further improve the system throughput, a *pre-processing* approach was investigated and was proved to be effective. Moreover, we proposed a queuing network to model the system that is useful for deciding on the size of the buffers, the number of platforms, and the optimal storage scheme.

For future studies, it would be most helpful to work out the queuing network. Other approaches for interleaving the incoming requests will also be essential to optimize the system.

VI. ACKNOWLEDGMENTS

This project is sponsored by ASTAR (Agency for Science, Technology, and Research), MPA (Maritime and Port Authority), and NTU (Nanyang Technological University). The authors would like to thank the staff of MPA, Dr. Song Tiancheng, Mr. Ong Ah Kiong, Mr. Yow Liang Keon, and Mr. Chan Chi Fei, for their valuable input.

VII. REFERENCES

[1] J.P.V.D. Berg, "Simulation study of an automated storage/retrieval system", *International Journal of Production Research*, vol. 38, no. 6, 2000, pp. 1339-1356.

[2] Y.A. Bozer, "Travel-time models for automated storge/retrieval systems", *IIE Trans.*, pp. 329-338, 1984.

[3] C.Y. Chen, W.J. Hsu, V.Y. Vee, P. Lu, and S.Y. Huang, "Automated storage and retrieval systems for container operation", *CAIS Technical Reports*, no. TR-HCTS02, 2001, pp. 1-16.

[4] C.Y. Chen, S.Y. Huang, W.J. Hsu, and Y.H. Hu, "AS/RS operation models for container terminals", *CAIS Technical Reports*, no. TR-HCTS08, 2002, pp. 1-17.

[5] P. Egbelu, "Framework for dynamic positioning of storage/retrieval machines in an automated storage/retrieval system", *International Journal of Production Research*, vol. 29, no. 1, 1991, pp. 17-37.

[6] M. Han, L. McGinnis, J. Shieh, and J. White, "On sequencing retrievals in an automated storage/retrieval system", *IIE Trans.*, vol. 19, no. 3, 1987, pp. 56-66.

[7] W.H. Hausman, L.B. Schwarz, and S.C. Graves, "Optimal storage assignment in automatic warehousing systems", *Management Science*, vol. 22, no. 6, 1976, pp. 629-638.

[8] B. Park, "Optimal dwell point policies for automated storage/retrieval systems with dedicated storage", *IIE Trans.*, vol. 31, 1999, pp. 1011-1013.

[9] M.J. Rosenblatt and A. Eynan, "Deriving the optimal boundaries for class-based automatic storage/retrieval systems", *Management Science*, vol. 35, no. 12, 1989, pp. 1519-1524.