CONWIP Control for Multiproduct Systems

Sarah M. Ryan  
Iowa State University

F. Fred Choobineh  
University of Nebraska

Abstract: To implement CONWIP control in multiproduct systems, planners must specify not only a total WIP level, but also its allocation to the different products. We describe two methods for setting individual product WIP levels. The first approach focuses on equitably meeting randomly arriving orders for the various products. We designed a heuristic allocation procedure to attain a throughput target. The second approach uses a nonlinear program that bounds the throughput for a given total WIP. The optimal WIP mix is extracted from the upper bounding solution. Both approaches are validated by simulation in numerical examples.

Introduction: Research on pull production control policies in general and CONWIP in particular has focused primarily on flow lines that produce one or more similar products. However, the benefits of a limit on work in process (WIP) also can be expected to accrue in multiproduct production systems. Consistent with the increased complexity of producing multiple products, implementing a CONWIP policy is significantly more complicated than in a single product flow shop. First, planners must determine a total WIP level to achieve high total throughput. Second, the total WIP must be allocated to the different products.

The proportion of total WIP allocated to each product, the WIP mix, that optimizes system performance generally will not be the same as the product mix. The product mix, specified by the proportions of the total demand on the system that correspond to each product, can be treated as a known parameter vector. But the WIP mix should be influenced by the products’ work contents and the capacities of processing resources. We view it as a vector of decision variables.

This paper summarizes two approaches for determining the total WIP and WIP mix. The first approach focuses on equitably meeting randomly arriving orders for the various products. The goal of the second approach is to achieve a high throughput consistent with the product mix.

Achieving a Specified Service Level: In a make-to-order system producing multiple product types, the objective is to provide a high customer service level that is equitable across product types, without maintaining excessive WIP. Define the following notation:

For product type \( r = 1, \ldots, R \):

\( \lambda_r \): Arrival rate of orders for type \( r \) products.

\( \alpha_r = \lambda_r / \sum_{r=1}^{R} \lambda_r \): Proportion of all orders over the planning horizon that are for type \( r \).

\( K_r \): Number of type \( r \) authorization cards in circulation (fixed WIP level for type \( r \)).

\( \omega_r(K_1, \ldots, K_R, \lambda_1, \ldots, \lambda_R) \): Proportion of orders for product type \( r \) that are not immediately satisfied.

\( \Omega \): Specified service level, i.e., the proportion of orders to be satisfied at once.

For a given product set with known manufacturing requirements and a given service level the problem of determining the
optimal number of authorization cards can be stated as problem (P).

\[
\begin{align*}
\text{Min } & N = K_1 + K_2 + \ldots + K_R \\
\text{(P) s.t. } & \omega_r(K_1, \ldots, K_R, \lambda_1, \ldots, \lambda_R) \leq 1 - \Omega, \forall r \\
& K_1, \ldots, K_R \geq 0, \text{ integer.}
\end{align*}
\]

While solving problem (P) is the goal, for several reasons we attack it indirectly. First, as highlighted in the notation, the proportions of orders that wait for fulfillment depend on the exogenous parameters \( \lambda_r \) as well as on the decision variables \( K_r \). When orders pull completed products from the system, the throughputs of individual products are limited by their order arrival rates. Secondly, to predict the values of \( \omega_r \), we must evaluate the performance of an open queuing network. On the other hand, by assuming high demand rates, we can model the system as a more tractable closed queuing network. Finally, in this closed network, assuming \( \theta_r = \alpha_r \Theta \) for \( r = 1, \ldots, R \), we can derive a theoretical maximum throughput, \( \Theta_{\text{max}} \), that can serve as the target for a card allocation heuristic.

