Workload control concepts in job shops

Land, Martin J.; Gaalman, Gerard J.C.

Workload control concepts in job shops: A critical assessment

Martin J. Land and Gerard J.C. Gaalman

tel: +31-50-363 7188/7196
fax: +31-50-363 7207
email: M.J.Land@eco.rug.nl

SOM theme A: Structure, Control and Organization of Primary Processes

Abstract

In the case of production environments with job shop characteristics, much research has been done on partial control such as priority dispatching. The development of comprehensive control concepts lags behind. However, the principles of workload control (WLC) have been elaborated to more comprehensive production control concepts. WLC concepts buffer the shop floor against external dynamics by creating a pool of unreleased jobs. The use of workload norms should turn the queueing of orders on the shop floor into a stationary process which can be characterised by an equilibrium.

This paper compares and discusses the concepts of WLC. Assumptions of stationarity implied in the workload norms are exposed. A subdivision of workload definitions is chosen as a starting-point to trace assumptions of stationarity. The assumptions highlighted relate to the shop floor situation and make demands upon the job release function. An obvious conflict between timing and balancing within the job release function leads to an examination of stationarity requirements on the job pool contents.

The analysis of stationarity requirements within existing production control concepts provides guidelines for developing production control concepts for job shops working under dynamic circumstances.

1. Introduction

Traditionally, the job shop is a type of production environment which can be found in mechanical industry, particularly in component manufacturing. More recently, semiconductor fabrication has led to job shop situations. Job shops are characterised by a wide variety of products with variable routings and processing times. Job shops have a functional layout with universal equipment. Production takes place according to customer specification and in small batches.

Typical job shops have to work under very dynamic circumstances, both internally and externally. External dynamics relate for instance to rush orders, the product mix and volumes demanded, while internal dynamics may relate to machine breakdowns, production rates, operator absenteeism, quality problems, production yields, etc. We call a modelled job shop dynamic if the probability distributions, which describe the variables, are non-stationary and change in the course of time.

The bulk of literature on job shops has been devoted to priority dispatching. Surveys show hundreds of priority rules to be applied on the shop floor *[Panwalkar & Iskander, 1977; Blackstone et al., 1982; Ramasesh, 1990]*. Another research field receiving much attention is the assignment of due dates *[e.g. Cheng & Gupta, 1989]*. Generally, research on priority dispatching and due date assignment does not consider comprehensive production control concepts, but isolates single elements of production control. The development of comprehensive control concepts still lags behind *[Hendry & Kingsman, 1989]*.

A starting point in the development of more comprehensive concepts has been the introduction of input/output control, first introduced by Wight *[Wight, 1970]*. Since then input/output control has been extended to a class of hierarchical capacity-oriented production control concepts for job shops *[Bertrand & Wortman, 1981; Tatsiopoulos, 1983; Bechte, 1988]*. These hierarchical concepts control workload, both at the level of order entry and the level of order release to the shop floor. The former level relates to all planned/accepted jobs, the latter relates to jobs on the shop floor. The control of workload on the shop floor creates a backlog/pool of orders waiting for release.

The pool is claimed to buffer the shop floor against external dynamics. With this claim, the class of production control concepts using workload control might be attractive for use in job shops which are subject to dynamic circumstances.

This paper assesses how workload control (WLC) concepts deal with the dynamics of job shops. Comparing existing WLC concepts, we expose underlying assumptions of stationarity and corrections for violated stationarity assumptions. In order to compare the concepts we consider the classical job shop model, consisting of a set of work stations, each station concerning one specific capacity type, required for one specific operation on a job. We do not restrict ourselves to the pure job shop, the common model in most simulation studies. That means, capacities of work stations are not necessarily balanced, and job routings are not completely random with equal probability for each work station to be visited in each stage of job progress.

Section 2 elaborates the WLC paradigm. The analysis of three WLC concepts from the release point of view in section 3 leaves us with three different workload definitions worth further investigation. It appears a useful starting point for our assessment in section 4, as the definitions and the corresponding workload norms expose the stationarity assumptions relating to the shop floor situation. The release function must provide for the stationary workload on the shop floor. Section 5 discusses the obvious conflict between a timing and a balancing function of job release which lead to an exposure of stationarity requirements relating to the job pool. We summarise our analysis for the three referenced WLC concepts by means of a table.

2. The workload control paradigm

WLC conceptualises the job shop as a queueing system. In front of each work station, an arriving job finds a queue of jobs waiting to be processed. The principle of WLC concepts is to control the length of these queues. The main instrument for this purpose is the release decision. The release decision allows a job to enter the queue of its first work station in the shop. Once released, a job remains on the floor until all its operations have been completed. The progress of jobs on the shop floor is controlled by priority dispatching at each work station.

WLC concepts do not release jobs to the shop floor if they are expected to cause queue lengths to exceed certain workload norms. It results in a *pool* of jobs waiting for release. As illustrated by figure 1 we refer to waiting time in the pool as the *pool time* and to the interval between release and completion of a job as the *shop floor flow time*. The shop floor flow time of a job can be subdivided into *station flow times*. The pool is a new object of control. Unrestricted acceptance of jobs at the entry could cause excessive pool times.

