Setting Real Time WIP Levels in Production Lines

Yang R.L., V. Subramaniam and S.B. Gershwin

Abstract— A new algorithm is developed for setting WIP level in production lines. It is a pull policy that determines when to authorize a job to be processed. Dynamic information in the system such as machine failures and repairs, and demand variations are taken into consideration to provide an efficient method to dynamically determine the WIP level. Two types of production lines are examined in this paper. The first is production lines which have few random interruptions and the second production line is exposed to frequent interruptions. Experiment results from these two types of production lines illustrate that the new algorithm generates higher service level with lower WIP compared to the CONWIP policy. In addition, the superiority of the new algorithm is more significant when the production line is exposed to frequent random interruptions.

Index Terms —WIP, production line, service level

I. INTRODUCTION

Control policies have great impacts on the management of the production systems. *Push* and *pull* policies are classified based on the way how production starts. Push policies schedule the production, while pull policies authorize it. Accordingly production systems which are controlled by these policies are referred to as push and pull systems respectively. In a pull system, the start of a job is triggered by the completion of another. The superiorities of pull over push systems have been studied in [1]-[3] and may be summarized as: (i) The reduced WIP in pull systems makes any environmental problem more noticeable than in the push systems, which facilitates the environment improvement. (ii) In terms of efficiency, the pull systems provide the same production rate with lower WIP level than push systems. (iii) It is evident that the variance of the flow time is less in a pull system than in an equivalent push system because the pull systems are able to regulate the fluctuation of the WIP level while push system are not able to. (iv) Pull systems are more robust than push systems because controlling WIP is more robust than throughput [4].

Popular examples of the push system are the material requirement planning (MRP) [5] approach and its successor manufacturing resource planning (MRPII). Problems are found with the MRP system may be briefly as: (i) MRP does not always generate flexible plans and chaos is resulted when it is late and (ii) MRP does not make decisions based on real capacity utilization. Detailed cases of pull systems may be referred to the literature review part. In this paper, we will provide a new algorithm to set the WIP level in a pull system and illustrate how it is able to adjust the WIP in response to changes in the production environment.

This paper is organized as follows: following this introduction, a literature review of various pull policies is provided in Section 2. Subsequently, the new control algorithm is described in Section 3. In section 4, a simulation study and the analysis of the results are presented. Finally, this paper concludes with a summary of the key findings and future work.

II. LIETERATURE REVIEW

There are numerous of research on the pull systems. Basestock policy is one classical example and it may be referred to [6]. It limits the amount of WIP between the production stage and the demand process. Each machine attempts to maintain a certain amount of WIP in its output buffer and this amount is called the basestock level [7] of the machine. Different basestock level may be achieved by setting different initialization card to various machines. However, the initialization is usually based on the forecasting of the demand and estimated production time times certain safety factor. It is liable to be affected by expert's experience and preference, and is not able to adapt to random interruptions and changing production environments.

Kanban is most known as the Toyota production system and it means card in Japanese. One survey paper on Kanban is [8]. [9] describes the Kanban in detail and compares it with push system by using performance measure of variance amplification. Minimal blocking, another Kanban model is started by [10] and it is stated that the minimal blocking model 'may be different from operation concept of a typical pull system'. In a minimal blocking system, the total inventory of a cell consists of a machine and its immediate upstream and downstream

(No page numberings)

buffers. In [11] opposed the findings of [2] by showing that the distribution of Kanbans among the stations has a significant impact on the system performance. [12] studies multiple part type Kanban systems by using a Brownian approximation model and the total number of Kanban that are shared by different part types is limited. Later on their research has been extended by [13] and in Chevalier and Wein's approach, the Kanban is not controlled directly. Other multiple Kanban research which studies the difference between decicated Kanban and shared Kanban (or hybrid Kanban) may refer to [14] - [16].

