A Nested Configuration of POLCA and Generic Kanban in a High Product Mix Manufacturing System

Oladipupo Olaitan[0000-0002-4604-2543] , Giuseppe Fragapane[0000-0001-8915-509X]
Erlend Alfnes[1111-2222-3333-4444] and Jan Ola Strandhagen[1111-2222-3333-4444]

[1] Norwegian University of Science and Technology (NTNU), Trondheim, Norway
Oladipupo.a.olaitan@ntnu.no

Abstract. The work presented in this paper is part of a project aimed at streamlining and improving the process flow at a leather furniture manufacturing company. The manufacturing throughput time is highly variable, and this has resultant effect on subsequent stages of the process, where the components produced are to be assembled. It would be highly beneficial to have a high level of predictability at the upstream stages of the production process, such that the downstream assembly operation can be planned in accordance with the expected arrival of components for specific product types. Research conducted in a previous phase of the project showed that the application of the CONstant Work In Progress (CONWIP) control mechanism to regulate inventory yielded significant improvements in the throughput time's mean and variation. However, as it is the case with tighter control of inventory in manufacturing, previously unrealised problems were exposed in relation to the selection of the product model to release into the CONWIP loop. This has significant impact on the balance of the distribution of workload across the system's workstations and among the multi-skilled teams at one of the workstations.

This research implements a nested configuration of the Paired-cell Overlapping Loop Of Cards with Authorisation (POLCA) and the Generic Kanban control mechanisms to achieve a balance of the workloads. This ensures a synchronised flow of the different product mix through the entire manufacturing system.
Keywords: POLCA, Generic Kanban, High-mix Manufacturing.

1 Introduction

The research presented in this paper is part of a project titled SØM4.0, which is aimed at streamlining and improving the process flow at a furniture manufacturing company. There is a significant amount of deviation between the value adding time and the throughput time. Additionally, the throughput time is highly variable, and this has resultant effect on subsequent stages of the process, where the components produced are to be assembled.

Research conducted in a previous phase of the project showed that the application of the CONWIP control mechanism to regulate inventory and a sequencing rule to control production yielded significant improvements in the throughput time's mean and variation [1, 2]. However, as it is the case with tighter control of inventory in manufacturing, previously unrealised problems were exposed in relation to the selection of the product model to release into the CONWIP loop. This has significant impact on the balance of the distribution of workload across the manufacturing stages and among the operator teams at one of the processing steps.

The aim of the research reported here is to implement a production control mechanisms that would provide a balance of workload, both across the system's workstations and among the teams at one of the workstations. This would ensure a synchronised flow of the different product mix through the entire system.

The rest of the paper is organised as follows. In the next section, a more detailed overview is provided of the case study company's challenges, previous works that have been done to overcome the challenges and the current challenges that are to be addressed in this work. In Section 3, related production control mechanisms that have been applied to address similar challenges are reviewed and discussed. Section 4 will describe a concept for achieving workload balance through a nested configuration of control mechanisms in the case study system. Finally, Section 5 will provide insights and conclusions on the nested configuration concept described in this work. It will discuss the benefits of such nested configuration of production control mechanisms and establish grounds for future work to further investigate the concept.

2 Overview of Case Study System

The case study company is a typical example of high-mix production. The product line consists of 36 different models, most of which are offered in two or three different sizes (small, medium and large), while other models are offered with two differential models. Furthermore, all the products are offered with a wide selection of materials and colours. The main variety explosion occurs right from the beginning of the production process where the model and materials are selected.

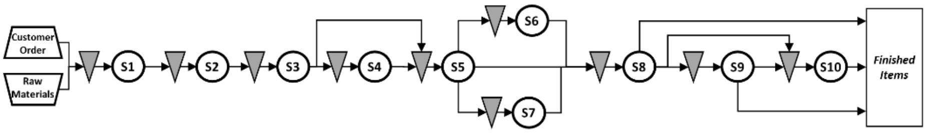

Fig. 1. Process Flow Chart

The leather pieces for the furniture are cut in Step 1 followed by them undergoing variety of sewing operations between Steps 2 and 10, which constitute the sewing section. This work focuses on Steps 2 to 10 (i.e. S2 to S10), which are the most labour intensive and value-adding processes of the whole production. The routing possibilities through these processing steps differ from one product model to another, as shown in Fig. 1.