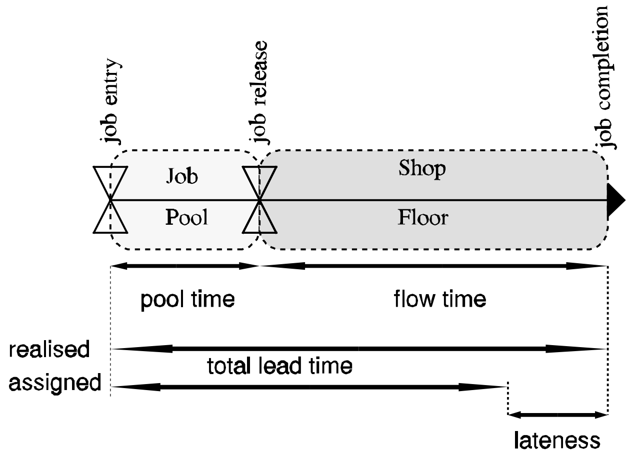

figure 1: lead time components

A hierarchical control concept emerges *[Kingsman et al., 1989]*, with three levels which respectively relate to job entry, job release and priority dispatching (figure 2). At each level, we distinguish two means of control, input control and output control. Input control regulates the allowance of jobs to the next stage, respectively accepting jobs for entry into the pool, releasing jobs to the shop floor, and dispatching jobs for processing (thus allowing a job to enter the queue of its next operation). On the output side, capacity management contributes to the control of workload through regulation of the outward flow, by means of respectively medium-term, short-term and daily capacity adjustments *[e.g. Park & Bobrowski, 1989]*. In addition due date assignment or due date acceptance takes place at job entry. This paper concentrates on the input side.

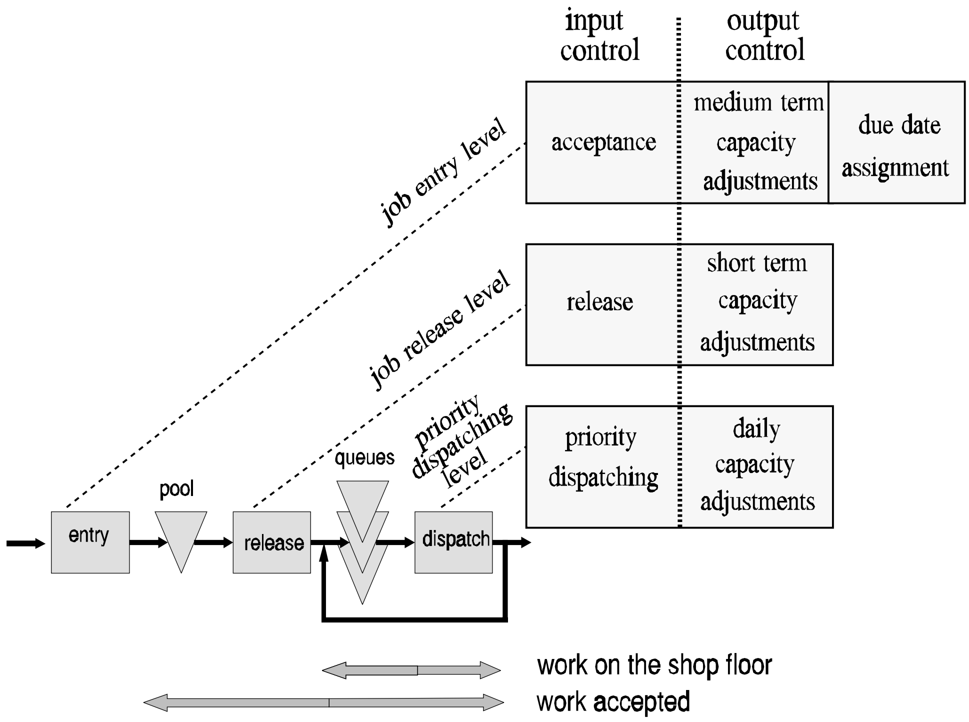

figure 2: The hierarchical WLC concept

The job entry level is very important, if one can influence the incoming orders. In that case, order acceptance and due date assignment/acceptance can support the release decision, providing it with a ’releasable’ set of jobs, thus keeping pool times small. In fact, the job pool between entry and release acts as the visualised imbalance between job supply and production capacities.

The role of priority dispatching in WLC is a very modest one, because the choice among jobs is limited due to short queues. Generally, WLC concepts favour dispatching priorities such as first-come-first-served (FCFS) which stabilise operation flow times or due date oriented priorities which correct progress differences among jobs. These kinds of priorities facilitate a good timing of job release.

At the release level the use of workload norms controls the work station queues. The control of queue lengths, resulting in short and predictable flow times, is the key to both lead time and due date performance *[Bertrand, 1983]*. However, the major strength of WLC concepts is withholding jobs from the shop floor, reducing average queue lengths. Besides a reduction of work-in-process, withholding jobs from shop floor has numerous additional advantages as it enables management to delay final production decisions *[Irastorza & Deane, 1974]*. It reduces waste due to cancelled orders, facilitates later ordering of raw materials, takes away the need of expediting of rush orders, etc. Fluctuations in the incoming order stream should be absorbed by the pool. Altogether, it should create a stable stationary situation on the shop floor.

Only restricting queue lengths is generally not sufficient. If average queue lengths decrease but variances do not, the idle time at work stations will increase. This situation is not allowable for the common job shop, where many work stations can be temporary bottlenecks. The loads of potential bottlenecks should be kept close to a norm level instead of below a norm level. The release function which aims at short queue lengths and a reduced variability of queue lengths is called *load-balancing* within this paper.

Simulations of release rules with limited balancing qualities often show a deteriorated lead time performance. This has made the influence of 'controlled release' a topic of scientific research *[Melnyk et al., 1991; Kanet, 1988]*. In practice, WLC concepts prove to have a positive effect on lead times *[Wiendahl, 1992]*, a result often attributed to improved 'shop floor transparency'.

In summary, WLC concepts try to create a situation on the shop floor of short and stable queues. A pool of unreleased jobs buffers the shop floor against external dynamics, the incoming non-stationary job stream. The queueing of jobs on the shop floor is turned into a stationary process. Release performs a key-role in reaching this stationary situation. It is the most elaborated function within WLC concepts. Therefore, we will compare and assess existing WLC concepts from the release point of view.