Another pull policy is the constant work in process (CONWIP) policy which is proposed in [1]. In a CONWIP system, new jobs are not allowed to enter the system unless the total number of jobs in the system drops below certain given level. The CONWIP policy controls production systems through CONWIP loops and a CONWIP loop consists of at least one machine and a buffer, and a preset WIP level. A CONWIP system may consist of more than one CONWIP loop, especially there exists multiple part types or multiple routings in the system. Comparison of the CONWIP with other pull policies is studied in [2]. [17] proposed a hybrid Kanban-CONWIP policy and compared it with CONWIP, Kanban and other policies based on a four machine production line by using performance measure of service level and the amount of WIP in the system. In addition, a simple adaptive production control method which is named as statistical throughput control (STC) for setting WIP level is proposed in [18] under CONWIP protocol.

III.SETTING WIP LEVELS FOR PRODUCTION LINES

Although studies show that it performs well compared to other pull policies [1]-[3], it is not flawless. The CONWIP policy requires one or more preset values. Refer to [19] for Multiple CONWIP loop systems. Heuristically, the more CONWIP loops are in the system, the more preset values are required. These preset values are based on experts' experience and preference. Moreover, when there is a fluctuation in the demand in the system, the preset value may become not accurate and is not able to reflect the true number of kanbans required in the system. Once this value is fixed, it is not changed until next time that system is examined. However, in real situation, there are many random interruptions such machine failures and repairs occurring in the system, while the decision making procedure of the CONWIP does not take into account all of these dynamic status. In this paper, we propose a new WIP setting algorithm for production lines and it is able to address the insufficiencies of the CONWIP.

The primary difference between CONWIP and Kanban systems is that the CONWIP *pulls* jobs into the front of the first machine in the production line and then *pushes* them between machines elsewhere in the line, while Kanban policy pulls jobs between all machines. CONWIP policy usually requires less preset WIP vales compared to Kanban. Therefore, in this paper we compare our new WIP setting algorithm with the CONWIP policy.

A production line is illustrated in Figure 1. There are n machines in this system and demand is satisfied from the finished goods (FG) in the last buffer. Our new algorithm pulls jobs at each machine by using a dynamically determined WIP level. A downstream system of a particular machine is defined as the part of the system that begins with this machine and terminates at the FG buffer. Virtual inventory of a machine is defined as the amount of WIP that in the corresponding downstream system and it is noted as ϕ_i, where i is the index of the machine. $DWIP_i$ is noted as the dynamically determined WIP level for machine i. The heuristic of this new algorithm is that machine i is permitted to pull jobs to process when $\phi_i <$ $DWIP_i$. In the following subsections, we will describe how to dynamically determine the $DWIP_i$ value.

$DWIP_i$ consists of two terms, namely, average demand term and demand variation term. The first term takes into account the average demand (AVG) and lead time (L_i, where i is the index of machines) which is defined as the estimated time between the jobs enters a particular machine and it reaches the FG buffer. L_i considers machine repair time when there is a failed machine in the downstream system. Our assumption allows more than one machine in the downstream system of M_i to fail at certain moment. The demand variation term utilizes lead time L_i in addition to the standard variation of the demand.

IV.SIMULATOIN RESULTS AND ANALYSIS

The purpose of this section is to study the characteristics of the new algorithm which is able to dynamically setting WIP for production lines. In this paper, the machines in the tandem line are assumed to have exponentially distributed times of failures and repairs. Each machine is assumed to process a part with mean μ minutes and a standard deviation of σ minutes, and it is modeled using a lognormal distribution with parameters μ and σ. The demand for parts is modeled using an exponential distribution. A simple four-machine-one-part-type flow line as illustrated in Figure 2 is studied in this paper. Machines M1, M2, M3 and M4 are the four machines in the production line. $DWIP_1$, $DWIP_2$, $DWIP_3$ and $DWIP_4$ are the four DWIP values to be dynamically determined in real time. Experiment parameters are listed in Table 1.

In this paper we study two types of production lines. The first is production lines which has few random interruptions and we set MTTF=100 and MTTR=1 respectively. The second is a production line which is

exposed to more random interruptions and we model it by setting MTTF=50 and MTTR=2 respectively.