For a solution procedure focused on throughput, define:

\[
\begin{align*}
\theta_r(K_1, \ldots, K_R) & \text{: System throughput of type } r \text{ products.} \\
\Theta(K_1, \ldots, K_R) & = \sum_{r=1}^{R} \theta_r(K_1, \ldots, K_R) \text{: Total system throughput.}
\end{align*}
\]

Then, given a proportion \( \beta < 1 \) of the theoretical throughput \( \Theta_{\text{max}} \), let \( \theta^*_r = \beta \alpha_r \Theta_{\text{max}} \) be the operating throughput level for type \( r \). Our goal is to determine card counts that enable the throughputs \( \theta_r^* \) of type \( r \) products in the closed network to achieve the targets \( \theta^*_r \):

\[
\begin{align*}
\text{Min } & N = K_1 + K_2 + \ldots + K_R \\
\text{(Q) s.t. } & \theta_r(K_1, \ldots, K_R) \geq \theta^*_r, \forall r \\
& K_1, \ldots, K_R \geq 0, \text{ integer.}
\end{align*}
\]

We designed a simple allocation procedure for solving problem (Q) [1]. An approximate extension to product-form queuing network analysis was used to evaluate the throughputs, \( \theta_r \). In tests on a small two-product system, this heuristic found optimal allocations of fixed total WIP levels, and was robust to different initial allocations. The WIP mix \( \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_R) \), where \( \gamma_r = K_r / N \), differed from the product mix, \( \alpha \), in some cases markedly. This WIP mix was then applied to solve problem (P) by fixing \( \gamma \), setting \( K_r = \gamma_r N \), and choosing the smallest \( N \) that satisfied the service constraint. The same queuing network analysis method was used to estimate the waiting probabilities, \( \omega_r \). In the small example system, this procedure found solutions very close to optimal (as found by exhaustive search). On a larger example with 5 product types processed on 10 stations, it also performed well compared with other card allocation heuristics.

**High Throughput with a Specified Product Mix:** The work described in the preceding section confirmed that uniform customer service across product types can be achieved by matching the throughput to the product mix. The WIP mix required to achieve the specified product mix depends on the processing requirements of the different product types and their queuing interactions at processing stations. In further work [2] we have developed a method for finding the WIP mix that applies to systems with fewer restrictions on the processing requirements. We specially adapted a performance evaluation technique and explored the impact of sequencing rules on the throughput and the WIP mix.
One rationale for using CONWIP control in a single product system is that WIP is easier to control than throughput. In multiple product systems, some researchers have suggested using a probabilistic or deterministic release policy to control the throughput mix while maintaining a constant level of total WIP. The system can be modeled as a single chain multiple class closed queuing network. The resulting observed WIP mix is highly dependent on the sequencing at each station. However, other research indicates that controlling individual WIP levels, as modeled by a multiple chain queuing network, leads to better performance. We use a single chain model to optimize the WIP mix but then apply it in a multiple chain control policy. The overall procedure is illustrated by Figure 1.

We adapted a linear programming method [3, 4] for evaluating the single chain queuing network. The model optimizes throughput subject to constraints on the total system population, “sampling equalities”, non-idling, stationary first and second moments of queue lengths, and nonnegativity restrictions. Maximizing and minimizing the objective function yields loose upper and lower bounds on the throughput. They can be tightened by adding linear constraints that embody preemptive priority sequencing. However, priority sequencing severely skews the WIP mix toward lower priority products. We feel it is unlikely to be adopted widely in practice. Instead, we developed nonlinear constraints that assume sequencing is independent of product type (for example, random or first-come-first-served). The nonlinear program provides increasingly tight bounds on throughput as the total system population (WIP) increases. The optimal WIP mix is extracted from the maximizing solution.

Tests on numerical examples indicate that the throughput bounds are quite tight. They encapsulate estimates obtained by simulation using a variety of sequencing rules, including priority rules. The examples demonstrate that priority sequencing has only a minor impact on the total throughput. When the WIP levels derived from the model are applied in a multiple chain control policy (i.e., with fixed WIP for each product type), the total throughput exceeds that for the single chain policy, and the throughput mix closely matches the specified product mix.

Conclusions: Planners often focus on high total throughput. But in a multiproduct system, the throughput must be balanced across product types to match the desired product mix. We have developed methods for finding a WIP mix that will achieve a specified product mix and have demonstrated that the resulting throughput mix fulfills orders for the various product types equitably. This work has also provided evidence that better overall system performance can be achieved by setting individual WIP levels than by maintaining a fixed total WIP and matching throughput to the product mix by order release.

References


Acknowledgement: This work was supported by NSF Grants DMI-9701403 (Ryan) and DDM-9400146 (Choobineh).

**Figure 1.** Procedure for planning WIP total and mix in the single chain model and applying them under multiple chain operational control.