3. Existing WLC concepts

In the preceding section we have seen that release should control the queue lengths in front of each work station. The queues must be short and stable, the load-balancing function. On the other hand, each job should be released timely with respect to its planned due date and expected flow time, the *timing* function.

Leaving out capacity decisions at the release level, two components of the release decision are distinguished: a *sequencing* decision and a *selection* decision. The sequencing decision can be described as the setting of priorities for jobs to be released, 'selection' decides whether a job will be released or not at some specific moment. Most WLC concepts focus the sequencing decision on timely release and create due date based sequences. Taking into account this sequence, release selects a set of orders that keep the workload of work stations at certain norms. These workload norms are the main instrument of workload control.

For three WLC-concepts we discuss the release decision, the workload definition applied, and the determination of the corresponding workload norms. In addition, we present some developments which provide new ideas for release procedures within WLC.

Bechte's WLC concept

The release procedure proposed by Bechte *[Wiendahl, 1987, 1995; Bechte, 1988, 1994]* builds on three parameters: a *release period*, a *time limit* and a *load limit*. The decision to release jobs is taken periodically, at the beginning of each *release period*. All jobs in the pool are sequenced in order of their *planned release date*. The planned release date is determined by backward scheduling from the job due date: *norm station flow times* for all work stations in the routing of the job are subtracted from its due date. All jobs within the *time limit* from their planned release date are candidates for release. In the established sequence, jobs are released, until the workload norm of a work station, the *load limit,* is exceeded for the first time. All other candidates visiting this station have to wait in the job pool until the next moment of release. The selection process goes on for the remaining candidates.

The workload considered in the concept of Bechte is the queue length at a work station (in units of processing time). The workload is controlled by the load limit. The load limit LL_s of a work station s consists of two components: the planned output during the release period and the planned queue length at the end of the release period. The actual output O_s during the release period and the actual queue length Q^E_s at the end of the release period satisfy the balance equation:

$$Q^E_s + O_s = Q^B_s + I_s$$

with Q^B_s: the queue length at the beginning of the release period

I_s: the input to the queue from jobs arriving during the release period

The release decision at the beginning of the release period must bring $Q^E{}_s + O_s$ at the norm level LL_s. The above balance equation is used. $Q^B{}_s$ is known at the moment of release, the queue input I_s is influenced by the jobs on the floor upstream of s and by the release of new jobs, see figure 3. Some of them will arrive at s, some will not. Bechte estimates the input during the release period by means of the *load conversion algorithm*:

> If the workload of station x with a planned output component PO_x reaches its limit LL_x, a fraction PO_x/LL_x of the workload is planned to pass the station. Therefore, the probability that job j in the queue of x passes station x during the release period is estimated by PO_x/LL_x. The probability that job j reaches the queue of work station s is the probability that job j passes all its remaining upstream stations (the set U_{js}). This probability Pr_{js} is estimated by the product $\prod_{u \in U_{js}} \left(\frac{PO_u}{LL_u} \right)$. Suppose the processing time of job j at station s is p_{js}, then the expected input to the queue of station s from all upstream jobs (the set J_{us}) is estimated by $\sum_{j \in J_{us}} Pr_{js} \cdot p_{js}$.

First, load conversion is applied to estimate the input to the queue from jobs actually on the shop floor. Next, new jobs are released and their input is estimated until the estimated workload of a work station reaches its load limit. Notice that, within the load conversion procedure, the actual upstream positions of jobs at the time of release have been taken into account.

The workload norm LL_s is derived from the norm station flow time $NSFT_s$ of the work station. It assumes the following steady-state relationship *[Wiendahl 1987, 1995]*: $LL_s = PO_s + \frac{PO_s}{T} \cdot NSFT_s$[1], with T: the length of the

[1] Here, $NFST_s$ is not determined as the norm station flow time for a job but for a unit of processing time

release period. The next step should be the determination of realistic norm station flow times, as they are essential elements of both workload norms and planned release dates. Nyhuis *[Nyhuis, 1992]* presents a theoretical approach to estimate realistic norm values for this concept. Till now, trial-and-error determination, step-wise lowering norms, seems to be most successful for practical situations.

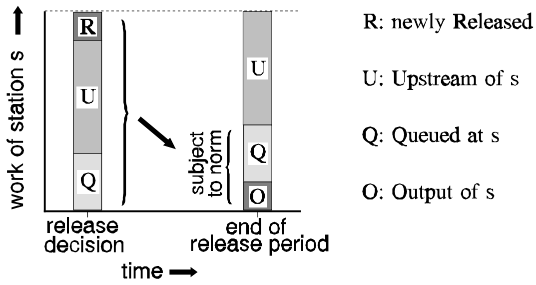

figure 3: Bechte: workload at the end of the release period estimated at the time of release

Bertrand's WLC concept

Bertrand developed a WLC concept for the diffusion department of a semiconductor plant *[Bertrand & Wortmann, 1981]*. Bertrand does not discuss the release sequence, but elaborates the workload norms extensively. The release decision is taken periodically and the release of jobs is allowed if the workload of each work station remains below its norm value.

The workload considered in this WLC concept differs from the workload considered by Bechte. The workload definition of Bertrand covers the processing time of all jobs on the shop floor which still have to be processed at the work station concerned. The corresponding workload norm consist of two components: the planned work station output during the release period and the planned quantity of work upstream or in the queue at the end of the release period. An extended balance equation can be used to determine the actual workload of a work station s at the end of the release period:

$$(U^E_s + Q^E_s) + O_s = (U^B_s + Q^B_s) + R_s$$

with U^E_s: the processing time (on s) of jobs upstream at the end of the release period

U^B_s: the processing time of jobs upstream at the beginning of the release period

R_s: the processing time of jobs released at the beginning of the release period

All other variables as defined before.