Figure 3 and 4 presents the experiment results which illustrate that the new algorithm always provides higher service level with lower WIP in the system. From Figure 3 we may observe that by keeping the same amount of WIP level, the new algorithm is able to generate a service level which is 0.5% higher than the CONWIP. For example, when the average WIP remains at 7, the generated service level by the new algorithm is 99.6% while the CONWIP achieves 99.1%. Although this is not a big different, it counts much when WIP in very expensive and keeping WIP is very costly while the service level required by customers is very high.

From Figure 4 we may observe that when the production line is exposed to more frequent random interruptions, the advantages of the new algorithm is more obvious. For example, to achieve service level of 98%, the new algorithm requires WIP level of 12, while the CONWIP is not able to provide same service level even when the WIP is increased to 14. This may be interpreted by the characteristics of the new algorithm. It adjusts the WIP level dynamically based on the real time system status. Every time when an available machine makes decisions whether or not to pull a job, it looks compares virtual inventory of the downstream system and the DWIP value which are determined in real time. Essentially this adds a feedback to system for decision making and provides an efficient control to the system.

V.CONCLUSIONS

In this paper we propose a new algorithm for setting WIP levels in production lines. This algorithm is a pull policy and it takes into account the dynamics of the system for decision makings. Experiment results illustrate that it outperforms the CONWIP policy by providing higher service level with lower average WIP level. The advantages of the new algorithm are more obvious when the production line is exposed to more frequent random interruptions. This new algorithm is still in infancy. Future research will be on its application to reentrant or multiple part type systems. Bottleneck cases will also be studied.

REFERENCES

[1] Spearman M.L., D.L. Woodruff, and W.J. Hopp, "CONWIP: a pull alternative to kanban", International journal of production research, Vol. 28, No. 5, 879-894, 1990.
[2] Spearman M.L. and M.A. Zananis, "Push and pull production systems: issues and comparisons", Operations research, Vol. 40, No. 3, 521-532, 1992.
[3] Hopp W.J. and M.L. Spearman, "Factory physics: foundations of manufacturing management", Burr Rigfe, IL: Irwin, 1996.
[4] Spearman M.L. and M.A. Zananis, "push and pull production systems: Issues and comparisons", Technical report 88-24, Department of industrial engineering and management sciences, Northwestern University, Evanston, IL, 1988.
[5] Vollmann T.E., W.L. Berry and D.C. Whybark, "Manufacturing planning and control systems", Homewood, IL: Richard D. Irwin, 1984.
[6] Clark, A.J. and H. Scarf, "Optimal policies for the multi-echelon inventory problem", Management science, Vol. 6, No. 4, 475-490, 1960.
[7] Kimball G., " General principles of inventory control", Journal of manufacturing and operations management, Vol. 1, No. 1, 119-130, 1988.
[8] Berkley B.J., "A review of the kanban production control research literature", Production and operations management, Vol. 1, No. 4, 393-411, 1992.
[9] Kimura O. and H. Terada, "Design and analysis of pull system, a method of multi-stage production control", International journal of production research, Vol. 19, No. 3, 241-253, 1981.
[10] So K.C and S.C Pinault, "Allocation buffer storages in a pull system", International journal of production research, Vol. 15, no. 2, 1959-1980, 1988.
[11] Gstettner S. and H. Kuhn, "Analysis of production control systems kanban and CONWIP", International journal of production research, Vol. 34, 3253-3273, 1996.
[12] Harrison J.M and L.M. Wein, "Scheduling networks of queues: heavy traffic analysis of a two-station closed network", Operations research, Vol. 38, 1052-1064, 1990.
[13] Chevalier, P.B. and L.M Wein, "Scheduling network of queues: heavy traffic analysis of a multistation closed network", Operations research, Vol. 41, 743-757, 1993.
[14] Buzacott J.A and G.J. Shanthikumar, "Stochastic models of manufacturing systems", Prenntice Hall, 1993.
[15] Duenyas I., "A simple release policy for networks of queues with controllable inputs", Operations research, Vol. 42, 1162-1171, 1994.
[16] Framinan J.M., R. Ruiz-Usano and R. Leisten, "Input control and dispatching rules in a dynamic CONWIP flow-shop", International journal of production research, Vol. 38, 4589-4598, 2000.
[17] Bonvik A.M., C.E. Couch and S.B. Gershwin, "A comparison of production-line control mechanisms", International journal of production research, Vol. 35, No. 3, 789-804, 1997.
[18] Hopp W.J. and M.L. Roof, "Setting WIP levels with statistical throughput control (SCT) in CONWIP production lines", International journal of production research, Vol. 36, 867-882, 1998.
[19] Hopp W.J. and M. L. Spearman, Factory physics, McGraw-Hill, 2nd edition, 2001.
[20] Yang R.L. and V. Subramaniam, "A control policy for manufacturing systems", to appear, 2005.
[21] Yang R.L. and V. Subramaniam, "A control policy for reentrant manufacturing systems", to appear, 2005.
[22] Personal communication with Prof. Gershwin S.B from MIT, 2004-2005.

