An Adaptive Kanban and Production Capacity Control Mechanism

Léo Le Pallec Marand, Yo Sakata, Daisuke Hirotani, Katsumi Morikawa, Katsuhiko Takahashi

To cite this version:

Léo Le Pallec Marand, Yo Sakata, Daisuke Hirotani, Katsumi Morikawa, Katsuhiko Takahashi. An Adaptive Kanban and Production Capacity Control Mechanism. 19th Advances in Production Management Systems (APMS), Sep 2012, Rhodes, Greece. pp.452-459, 10.1007/978-3-642-40352-1_57 . hal-01472276

HAL Id: hal-01472276

https://inria.hal.science/hal-01472276

Submitted on 20 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

An Adaptive Kanban and Production Capacity Control Mechanism

Léo Le Pallec Marand, Yo Sakata, Daisuke Hirotani, Katsumi Morikawa and Katsuhiko Takahashi*

Department of System Cybernetics, Graduate School of Engineering, Hiroshima University, 1-4-1, Kagamiyama, Higashi-Hiroshima, 739-8527, Japan
* takahashi@hiroshima-u.ac.jp

Abstract. This paper proposes an adaptive kanban and production capacity control system as a new production planning and control mechanism for *Just-in-Time* environments to minimize the long term average inventories, Work-In-Process, backorders and operating costs. It is based on the adaptive kanban system proposed by Tardif and Maaseidvaag (2001), but dynamically adjusts both the number of kanbans and the level of production capacity with respect to inventories and backorders. It is expected to be more resistant to changes in the demand than previous *pull* ordering mechanisms. We present how to evaluate its performance for the case of a single-stage, single-product manufacturing process with exponential processing times and demand following a Poisson process. Simulation results under variable demand means are presented.

Keywords: Adaptive production capacity, Adaptive kanban system, Markov analysis

1 Introduction

Just-In-Time ordering systems, such as the Kanban and Conwip mechanisms, are used for reducing the costs associated with inventories and work-in-process (WIP) while maintaining a high level of service for customers. These systems authorize the release of a production order only when a demand is received from the customer and buffers are allocated to each inventory point to absorb changes in demand or production. However, this is for absorbing stable changes only, if unstable changes appear the allocated buffers cannot absorb them. Therefore *pull* systems are most efficient in environments with stable demand and processing times. The traditional approach is to optimize the number of cards used to control the release of production orders over a period based on demand forecasts. However, using forecasts contradicts the *Just-In-Time* principles and setting the number of cards over a period prevents timely readjustments if environmental conditions change quickly. Therefore, various mechanisms have been proposed to dynamically adjust WIP levels in order to increase the performance of *pull* systems in environments facing demand variability [2].

adfa, p. 1, 2011.
© Springer-Verlag Berlin Heidelberg 2011

The dynamic card controlling mechanism for CONWIP systems [1] adds or subtracts production cards when the manufacturing system's throughput (in make-to-order environments) or service level (in make-to-stock environments) is less or more than a production target. The reactive kanban system [4] dynamically adjusts the number of kanbans, using control charts for monitoring demand inter-arrival times and detecting unstable changes in the demand distribution.

The adaptive kanban system (AKS) was proposed [6] for a single stage manufacturing process. It does not monitor inter-arrival or output times but finished goods and backorders levels. Based on the monitored levels, additional kanban cards are released and captured in addition to initial kanban cards. For setting the AKS, Genetic Algorithm- and Simulated Annealing-based heuristics were compared [3]. Also, the AKS was adapted for a two-stage production line, and the benefits of simultaneously or individually adjusting the number of kanbans at each stage were compared [5].

These systems dynamically adjust the number of production cards, but the amount of WIP is not the only parameter impacting service levels and operating costs. The amount of production capacity also impacts service levels (by conditioning the manufacturing system's throughput) and operating costs (having more capacity than needed is a waste). The mechanism presented here, based on the AKS, dynamically adjusts not only the amount of WIP but also the throughput by changing the amount of active resources. It is expected to minimize inventories, WIP and operating costs while maintaining a high level of service and be more resistant to changes in the demand.

Section 2 presents our Adaptive Kanban and Production Capacity Control system and how to evaluate its performance. We study its performance compared to the Adaptive Kanban system in Section 3 and conclude with brief comments in Section 4.