At the moment of release the right-hand side of this equation is completely known. The processing times of all jobs which are newly released are the input to the workload. Thus, the release of new jobs directly influences the workload (see figure 4). The release decision can be made without a sophisticated estimation procedure.

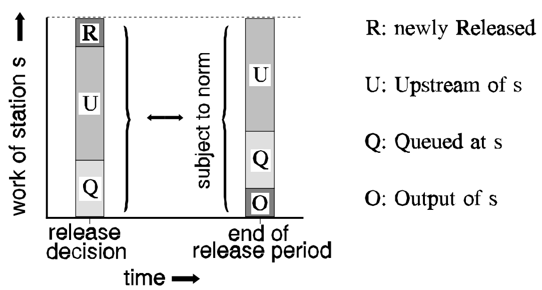

figure 4: Bertrand: the workload subjected to the norm is known exactly upon release

The determination of correct norm values is more complex for this workload definition. It is especially difficult to determine an accurate norm for the quantity upstream at the end of the release period. Bertrand calculates the norm as follows:

> If the flow time of a job j at a work station s equals the norm station flow time $NSFT_s$, j will queue at this work station for $NSFT_s$ time units. According to the workload definition, job j will be part of the workload of station s during its stay at stations upstream of s (the set U_{js}) as well. So, the norm pre-station flow time of job j is $NPFT_{js} = NSFT_s + \sum_{u \in U_{js}} NSFT_u$. As long as job j is part of the workload, it increases the workload by its processing time p_{js}. In the course of time job j will contribute $NPFT_{js} \cdot p_{js}$ to the cumulative workload of s. Now, Bertrand uses a set of jobs J which is supposed to be a good representation of the total population of jobs. All jobs of J together create a cumulative workload of $\sum_{j \in J} NPFT_{js} \cdot p_{js}$. If the average output per release period of length T equals the planned output PO_s, it takes $\frac{\sum_{j \in J} p_{js}}{PO_s}$ periods or $\frac{\sum_{j \in J} p_{js}}{PO_s} \cdot T$ time units to process all the jobs of J. The planned average workload during this interval will be $\frac{PO_s}{T} \cdot \frac{\sum_{j \in J} NPFT_{js} \cdot p_{js}}{\sum_{j \in J} p_{js}}$, which completes the calculation of the second norm component. It appears to be the product of the planned utilisation level and a weighted average of the norm pre-station flow times. Finally, Bertrand adds the planned output component and sets the norm to

$$PO_s + \frac{PO_s}{T} \cdot \frac{\sum_{j \in J} NPFT_{js} \cdot p_{js}}{\sum_{j \in J} p_{js}}.$$

In principle, this norm calculation applies to all work stations. For low-utilised stations a workload slightly higher than this norm is allowed if the actual job mix gives reason to it. Here, the effect on flow times will be small. Notice that the norm value calculated increases with the number of upstream stations. Roughly speaking, the norm value depends on the average work station position within *J*. The norm accounts for the work station position within a presupposed set of jobs. The release decision does not use information on the actual upstream positions of jobs at the moment of release. As for Bechte, the determination of realistic norm station flow times is open to question.

Tatsiopoulos' WLC concept

Tatsiopoulos *[Tatsiopoulos, 1983]* developed a WLC concept for a small subcontracting component manufacturer. The concept has been elaborated by Kingsman and Hendry *[e.g. Kingsman et al., 1989]*. The concept formalises three ways of job release [Hendry and Kingsman, 1991]. The common *push release* takes place periodically, *intermediate push release* can be forced by rush orders or orders with retarded material availability, and an *intermediate pull release* can be triggered from the floor when a foreman sees his station threatened by unplanned idleness. The periodic release decision considers the orders in the sequence of their planned latest release date. The calculation of the planned release dates is rough compared with Bechte. For each job the same norm shop floor flow time is subtracted from the job due date. The release of jobs is allowed unless a workload norm is exceeded, which applies to the intermediate pull releases as well. Additionally, a minimum workload is suggested. Unfortunately, both the use and the calculation of the minimum norm are not further elaborated.

Commonly, this WLC concept applies the same extended workload definition as the concept of Bertrand. We restrict ourselves to another workload definition, applied in the WLC system implemented by Tatsiopoulos *[Tatsiopoulos, 1983]* and also mentioned in *[Hendry & Kingsman, 1988]* and *[Tatsiopoulos, 1993]*. This definition covers all work on the shop floor, even work completed at the work station concerned. For each work station a norm is set for the accumulated processing times of jobs upstream, job in the queue, and jobs downstream. The corresponding actual workload satisfies the following balance equation (see figure 5):

$$(U^E_s + Q^E_s + D^E_s) + C_s = (U^B_s + Q^B_s + D^B_s) + R_s$$

with D^E_s: the processing time (on *s*) of jobs downstream at the end of the period

D^B_s: the processing time of jobs downstream at the beginning of the period

C_s: the processing time of jobs which leave the shop during the release period

All other variables as defined before.

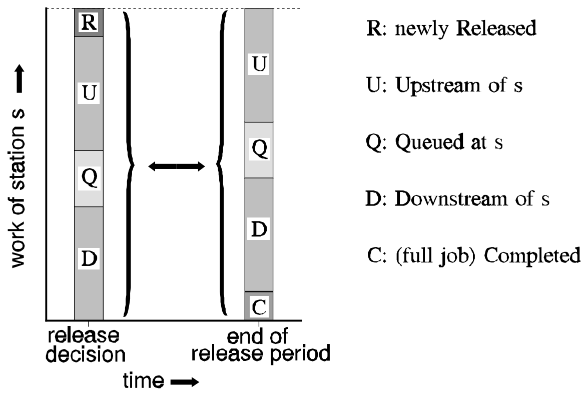

figure 5: Tatsiopoulos: a broad workload definition including downstream work

Again all right-hand side components are known at the moment of release. The WLC concept does not clarify whether the shop output C_s from jobs fully completed during the release period is included in the workload norm. Notice that the workload definition further simplifies keeping up with the actual workload as it avoids the need for data regarding the completion of single operations. The completion of the job can be reported when it leaves the shop floor.