2.1 Background

The company previously applied a Push control mechanism to release items from the cutting into the sewing section, followed by applying a FIFO rule to sequence their processing at the workstations. Barring any capacity constraints, this should ensure that the products are completed as planned in the production order. However, the direct release of semi-finished items from the leather cutting section (Step 1) into the downstream sewing section created a constant need for human intervention to re-sequence and reshuffle the in-process items. A personnel had to monitor the workloads at the different teams and redistribute waiting items to alternative teams that had spare capacity for processing.

To overcome the problem highlighted above, a CONWIP control was implemented in the sewing section, which meant that a limit was set on the number of WIP allowed in the sewing section and an intermediate buffer created between the two sections. The lower amount of WIP, judging by Little's law [3], reduced the throughput time [1, 2]. The intermediate buffer served decoupling purposes between the flows out of the cutting section and the releases into the sewing section. As a result, the resequencing became easier, as the number of items that had to be re-ordered reduced, and the level of *chaos* in the sewing section reduced, although this is not directly quantifiable.

However, although the lower WIP level reduced the effort needed for reshuffling, it did not ensure a balanced workload distribution across the teams at S8. The release of items from the intermediate buffer into the sewing section did not consider the current level of inventory for individual product models or operator teams at S8. Prior to the implementation of the CONWIP, there were excessive items in the sewing section which ensured that the workstations and operator teams were consistently utilised. With the CONWIP, it became necessary to ensure that the limited number of items released into the sewing section consisted of a balanced distribution of workload across the workstations and the operator teams at S8.

Making use of additional information, such as the capacity availability and level of inventory downstream, can significantly improve item release decisions. And, as reported in a previous study, the release method into a system is more crucial than the sequencing or prioritisation rule at individual workstations, such that if the release rule functions well, the setting of sequencing rule at individual workstation becomes less effective or even counterproductive [4, 5]. Hence, the aim of the current work reported here is to implement control mechanisms that will synchronise the release of items into the sewing section with the current state of the downstream workstations, particularly the operator teams of Step 8. This shares some of the objectives of existing production control techniques which are going to be discussed in the next section.

3 Review of Literature on Related Control Mechanisms

The card-based mechanisms that are available in literature can be categorised into unit-based or load-based, depending on if their signal card represents the physical inventory implication or the workload impact of releasing an item for processing. The original card-based mechanism was the traditional Kanban control strategy of Toyota Production System, and it was unit-based [6]. Other unit-based mechanisms are the CONWIP [7], the Generic Kanban [8], the paired-cell Overlapping Loops of Cards with Authorisation (POLCA) [9] and many others. Load-based mechanisms adapted workload control concepts [4, 10] into card-based control mechanisms. As such, some of the existing load-based mechanisms have retained the same control logic of previously existing unit-based mechanisms, but with load-based interpretation of their cards. Examples of such load-based mechanisms are the Constant Load (CONLOAD) [11], Control of Balance by Card Based Navigation (COBACABANA) [4] and Load-based POLCA [12], which can be said to be the load-based versions of the CONWIP, Generic Kanban and the POLCA respectively. The aim of these load-based implementations was to take into account that different products might have different workload impacts on a manufacturing resource, especially in high-mix systems in which the processing time requirements of the system's products vary significantly. Unit-based mechanisms, on the other hand, would not adequately quantify the difference in the load impacts of releasing a unit of a short processing and a high processing product into a system.

However, load-based mechanisms are generally less used than the unit-based mechanisms, because of challenges that have to do with accurate quantification of the load impact of released item units, as well as the complexity of their software and hardware requirements for implementation and execution [13]. As evident in the COBACABANA, for which two methods of representing workload have been reported [4, 14], it can be difficult to find a trade-off between requiring many cards to precisely represent different product workload ranges and reducing the number of cards with the implication of a less precise load representation.

Furthermore, some of the latter control mechanisms have been adapted to overcome the limitations of previous mechanisms in non-repetitive, dynamic environments, by incorporating necessary flexibility in the logic for the transmission of the cards. It is often easier to adapt a control mechanisms to an environment – even if it results in some loss of effectiveness – than to adapt the environment to the strategy [8]. Some of the common adaptations are the use of generic (or centralised) control of cards as found in the Generic Kanban and the Generic POLCA, which are adaptations from the traditional Kanban and POLCA respectively. In the original control mechanisms, an item obtains the authorisation cards for processing at the workstations one at a time; while in the generic adaptations, an item must obtain all the authorisation cards it requires for its required processing steps before it is released into the system. The aim is to ensure that an item has obtained authorisation cards for all the workstations, to indicate the availability of capacity at all the stages at which it is to be processed, before it is released at all into the system. Hence, once an item is released, subsequent queueing in between downstream stages is minimised because it does not have to wait for authorisation cards and, as a consequence, the system inventory level is reduced and replaced with card inventories instead. Another benefit of the generic control of authorisations cards is that it makes the control mechanisms applicable to systems with flexible routing, since each product model type can dynamically select the cards for the workstations that belong to its route.