2 Adaptive Kanban and Capacity Control System

For simplicity we restrain our attention to the case of a single-stage, single product manufacturing process MP with S identical parallel servers. We assume demand to follow a Poisson process of rate λ_D and an exponential processing time at each server with rate μ_P. These assumptions are quite restrictive but will allow us to use Markovian analysis for estimating the performance of the system and utilized in the literature [1, 3, 5, 6]. We further assume there is no shortage of raw parts, which can be considered true for sufficiently strict service level requirements on the preceding production stage. With the assumption of exponential processing times, parts exit MP with S identical parallel servers according to a state-dependent Markovian process of rate $\mu_P(n)$, where n is the number of orders in MP; $\mu_P(n) = n * \mu_P$ if n smaller than S and $\mu_P(n) = S * \mu_P$ if n greater than S.

For the single-stage, single product manufacturing process, we propose an adaptive kanban and production capacity control system (AKCS) based on the AKS. The complete mechanism of AKCS is described in Fig. 1. The decision variables are the number of kanban cards K in original kanban system, the number of additional kanbans E, a release threshold R and a capture threshold C in AKS and the number of flexible servers M. By setting the decision variables, the total cost Z minimized. The total cost

Z consists of the cost for holding inventories and Work-In-Process I, the backorders cost B, the cost for switching flexible servers F, and the cost for operating the manufacturing process O, and the unit costs are defined as h, b, f, and d.

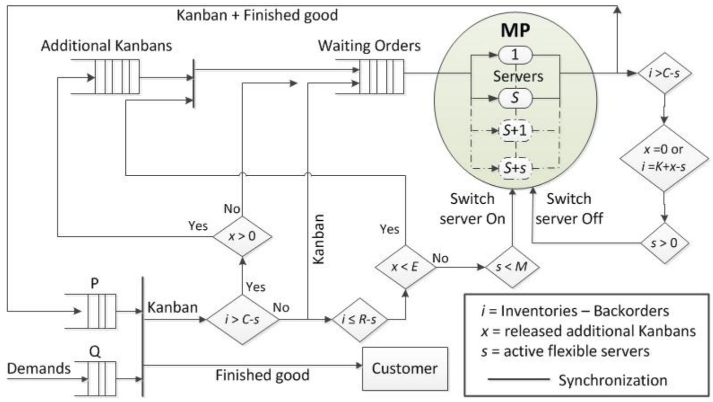

Fig. 1. Adaptive Kanban and Production Capacity Control mechanism

As in the kanban system, all of K kanbans are attached to finished parts stocked in queue P one by one. When a finished part is delivered to the customer, the attached kanban is removed and sent to MP as a production order. If P is empty, new demands are stocked in queue Q as backorders until a part is processed through MP and delivered to the customer.

Furthermore, as in AKS, in addition to K initial kanbans, up to E additional cards can be released into the manufacturing system or captured whenever inventories reach a release threshold R or a capture threshold C. When a demand arrives and inventories are equal to R or less, not only an initial kanban but also an additional kanban are released in the manufacturing system as production orders if it is available. If inventories are greater than C when a demand arrives and some additional cards have already been released, the kanban attached to the finished good delivered to the customer is captured.

In our Adaptive Kanban and Capacity Control system (AKCS), not only the number of kanbans but also the production capacity of the manufacturing process is adjusted, by increasing or decreasing the number of active servers in MP. Then, in addition to the initial number of S fixed servers, M flexible servers can be switched on and off to adjust the capacity of M. In the AKCS, finished parts level and on-off of flexible server will be controlled, and costs not only for inventory holding and backorders but also for operating flexible servers will be suppressed.

Increasing capacity is synchronized with demand arrivals while reducing capacity is synchronized with finished parts exits from MP. The flexible servers are switched on when inventories reach R but all additional kanbans have already been released. Flexible servers are switched off when inventories outnumber C but all additional kanbans have already been captured or when the number of busy servers becomes too small (equal to s). For the system to be able to return to its initial state K must be greater than C, and C equal or greater than R. For suppressing costs for inventory holding and backorders, adapting flexible servers is more effective than adapting additional kanbans, however, it leads to an increase in the cost for switching flexible servers. Then, adapting flexible servers only after adapting additional kanbans is considered for suppressing the cost. This system is equivalent to the AKS if M=0, and to the traditional kanban system (TKS) if M=0 and E=0.

Let i be the amount of inventories minus backorders, s the number of active flexible servers and x the number of additional kanbans released at the time. The proposed control rule for adjusting the number of kanbans and flexible servers is shown as follows.