Yang R.L. is a Ph.D. in National University of Singapore. She obtained M.S. in innovation of manufacturing systems and technology (IMST) from Nanyang Technological University. She studied for M.E. in robot control and intelligent control and B.A. in E.E. in Tongji University, China. Her research field is in control policies for manufacturing systems.

V. Subramaniam is an assistant professor from National University of Singapore. He holds Ph.D. and M.S from MIT. He obtained M.E. and B.S. from National University of Singapore. His research area is scheduling of manufacturing systems, scheduling of job shops, manufacturing cells, analysis and modeling of manufacturing systems.

S.B Gershwin is a senior research scientist from department of Mechanical Engineering of Massachusetts Institute of Technology. He received the B.S. degree in Engineering Mathematics from Columbia University, New York, in 1966; and the M.A. and Ph.D. degrees in Applied Mathematics from Harvard University, Cambridge, Massachusetts, in 1967 and 1971.
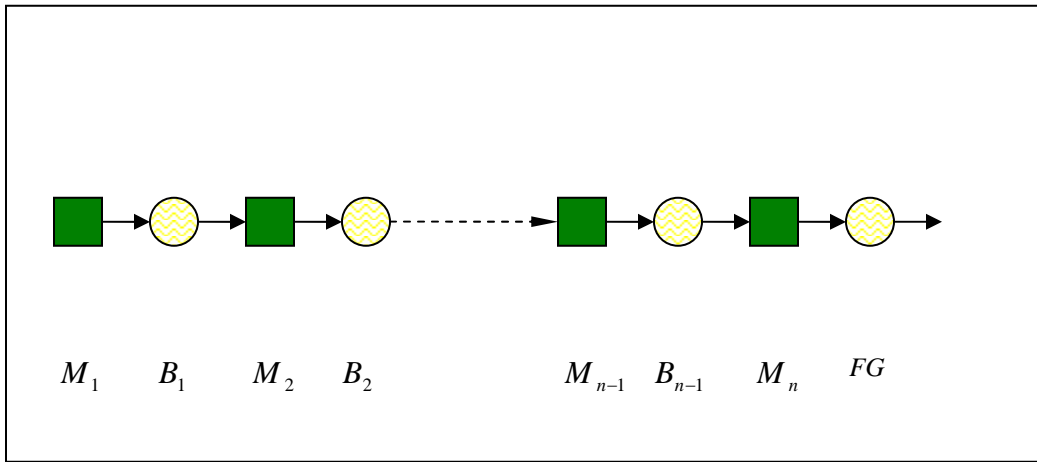
(No page numberings)
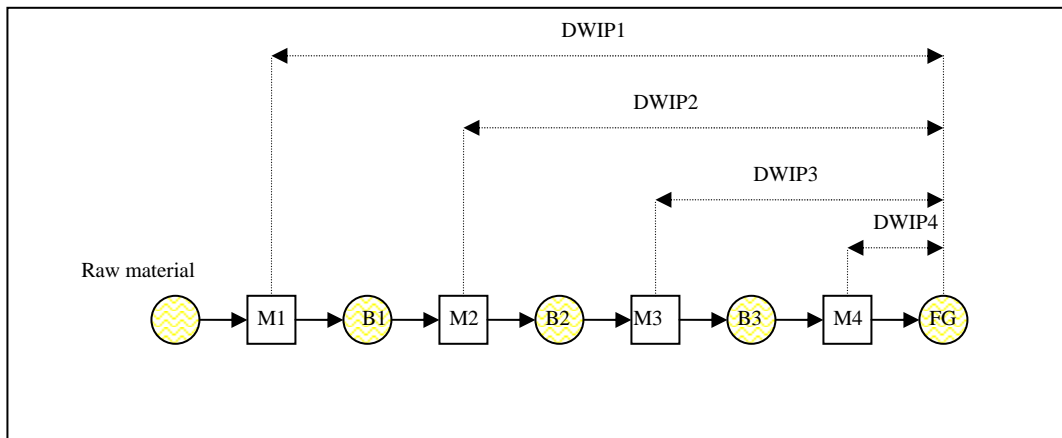
Figure 1: A production line