Hendry and Kingsman suggest that the workload definition enables the use of the same norm value for each work station *[Hendry & Kingsman, 1988]*. This norm value is proportional to the maximum acceptable shop floor flow time.

Other methods of controlled release

Other ideas for release procedures within WLC are provided by *[Glassey & Resende, 1988]* and research of Wein *[Wein, 1990; Wein, 1992; Wein & Chevalier, 1992]*. Both suppose continuous release opportunities. As a consequence, job release takes place whenever a workload falls below its norm, instead of periodic replenishments. Both studies assume explicitly that all random variables are stationary.

The *starvation avoidance* policy of Glassey & Resende focuses at the avoidance of idle time at a bottleneck station. The policy is only elaborated for the very simplified situation of a flow shop with only one job type and one bottleneck station. However, their workload definition is an interesting contribution to the spectrum we recognised. It includes the processing time of a job in the workload of the bottleneck station, when the job's remaining processing time upstream is below a *critical time factor*. Thus, jobs in the queue and part of the jobs upstream of the work station are included.

Wein applies the same workload definition as Bertrand, the accumulated processing times of jobs upstream and in the queue of the work station. The difference between the concept of Wein and the previous ones can be found in the release sequence and the use of norms. The release procedure is elaborated for a situation with two work stations. Wein combines norms for the absolute workloads of both stations with norms for the ratio between the workloads. Figure 6 graphically depicts the workload conditions which require new releases. The shape of the area requiring new releases differs from the common rectangle area which results from absolute norms. The cut resulting from the ratio norms represents the principle that a better ratio between the workloads allows for lower workloads. Wein primarily sequences the jobs in order of their due date. But, when the difference between two due dates is below a certain limit, priority is given to the job which best restores the workload ratio between the stations, that is the job which contributes most to the smallest workload. Thus, the control policy manipulates the shop situation in the direction of the internal corner *c* of the shaded region in figure 6. Point *c* represents the smallest combination of workloads required.

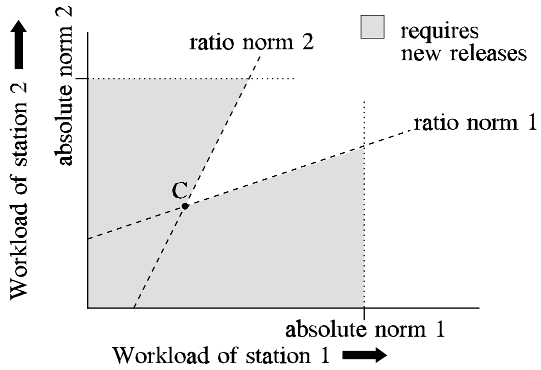

figure 6: Workload conditions requiring release in the concept of Wein.

4. Workload definitions and shop floor stationarity

The discussion of WLC concepts highlights both differences and similarities between WLC concepts. Especially the use of different workload definitions and corresponding norms is worth further investigation. All WLC concepts use norms for the quantity of work allowed on the shop floor. As jobs are released periodically, norms are set for the desired situation at the end of the release period. More precisely, norms are set for each work station on the shop floor, as the WLC concepts aim to control the queue length in front of each work station. Though principally the objective is to control the load in the queue of each station, we observe the use of extended workload norms. Bertrand includes the work content of work upstream and Tatsiopoulos all work on the shop floor, both upstream and downstream (see figure 7).

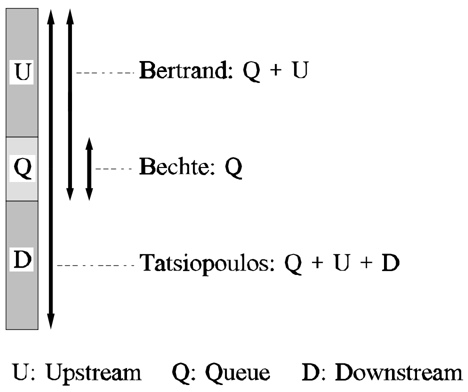

figure 7: The subdivision of workload definitions

The reasons for extensions stem from practical perspectives. Our discussion of the Bechte concept highlights that restricting the workload to the work station queue requires estimations of the input to the queue. Since all jobs upstream are candidates, a workload that includes all upstream work eliminates the need of input estimations. Upon release, a job contributes immediately to the workloads. If downstream work is also incorporated, the workload of each station will only change at job releases and job completions, and the release decision no longer requires a record of operations completed at each work station.

If an extended workload norm should control the load in a queue during the release period, assumptions will be inevitable. For each type of workload addressed by the workload norm we expose the underlying assumptions:

Bechte: queue only

In general, the contents of the work station queues are not directly influenced by the release of jobs. Jobs may have to pass other (upstream) stations first. Since the arrival of jobs is influenced by many uncertain factors, simplifying assumptions are necessary to obtain a simple estimation procedure. The WLC concept of Bechte accounts for the actual upstream positions of jobs at the time of release. So contrary to the other concepts, this concept does not make assumptions about these positions. The assumptions of Bechte are restricted to the flow of jobs during the release period as it is estimated by load conversion, and to the actual volume of the workloads upon release. The load conversion procedure evoked a number of criticisms, criticisms for the larger part published in German literature *[Adam, 1988; Adam, 1989; Häfner, 1992; Hansman, 1993; Knolmayer, 1991; Greiner, 1989]*. Without going into detail, we might say that most of these criticisms relate to the assumption of unrestricted divisibility of the workload:

1) The estimated probability PO_x/LL_x that a job passes a station x neglects the fact that each job as a whole must pass the station during the release period and not a fraction of its processing time. If its processing time is large, the actual probability will decrease.