However, the global control of cards, can lead to blocking in high-mix systems in which products have different routing possibilities. Because items obtain all the card for the workstations along their routes before being released into the system, it could become a problem for other products. If for any reason an item becomes stuck at upstream stations that it does not share with other items, it will continue to hold on to downstream authorisation cards that could be directly used by other items that are processed through different upstream workstations. This could lead to starvation of the downstream workstations, if there is no card available to authorise the release of those other items. Methods, such as continuous monitoring and direct release of items to starving workstations even if it violates their workload limits [15], and the differentiation of indirect from direct loading , have been proposed to avoid this blocking.

This problem has been described in other studies on divergent systems [14] as premature idleness of the downstream stations, and recommendations have been made on how to overcome the problem [16]. Another investigation into this problem suggests that a proper setting of the load limit at critical loops can be used to mitigate it [13].

4 Nested Implementation of Multiple Control Mechanisms

Based on the above it is clear that control mechanisms that operate generic control cannot be applied in the type of highly variable routing system of this work. It is for this same reason that the Generic Kanban, the GPOLCA and the COBACABANA cannot be applied in the system considered here. However, it is possible to apply elements of the Generic Kanban in a way that takes advantage of its workload regulating attribute to balance the workload across the operator teams of S8. It has been previously demonstrated in existing works that COBACABANA, which has a similar logic to the Generic Kanban, is effective at balancing workload across intermediate stages of a manufacturing system [13, 17]. This attribute, through the Generic Kanban, will be extended in this paper to balancing workload across the lateral workstations at the same workstation, i.e. for the operator teams of S8 in the case study system.

In a nested configuration of two control mechanisms, the POLCA will be applied across the system, while a Generic Kanban will be used to synchronise the release of items into the system with the work rate of the operator teams at S8. The steps involved in setting up this nested configuration, are as described in the following sub-sections.

4.1 Implementation of POLCA

POLCA has been reported to be suitable to high-mix, low-volume environments in which the traditional Kanban control mechanism would result in the proliferation of inventory. This is because the traditional Kanban control mechanism would need to keep stage level product specific base stock for each of the products produced in the system [18]. POLCA uses its cards to signal the availability of capacity downstream and release parts for processing, unlike the traditional Kanban control mechanisms which uses its cards to signal the need for the transfer of a specific part type downstream to fulfil a demand or to replenish stock. POLCA keeps inventory of cards instead of physical parts, and as a result is able to avoid the need to keep inventory for each specific product type. The above contrasts between the two have only been made to enable the reader understand POLCA within the context of the more widely known traditional Kanban control, and to describe some of the reasons why the traditional Kanban is not being considered in this work.

The first step taken in the implementation of POLCA in the case study system is to identify the possible paired cells formation for the eight different processing routes, as shown in Table 1. As shown in the table, there are between 3 to 5 paired cells that can be formed from each of the routes.

Table 1. Routing Information with Paired Cells

Route	Steps	Number of Paired Cells
R1	S2-S3-S5-S8	3
R2	S2-S3-S4-S5-S6-S8	5
R3	S2-S3-S5-S7-S8	4
R4	S2-S3-S5-S8-S9	4
R5	S2-S3-S5-S7-S8-S9	5
R6	S2-S3-S5-S8-S9-S10	5
R7	S2-S3-S5-S8-S10	4
R8	S2-S3-S5-S7-S8-S10	5

Next, the possible paired cells across the routes are listed long with the routes to which they apply and the total number of such routes, as shown in Table 2. The aim of this information is to ensure that the paired cells are set up with adequate cards to correspond to the number of routes they serve, and the expected production volume of the products processed along the routes. This is because, as earlier discussed in Section 3, a previous study has observed that setting appropriate card limits at such crucial points in the system is important to prevent the possibility of the system getting blocked.