Control rule for additional kanbans:

$$\begin{cases} \text{If } i \le R-s \text{ and } x < E \text{ when a demand arrives;} & \text{release a kanban } (x=x+1) \\ \text{If } i > C-s \text{ and } x > 0 \text{ when a demand arrives;} & \text{capture a kanban } (x=x-1) \end{cases}$$

Control rule for flexible servers:

$$\begin{cases} \text{If } i \le R-s \text{ and } x = E \text{ and } s < M \text{ when a demand arrives;} & \text{switch on a server } (s=s+1) \\ \text{If } i > C-s \text{ and } x = 0 \text{ and } s > 0 \text{ when a part exits MP;} & \text{switch off a server } (s=s-1) \\ \text{If } i = K+x-s \text{ and } s > 0 \text{ when a part exits MP;} & \text{switch off a server } (s=s-1) \end{cases}$$

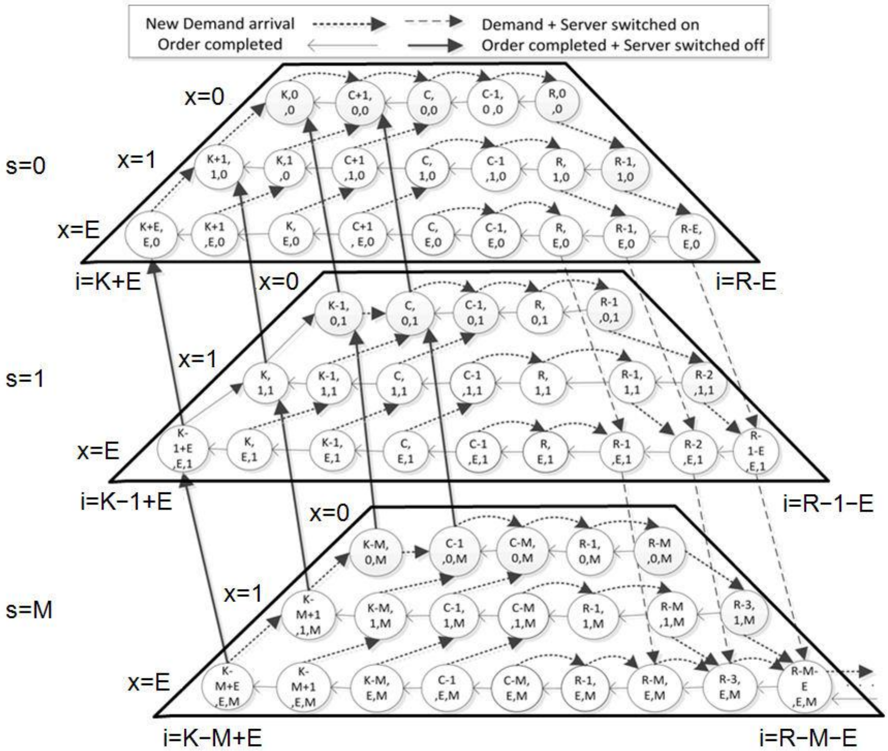

Fig. 2. Markov Chain for E=2,M=2, C>R

Under the assumption of exponential processing times and Poisson demand arrivals, we can model this system as a Markov Chain in terms of states (i, x, s). An example of the Markov Chain of the proposed AKCS is shown in Fig. 2.

If the maximum throughput of MP is greater than the average demand rate then the steady state probability $P(i, x, s)$ of each state exists and can be calculated. Summing the balanced equations of all states (i, E, M) where $i < R–E–M$, we can prove Eq. (1).

$$P(i,E,M) * \mu_P(i,x,s)/\lambda_D \quad = P(i+1,E,M) \tag{1}$$

Therefore the balanced equations of all states (i, x, s) where $K–s+x < i < R–s–x$ form a finite linear system of N independent equations with N unknown variables $P(i, x, s)$. This system is easy to solve and once the steady probability of each state (i, x, s) has been calculated, and we can estimate the following total cost Z of the system.

$$Z = I + B + F + O \tag{2}$$

Because finished parts and WIP are attached to a kanban, the inventory holding cost I is equal to the average number of cards in the ordering system multiplied by the unit cost h set to 1.

$$I = h * K + h * \sum_{s=0}^{M} \left(\sum_{x=0}^{E} \sum_{i=R-s-x}^{K-s-x} x * P(i,x,s) \right) \tag{3}$$

The backorders cost B is equal to the average amount of backorders multiplied by the unit cost b.

$$B = b * \sum_{s=0}^{M} \sum_{x=0}^{E} \sum_{i=0}^{\infty} i * P(-i,x,s) \tag{4}$$

The cost F is equal to the average number of times production capacity is adjusted, multiplied by the unit cost f. The first part of the formula is the number of times a server is switched on. The second and last parts are the number of times a server is switched off when $i > C–s$ and x=0 or when $i=K+x–s$.