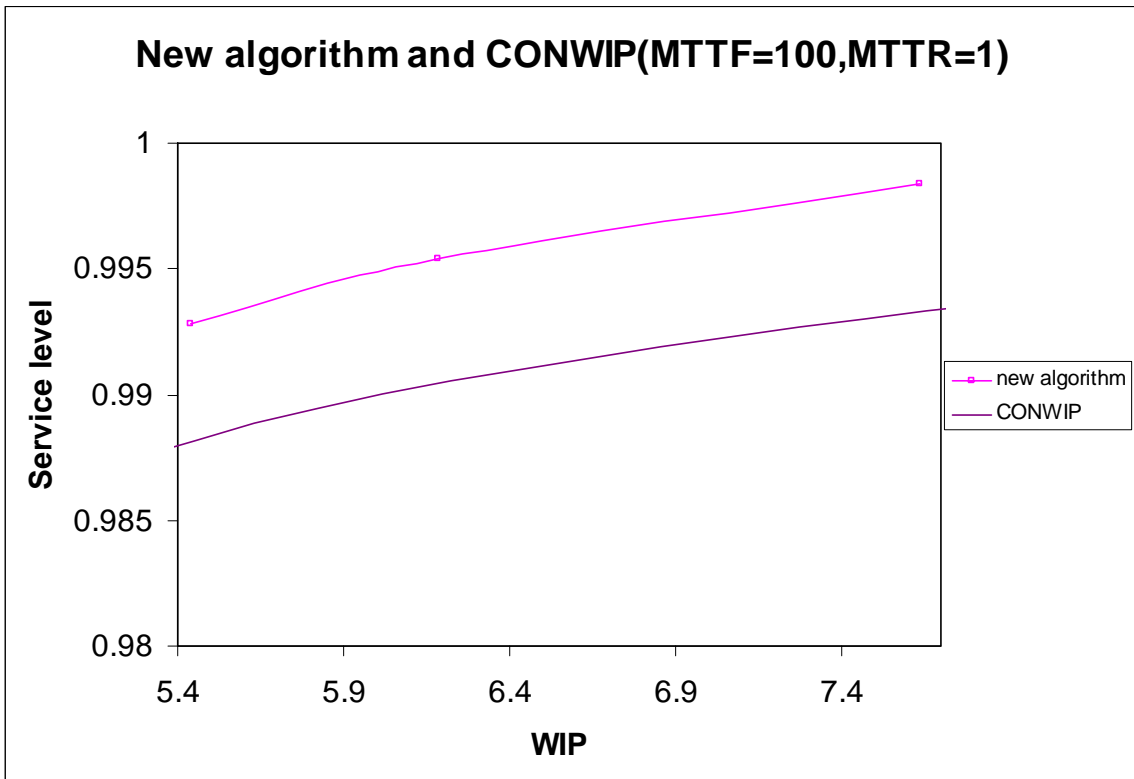Figure 2: A four machine production line for the new algorithm