Table 2. Paired Cells and Routes Served

Paired Cells	Routes Served	Number of Routes Served
S2-S3	All	8
S3-S4	R2	1
S3-S5	All, except R2	7
S4-S5	R2	1
S5-S6	R2	1
S5-S7	R3, R5, R8	3
S5-S8	R1, R4, R6, R7	4
S6-S8	R2	1
S7-S8	R3, R5, R8	3
S8-S9	R4, R5, R6	3
S8-S10	R7, R8	2
S9-S10	R6	1

The card allowance for the paired cells can be determined using the same formula in [19], but with the inclusion of the average lead time across the different product models that share the paired cells, as well as the expected production volume for all the product models during the planning horizon, in the original formula as expressed in equation (1).

$$Number\ of\ cards\ for\ paired\ cell\ i/j\ = \left(LT_i + LT_j\right).\mathrm{x}\ (\frac{NUM_{i,j}}{D}) \quad (1)$$

Where LT_i and LT_j are the estimated average lead times for the two cells (workstations) over the planning period of length, D, and $NUM_{A,B}$ is the total number of items that go from workstation i to j during the same period. Here, this will be the total number across the product model types that use the paired cells.

4.2 Implementation of Generic Kanban Control

At this point, the unbalanced workload across the teams at S8 has not been directly addressed, even if the POLCA's balancing of workload across workstation would have had indirect impact in regulating the flow of items into S8 [13, 17]. Therefore, a Generic Kanban control mechanism is implemented to balance workload among the operator teams of S8. The Generic Kanban operates a global assignment of cards and sets limits on the number of WIP items per workstation, which are monitored globally through centralised display boards by a planner to determine the possibility to release newer items into the system. Unused cards, denoting available allowance from the WIP limit at each station, are displayed on the board. A new item can only be released into the system if there are spare cards on the board for each of the workstations at which it is to be processed. Once the item is released, cards for each of the workstation in its processing route are removed from the board and attached to it. As it completes each of the operations, the corresponding cards are released from the item and sent back to the centralised board to signal the availability of capacity.

Table 3. Cards Settings and Generic Kanban Display for Step 8 Teams

Team	Number of Cards Set	Cards currently Occupied	Free cards displayed on Generic Kanban Board
A	35	30	5
B	26	26	0
C	28	24	4
D	32	30	2

As earlier reported, and as shown in Table 3, there are four operator teams responsible for carrying out the processing required for the different models at S8. The teams are skilled in the processing of different product models, with each team capable of processing more than one product model. Likewise, some of the product models have more than one team that is capable of processing them.

The implication is that the POLCA paired cells involving S8 (i.e. S5-S8, S6-S8, S7-S8, S8-S9 and S9-S10) will remain as they are, but items released into S8 will be routed to the available team that is qualified to process them. This means there will be four virtual POLCA loops (8A, 8B, 8C and 8D) into S8 and out of it, each representing the flow in and out of the teams within S8. Each virtual loop will have a set number of authorisation cards and the unavailability of cards for any of them would stop the release of products requiring the connected team into S8, but this won't prevent other teams from receiving items for which they have spare authorisation cards. Also, the information about the number of available authorisation cards for each of the virtual loops will be used globally as Generic Kanbans to select and release items for operator teams that have spare cards. As shown in Table 3, a Generic Kanban board showing the number of free cards at each of the operator teams is displayed to the planner to support the decision on the type of product model to release from the intermediate buffer into the sewing section. This ensures that the planner releases product model types upstream in synchronisation with the work rate of the teams at S8.

Therefore, the factors that have to be considered when a new item is to be released into first workstation of the sewing section (i.e. S2) are the availability of a S2-S3 POLCA card and the availability of spare cards on the board for one of the operator teams that are capable of processing the item. From then on, the progress of an item downstream through its required workstations is controlled locally by the availability of POLCA authorisation cards for the paired cells involved.

4.3 Insights, Conclusions and Future Work

The step by step practical implementation of POLCA in a case study manufacturing system, which has variable routings and high product mix, gives a good insight to industry practitioners on how POLCA can be implemented in systems with complex product routings. It shows how paired cells can be created from the system's workstations, while reducing the complications resulting from the variety of product routings, or from the common use of the same workstations along multiple product routes. If physical distance between workstations does not make exchange and synchronisation of cards between paired workstations difficult, then POLCA should be practicable in such systems, as presented.

It has demonstrated how the benefits of different control mechanisms can be combined in one system, through a nested configuration of POLCA and the Generic Kanban control mechanisms. This was necessary for the case study system of this research, in order to also achieve a balanced distribution of work load across the multi-skilled teams of a workstation. This extends existing knowledge from previous works which have mostly focussed on balancing workload across intermediate workstations of a manufacturing system. It should be mentioned that other concepts, such as Workforce training and Staffing and Balanced product-mix release methods, have been used to achieve workload balance at operations involving teams. However, such static methods would only offer limited solutions to the dynamic change in product mix and routing involved in the system considered here.