2) The product form of Pr_{js} suggests that the probability to pass a station is independent of the number of stations already passed during the release period, though each operation will delay a job, at least for its processing time. The probability that a specific job reaches a station might be estimated more accurately by considering its planned pre-station flow times in relation on the length of the release period.

If the workloads are large relative to processing times, the estimations show increased accuracy. Ambiguously, the paradigm of workload control forces the workload in the opposite direction. Since it aims at small workloads, workload control will lead to a more restricted divisibility. We observe other assumptions which relate to the volume of the actual workload implied in the estimated probability PO_x/LL_x:

1) The actual output O_x of the supplying stations should equal the planned output PO_x.

2) The actual workloads estimated from $Q^B{}_x + I_x$ should equal the norm LL_x.

Though the assumptions may seem obvious, the fact that they should hold for each work station, imposes strong requirements to the set of jobs on the floor.

Bertrand: queue and upstream included

It has been shown that the inclusion of upstream work in the workload norm avoids the need of predicting queue inputs at the moment of release. However, the easier determination of the actual workload at the moment of release rebounds upon a more complicated determination of the workload norm. It is possible to disaggregate Bertrand's second norm component into an element for the load upstream and an element for the load in the queue. The actual workload elements $U^E{}_s$ and $Q^E{}_s$ should correspond with their norm parts.

Otherwise, the queue might be idle while the workload is at its norm level. So, a stable composition with respect to the shares of the upstream and the queue elements of the workload is assumed. Bertrand does not check this assumption at the moment of release.

These general considerations about the workload composition can be elaborated in more detail. The calculation of the workload norm reveals the detailed assumptions. The norm is calculated for a specific set of jobs *J* which is supposed to be representative for the future portfolio of jobs. One might wonder what happens if the portfolio on the shop floor is going to differ from *J*. The calculation shows that the workload norm of a station *s* increases with the number of (upstream) stations to be visited before *s*. If the jobs on the shop floor visit *s* later on average than the jobs of *J* do, the actual load upstream should exceed its norm component in order to provide *s* with the planned queue load. The actual upstream positions of jobs within the workload of *s* are not checked. Thus, one must assume that the relative position of each work station within the actual job mix on the floor varies little, and that on average it equals its position within the presupposed set of jobs *J*. More exactly, the assumption relates to the pre-station flow time characteristics of the jobs on the floor. The actual mix of pre-station flow times weighted by the processing times of jobs should be stationary, as the calculated norm component $\frac{PO_s}{T} \cdot \frac{\sum_{j \in J} NPFT_{js} \cdot p_{js}}{\sum_{j \in J} p_{js}}$ shows.

Bertrand corrects the calculated norms. The correction allows for exceeding the norms of low-utilised stations. It provides room for deviations from the norm mix *J*. However, we note that the corrections only seem suitable for small deviations around the means. They will not be adequate for a shift of the mean or heavy incidental disturbances.

The idea of Glassey and Resende to control the quantity of work within some time distance from a work station adopts a middle course between the workload definitions of Bechte and Bertrand. As within the concept of Bertrand, the workload is an aggregate of the actual queue contents and future contents, and no estimation of the queue input is made. In contrast with Bertrand, Glassey and Resende account for the upstream position for jobs at the time of release, but in a less detailed way then Bechte. Since the release of a job only affects the workload of its first work stations directly, one should assume that the set of jobs released provide their downstream stations with a stable future load. The non-periodic release decision enables a fast reaction to load deviations.

Tatsiopoulos: queue, upstream and downstream included

Control of the work upstream of a work station *s* bears importance for the control of the work station queue, as the work upstream incorporates the future queue load of station *s*. Including work downstream of *s* in its workload norm does not contribute to the control of the queue. The jobs concerned already passed the queue of *s*. Any fluctuations of the actual downstream component will needlessly influence the release decision. An actual downstream component exceeding its corresponding norm component causes the decision to slow down the release of jobs. As a consequence, the work station considered may get idle even when there is plenty of work to be released. Thus, one should assume that on average the actual downstream component equals the norm component and that its fluctuations are be limited.

Violations of this assumption may have serious consequences. Since the downstream component will be relatively large for a station performing preparatory operations (i.e. a gateway station), these stations are particularly threatened by (unnecessary) idleness. Tatsiopoulos obviates this problem by providing gateway stations with the possibility of pull-release.

The workload norm no longer depends on the average position of a work station in the mix of jobs. No assumptions have to be made regarding the work station position in the actual set of jobs on the floor. A shift of the average work station position will have no effect in the long term.

In general, the WLC concepts must realise a stable input to the work station queue in order to control its length. Independent of the type of workload norm, it is in general not possible to influence this input directly by the release of jobs. One depends on assumptions regarding the input to the queue. First, the output rate of supplying stations is assumed stable. Stationarity of work station capacities is a prerequisite for stable output rates. Second, the mix of jobs released must have stationary characteristics. Bechte accounts for the actual upstream positions of jobs to estimate the input during the release period and his assumptions are restricted to the load conversion estimation procedure. Bertrand and Tatsiopoulos do not check the actual positions of jobs at the beginning of the release period. They assume that the extended workload which is subjected to their norms provides the planned input to the queue. These assumptions will be violated if the characteristics of the actual workload on the floor differ from the characteristics supposed within the norm calculation. All WLC concepts use a relationship between workloads and planned flow times. These relationships only hold in a stationary situation and with the assumption of all work stations loaded up to their workload norm. In summary, we may say that stationarity of characteristics is assumed for both jobs and capacity on the shop floor.

Realising the stationarity of capacity characteristics goes beyond the span of control of WLC concepts, realising stationarity of characteristics for the jobs within the workload is claimed to fall within. Till now we did not address the question whether, even if the characteristics of the actual workload have been checked, a release policy will be able to provide the work stations with the required stationary workload. The next section assesses the ability of the release policies to create the required workload.