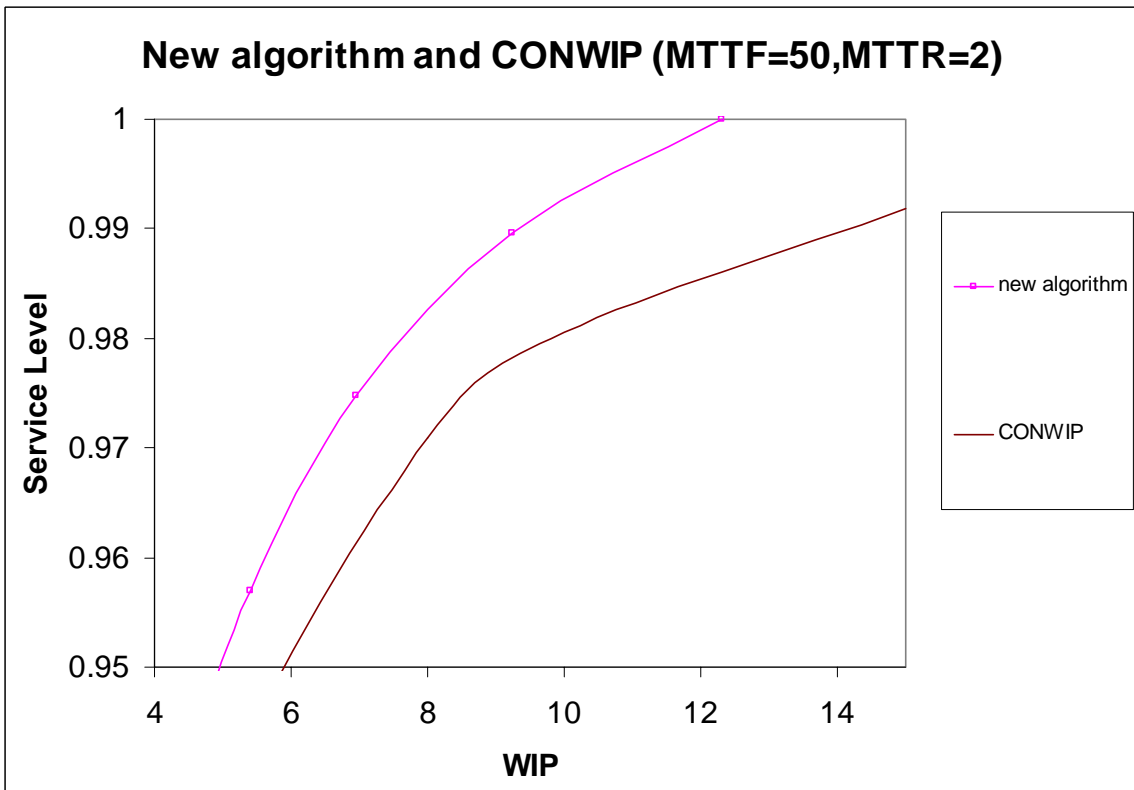Figure 3: Results for production line with infrequent random interruptions

Figure 4: Results for production line with more frequent random interruptions

(No page numberings)

Table 1: Experiment parameters

Machine	$M_1 \sim M_4$
MTTF	Two exponential distributions are used with mean of 100 and 50 respectively.
MTTR	Two exponential distributions are used with mean of 1 and 2 respectively.
Process Time	Log normally distributed with mean of 0.98 and standard deviation of .02mins.
Demand	Exponentially distributed with mean of 1min.
CONWIP Value	CONWIP value varies from 1 to 15.