The future aim of this research is to implement the nested POLCA-Generic Kanban configuration and measure its performance. This would involve using simulation models to conduct experiments for configuring the card settings for the paired cells. Existing techniques, such as Evolutionary algorithms and metaheuristics algorithms, that have been successfully applied in simulation-optimisation for setting Kanban numbers in Kanban related literature will be adopted for this purpose.

Acknowledgement

This research has been carried out as part of SØM4.0 project, as funded by the Norwegian Research Council's BIA programme. The authors wish to thank the project partners for facilitating this work.

References

[1] Q. Yu, G. H. Brekken, and M. M. Eide, "A Simulation Enhanced VSM Approach for high-Mix Manufacturing Environment," presented at the International Workshop of Advanced Manufacturing and Automation, University of Manchester, UK, 2016.
[2] O. A. Olaitan, Q. Yu, and E. Alfnes, "Work In Process Control for a High Product Mix Manufacturing System," in *The 50th CIRP Conference on Manufacturing Systems*, Taichung, Taiwan., 2017.
[3] J. D. C. Little, "A Proof for the Queuing Formula: $L = \lambda W$," *Operations Research,* vol. 9, pp. 383-387, 1961.
[4] M. J. Land, "Cobacabana (control of balance by card-based navigation): A card-based system for job shop control," *International Journal of Production Economics,* vol. 117, pp. 97-103, 2009.
[5] M. J. Land, *Workload control in job shops, grasping the tap*: Labyrint Publications Ridderkerk, 2004.
[6] T. Ohno, "The origin of Toyota production system and kanban system," in *Proceedings of the International Conference on Productivity and Quality Improvement*.
[7] M. L. Spearman, D. L. Woodruff, and W. J. Hopp, "CONWIP: a pull alternative to kanban," *International Journal of Production Research,* vol. 28, pp. 879-894, 1990.
[8] T. M. Chang and Y. Yih, "Generic kanban systems for dynamic environments," *International Journal of Production Research,* vol. 32, pp. 889-902, 1994.
[9] R. Suri, "Don't push or pull-POLCA," *APICS The Performance Advantage,* vol. 8, pp. 32-39, 1998.
[10] W. BECHTE, "Load-oriented manufacturing control just-in-time production for job shops," *Production Planning & Control,* vol. 5, pp. 292-307, 1994.
[11] O. Rose, "CONLOAD—a new lot release rule for semiconductor wafer fabs," in *Proceedings of the 31st conference on Winter simulation: Simulation---a bridge to the future-Volume 1*, 1999, pp. 850-855.
[12] N. Vandaele, I. Van Nieuwenhuyse, D. Claerhout, and R. Cremmery, "Load-based POLCA: an integrated material control system for multiproduct, multimachine job shops," *Manufacturing & Service Operations Management,* vol. 10, pp. 181-197, 2008.
[13] N. Ziengs, J. Riezebos, and R. Germs, "Placement of effective work-in-progress limits in route-specific unit-based pull systems," *International Journal of Production Research,* vol. 50, pp. 4358-4371, 2012.
[14] M. Thürer, M. J. Land, and M. Stevenson, "Card-based workload control for job shops: Improving COBACABANA," *International Journal of Production Economics,* vol. 147, pp. 180-188, 2014.
[15] M. Thürer, M. Stevenson, and C. W. Protzman, "COBACABANA (Control of Balance by Card Based Navigation): An alternative to kanban in the pure flow shop?," *International Journal of Production Economics,* vol. 166, pp. 143-151, 2015.
[16] B. Oosterman, M. Land, and G. Gaalman, "The influence of shop characteristics on workload control," *International journal of production economics,* vol. 68, pp. 107-119, 2000.
[17] R. Germs and J. Riezebos, "Workload balancing capability of pull systems in MTO production," *International Journal of Production Research,* vol. 48, pp. 2345-2360, 2010.
[18] R. Suri and A. Krishnamurthy, "How to Plan and Implement POLCA–A Material Control System for High Variety or Custom-Engineered Products," *Center for Quick Response management Technical Report,* 2003.
[19] A. Krishnamurthy and R. Suri, "Planning and implementing POLCA: a card-based control system for high variety or custom engineered products," *Production planning and control,* vol. 20, pp. 596-610, 2009.