5. The timing/balancing conflict and pool stationarity

The preceding section pointed out that all WLC-concept make assumptions which relate to mix of jobs on the shop floor. Since the release decision should provide the shop floor with this mix, the assumptions impose requirements on the release decision. A minimum requirement within all WLC-concepts is that the volumes of the actual workloads upon release equals the workload norms.

Stable workloads, equal to norm values, should be guaranteed by the load-balancing function of release. Only if load-balancing functions well, the queues of work stations will be stable. Stable queues should keep flow times at their planned level. Planned flow times in turn determine the planned release date of a job. So, a precise timing of the release moment of a job depends on stable flow times. As a consequence, this timing-function of release depends on an effective load-balancing function to realise a good due date performance. The question is whether a good timing of job release also allows for sufficient load balancing. By assessing the release procedures of the WLC concepts, we will determine under which conditions the load-balancing and timing function co-operate.

We argue that the referenced concepts deal with order release in a one-sided way. Accurate timing is provided by the sequence in which jobs are considered for release. But, the load-balancing qualities are limited. The release procedures fit jobs into the workload in the predetermined sequence of planned release dates. Once a job fits, its release will not be reconsidered. This can be seen as a greedy algorithm. As a result, some workloads might be far below their norm, because the workload of one station reaches its norm. The release sequence could have been reconsidered in order to approximate the complete set of norm values more closely. In particular if the accepted order portfolio requires high utilisation levels, WLC-concepts may require better balancing properties to create sufficient throughput capability *[Land & Gaalman, 1994]*. An example of a completely balance-oriented approach is presented by *[Shimoyashiro et. al, 1984]*.

The release policy of Wein can be a first step in the development of more powerful release policies. It shows better balancing properties than the release policies of the referenced WLC concepts and carefully weighs balance requirements against job due dates. The policy does not require each station to be loaded up to a fixed norm. Instead it allows small fluctuations of the ratio between workloads, but the better this ratio, the smaller the volumes of workload required.

The job pool makes the balancing function less sensitive to the dynamics of the incoming order portfolio. A larger pool increases the choice of jobs to fill workload gaps. That way, the capacity requirements of the incoming stream of jobs are smoothed by the pool. It depends on the size of the pool to what extent fluctuations can be absorbed. However, a larger pool increases pool times and deteriorates lead time performance. Thus, lead time requirements restrict the size of the pool. This restriction may create a conflict between the load-balancing and timing function of release. At a certain moment jobs require release according to their planned release date. If the set of jobs requiring release do not fit into the workload norms, jobs will be delayed until the next moment of release and due date performance will deteriorate. Only if the load contribution of the job set requiring release does not show excessive peaks for any work station, conflicts between the timing and balancing functions can be avoided. So, a certain stationarity of the job pool contents must be required. Melnyk et al. *[Melnyk et al., 1992]* discern the same problem. Their simulation results indicate a more effective release, when release is preceded by smoothing of the workload.

Only small fluctuations around stable means can be absorbed by the pool. Strong dynamics (instable means, etc.) related to the incoming order portfolio will not create sufficient stationarity within the job pool. Existing WLC-concepts confronted with strong dynamics of the incoming stream of orders will depend on either high flexibility of capacity or possibilities to reject stationarity disturbing orders at the entry level. Till now, output control and order acceptance have been the least elaborated elements of the WLC concepts. An exception should be made for recent research on order acceptance by Hendry and Kingsman *[Hendry & Kingsman, 1993]*.

Of course, norm values can be adjusted continuously in dynamic situations. With the help of linear programming techniques, Zäpfel and Missbauer *[Zäphel & Missbauer, 1993]* determine new norm values, whenever dynamics of the incoming order stream give rise to this. Even if adequate determination of optimal norm values is possible, this will lead to cumbersome and nervous procedures for job shops which are exposed to strong dynamics.

During a short time interval, we might assume a stationary situation. Even then, it is still questionable whether the actual workloads must be exactly adjusted to a norm value upon release. Also in a stationary situation, workloads fluctuate without deteriorating performance: reacting to each deviation from the norm might lead to over correction. Instead, we might release constant quantities of work and only correct these quantities for fluctuations that exceed some ’normal-variance-based’ bounds. Such bounds may be able to handle an increased range of dynamic fluctuations without causing over correction, as the norm adjustments of Bertrand aim at reducing over correction of small load fluctuations. The above approach has proven its value in the field of statistical quality control.

6. Conclusions and suggestions for further research

WLC concepts buffer the shop floor against external dynamics by creating a pool of unreleased jobs. The use of workload norms should turn the queueing of jobs on the shop floor into a stationary process. Here, the release decision performs a key-role. WLC concepts translate the term ’control’ to ’maintenance of workload norm levels’.

However, each type of workload norm brings about a series of stationarity assumptions. Roughly speaking, WLC concepts assume stationarity of the shop floor situation. They depend on a certain stationarity of the job pool contents to create this stationary situation. Otherwise, the release decision will be confronted with conflicts between its load-balancing and its timing function.

Though a large pool buffer may protect the shop floor against external dynamics, it puts high pressure on lead times. Consequently, the WLC concepts correct for violations of internal stationarity assumptions, adjusting norms before release, or afterwards with intermediate releases. Table 1 summarises the different workload norms, the assumptions and the formalised corrections.

The question arises whether all stationarity assumptions are necessary. Might it be possible to incorporate the reactions to dynamics in the frame of the control concepts? Continuous adjustment of norm values is a cumbersome procedure, since even the determination of accurate norm values is a complex decision, not yet crystallised. The many job shops exposed to strong dynamic circumstances require control concepts that handle dynamics in a more natural way. This provides an interesting domain for further research.