$$F = f * \sum_{s=1}^{M} \left(\sum_{i=R-s-E}^{R-s} \lambda_d * P(i,E,s-1) + \sum_{i=C-s}^{K-s} \mu_P(i,0,s) * P(i,0,s) + \sum_{x=1}^{E} \mu_P(K-s+x,x,s) * P(K-s+x,x,s) \right) \tag{5}$$

The cost O is equal to the average number of active servers, multiplied by the unit cost d.

$$O = d * S + d * \sum_{s=0}^{M} \left(s * \sum_{x=0}^{E} \sum_{i=R-s-x}^{K-s-x} P(i,x,s) \right) \tag{6}$$

3 Performance Study

3.1 Expected Performance under Stable Demand.

We consider the single stage, single product manufacturing process described in section 3, with μ_P = 0.15, λ_D = 0.2 and the following cost parameters: b=1000, f=1000, d=1. Table 1 presents the best results obtained with the AKCS for S+M varying from 4 to 6 compared to the best result obtained using the TKS and AKS. Each time, parameters K, E, R, and C were chosen for fixed values of S and M after conducting an exhaustive search over the space of possible values. Here, the optimum number of servers with the AKS and TKS is 4. The AKCS yields better performances only if more than 4 servers are available and the gains are very small. However, cost parameters have a strong impact on performance and this is only a result under the condition of d=1 and f=1000. Cost parameters could be changed to advantage the AKCS (smaller f, bigger d). Also Table 1 shows results under stable environment while the AKCS performs much better in unstable environments where there is a great need for adjusting both WIP and capacity (see Section 3.2).

Table 1. Performance under stable demand; d=1, f=1000, b=1000

	K,E,R,C	*S*	*M*	*S+M*	*I+B*	*F*	*O*	Total	Gain
TKS	7,0,0,0	4		4	7.959		4	11.959	
AKS	6,1,5,6	4		4	7.926		4	11.926	
AKCS	8,0,0,7	3	1	4	9.089	0.387	3.007	12.483	
	7,1,0,5	4	1	5	7.687	0.170	4.002	11.859	0.57%
	7,1,0,5	4	2	6	7.614	0.229	4.002	11.844	0.68%

For a fixed total number of servers, using more flexible servers leads to bigger inventories, WIP and backorder costs but smaller operating costs. Using Markov analysis we study the impact of the cost ratios f and d on the performance of the AKCS. Demand is considered stable with λ_D=0.2. Fig. 3 presents the space (f, d) where the local optimum configuration of the AKCS uses M flexible servers, when b=1000 and S+M=5. As d increases Op increases and using more flexible servers becomes more interesting. As f increases the cost for adjusting capacity increases and performance is reduced. Therefore, if the penalty for switching servers is not prohibitive and d is expensive using more flexible servers may reduce the total cost and improve performance, but if f is too high and d too small, then increasing M reduces performance.

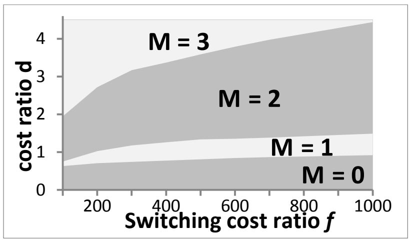

Fig. 3. Impact of the cost ratios f and d on the choice of M, for S+M=5 and b=1000.

Fig. 4 illustrates the impact of parameter *M* on the total cost, for *b*=1000 and various values of *f* and *d*, when the total number of servers is constant. We observe that the expected gain from the AKCS compared to the AKS is very dependent on the cost ratios. For *f*=900 and *d*=0.5, using more flexible servers is counterproductive as the gains from the reduced *O* are too small to balance *F* and the increased *I* and *B*. As *d* increases so does *O* but the expected gain from using more flexible servers becomes significant and the effectiveness of increasing *M* increases. When *d*=3.5, if *f* varies from 900 to 300 then *F* decreases and the effectiveness of using more flexible servers increases. We observe that the AKCS outperforms the AKS only if the reduced *O* outweighs the increase in *I* and *B* and if *F* is not too high, but if these conditions are met, it is able to reduce the total cost.

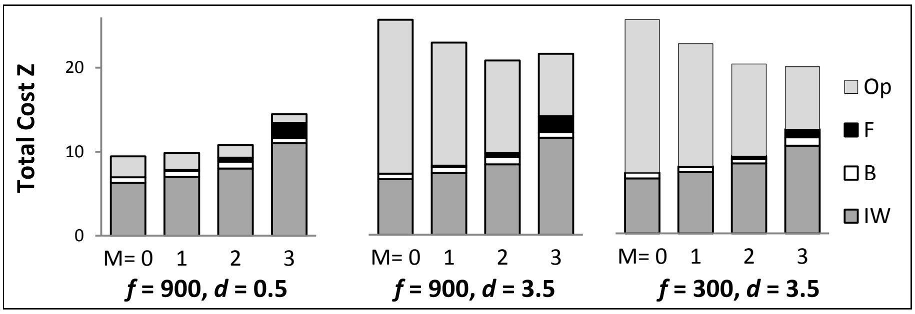

Fig. 4. Impact of *M* on the total cost for *S*+*M*=5, *b*=1000 and various *f* and *d*.

3.2 Expected Performance under Unstable Demand

The objective of this paper is to propose a Production Planning and Control system more resistant to changes in the demand than existing mechanisms. To assess its performance under unstable demand we conducted a discrete event simulation of a random walk through the Markov Chain. We considered the same manufacturing process and cost parameters as in the preceding study but unstable demand was introduced as follows: the mean of the demand inter-arrival times is set to 3 during the first 25 time units, 5 for the next 25 time units and 7 for the next 25 time units. This pattern repeats itself for the duration of the simulation. We considered a simulation length of

Table 2. Simulation under unstable demand mean; *b*=1000, *f*=1000, *d*=1

	K,E,R,C	*S*	*M*	*S+M*	*I+B*	*F*	*O*	Total	Gain
TKS	7,0,0,0	4		4	55.312		4	59.312	
	9,0,0,0	6		6	9.928		6	15.928	
AKS	6,1,5,6	4		4	58.766		4	62.766	
	8,1,6,7	6		6	9.671		6	15.671	
AKCS	8,0,0,7	3	1	4	67.589	0.067	3.195	70.851	
	7,1,0,5	4	1	5	31.833	0.041	4.053	35.927	
	7,1,0,5	4	2	6	25.540	0.063	4.070	29.672	
	10,0,6,8	3	3	6	10.437	0.699	3.462	14.597	6.9%
	9,0,4,5	4	3	7	9.493	0.387	4.110	13.989	10.7%

1,000,000 time units after a warm up of 100,000 time units. Table 2 presents the results of the best settings from Table 1 and new best settings. Parameters *K*, *E*, *R*, and *C* were chosen for fixed values of *S* and *M* after conducting an exhaustive search over the space of possible values. We observe that our proposed mechanism is much more interesting under unstable demand, as the performance gain compared to the AKS is now 6.9% with the same total number of servers and 10.7% with one more server.

4 Conclusion

This paper considered unstable changes in product demand and proposed an Adaptive Kanban and Production Capacity Control system (AKCS) that adjusts both Work-In-Process and production capacity for minimizing the total operating costs while maintaining a high level of service. The AKCS is able to outperform existing ordering systems under certain conditions and is especially interesting in environments facing demand variability where the cost for adjusting production capacity is acceptable and operating costs relatively high. It would be interesting to assess its performance on multi-stage, multi-product environments. The control rule relies only on monitoring inventories and therefore should be easy to implement.

In the present version, production capacity is adjusted only after adjusting the number of kanbans. Instead, releasing a different number of additional kanbans for each capacity level may increase the global performance by better adjusting the amount of WIP to the current manufacturing process capacity and reducing the risk of switching resources on and off in a short succession.

Acknowledgements. This work was partially supported by the Grant-in-Aid for scientific research of Japan Society for the Promotion of Science in 2010-2012.

References

1. Framinan J.M., González P.L., Ruiz-Usano R.: Dynamic Card Controlling in a Conwip System, Int. J. Prod. Econ. 99, 102-116 (2006)
2. Lage, J.M., Godinho, F.M.: Variations of the Kanban System: Literature Review and Classification, Int. J. Prod. Econ. 125, 13–21 (2010)
3. Shahabudeen P., Sivakumar G.D.: Algorithm for the Design of Single-stage Adaptive Kanban System', Comput. Ind. Eng. 54, 800-820 (2008)
4. Takahashi K., Morikawa K., Nakamura N.: Reactive JIT Ordering System for Changes in the Mean and Variance of Demand, Int. J. Prod. Econ. 92, 181-196 (2004)
5. Takahashi K., Morikawa K., Hirotani D., Yorikawa T.: Adaptive Kanban Control Systems for Two-stage Production Lines, Int. J. Manuf. Tech. Manag. 20, 75-93 (2010)
6. Tardif V., Maaseidvaag L.: An Adaptive Approach to Controlling Kanban Systems, Eur. J. Oper. Res. 132, 411-424 (2001)