Even, if we suppose temporary stationarity, the existing WLC concepts, with their continuously changing release quantities, neglect the normal variability of stationary characteristics. Statistical quality control has embraced control concepts that only react to excessive variability or shifting means, the real out-of-control situations. Statistical production control concepts like workload control, might gain applicability by adopting this approach.

	Bechte	Bertrand	Tatsiopoulos
workload subjected to norm (for release period T):	station queue at the end of T	station output during T + queue + upstream work at the end of T	(shop floor output during T) + queue + upstream work + downstream work at the end of T
feedback frequency:	each completed operation	each completed operation	each completed job
norm determination	directly derived from norm station flow times	derived from presupposed job mix and their norm pre-station flow times	equal to maximum norm shop floor flow time
upstream position of jobs versus work station position:	actual upstream positions of jobs used for estimation of queue input; workload norm independent of work station position	actual position not checked; workload norm depends on position of work station within job mix	actual positions unknown; workload norm independent of work station position
assumptions:	- each station loaded up to its norm; - simplifying assumption to estimate queue inputs	- each station loaded up to its norm; - smooth queue input from upstream jobs; - actual job mix corresponds with pre-supposed mix with respect to pre-station flow times	- each station loaded up to its norm; - smooth queue input from upstream jobs; - stable downstream workload component
corrections:	no formalised corrections	norm adjustment allowing small load fluctuations for low-utilised stations	intermediate pull-release for idling work stations

Table 1: Analysis of the concepts of Bechte, Bertrand and Tatsiopoulos

References

Adam, D; (1988); **Die Eignung der belangsorientierten Auftragsfreigabe für die Steurerung von Fertigungsprozessen mit diskontinuierlichtem Materialfluß** (in German). Zeitschrift für Betriebswirtschaft, Vol 58, No 1, 98-115.

Adam, D; (1989); **Probleme der belastungsorientiertien Auftragsfreigabe** (in German). Zeitschrift für Betriebswirtschaft, Vol 59, No 4, 443-447.

Bechte, W; (1988); **Theory and practice of load-oriented manufacturing control**. International journal of production research, Vol 26, No 3, 375-395.

Bechte, W; (1994); **Load-oriented manufacturing control, just-in-time production for job shops**. Production planning & Control, Vol 5, No 3, 292-307.

Bertrand, JWM; (1983); **The use of workload information to control job lateness in controlled and uncontrolled release production systems**. Journal of operations management, Vol 3, No 2, 79-92.

Bertrand, JWM; Wortmann, JC; (1981); **Production control and information systems for component-manufacturing shops**. Dissertation, Elsevier, Amsterdam, The Netherlands.

Blackstone, JH; Phillips, DT; Hogg, GL; (1982); **A state-of-the-art survey of dispatching rules for manufacturing job shop operations**. Int. journal of production research, Vol 20, No 1, 27-45.

Cheng, TCE; Gupta, MC; (1989); **Survey of scheduling research involving due date determination decisions**. European journal of operational research, Vol 38, No 2, 156-166.

Glassey, CR; Resende, MGC; (1988); **Closed-loop job release control for VLSI circuit manufacturing**. IEEE transactions on semiconductor manufacturing, Vol 1, No 1, 36-46.

Greiner, T; (1989); **Belastungsorientierte Auftragsfreigabe - nicht immer die beste Lösung** (in German). AV 26 (1989) 2, 65-68 Carl Hanser Verlag, München.

Häfner, H; (1992); **Ein Warteschlangenansatz zur integrierten Produktionsplanung** (in German). Schriften zur quantitatieven Betriebswirtschaftslehre, Physica-Verlag, Heidelberg.

Hansmann, K-W; (1993); **Capacity oriented production control for a job shop production**. Preprints "Management and new prod. systems", EIASM, Londen 5/6-4-93.

Hendry, LC; Kingsman, BG; (1988); **A decision support system to manage shop floor throughput times in make-to-order companies**. Paper presented at the 5th Int. working seminar on production economics, Igls, Austria.

Hendry, LC; Kingsman, BG; (1989); **Production planning systems and their applicability to make-to-order companies**. European journal of operational research, Vol 40, No 1, 1-15.

Hendry, LC; Kingsman, BG; (1991); **A Decision Support System for Job Release in Make-to-order Companies**. Int. J. of operations & production management, Vol 11, No 6, 6-16.

Hendry, LC; Kingsman, BG; (1993); **Customer enquiry management: part of a hierarchical system to control lead times in make-to-order Companies**. Journal of the operational research society, Vol 44, No 1, 61-70.

Irastorza, JC; Deane, RH; (1974); **A loading and balancing methodology for job shop control**. AIIE Transactions, Vol 6, No 4, 302-307.

Kanet, JJ; (1988); **Load-limited order release in job shop scheduling systems**. Journal of operations management, Vol 7, No 3, 44-58.

Kingsman, BG; Tatsiopoulos, IP; Hendry, LC; (1989); **A structural methodology for managing manufacturing lead times in make-to-order companies**. European journal of operational research, Vol 40, No 2, 196-209.

Knolmayer, G; (1991); **A widely acclaimed method of load-oriented job release and its conceptual deficiencies**. In: Fandel, G and Zäphel, G (eds.): Modern production concepts, Springer Verlag, Berlin.

Land, MJ; Gaalman GJC; (1994); **The choice of performance indicators for job shops;** in: K.W. Platts et al. (eds.), Operations Strategy and Performance, 101-106, Cambridge (UK).

Melnyk, SA; Ragatz, GL; Fredendall, L; (1991); **Load Smoothing by the Planning and Order Review/Release Systems: A Simulation Experiment**. Journal of Operations Management Vol 10, No. 4, Oct 1991, 512-523.

Nyhuis, P; (1992); **Logistische rationalisierungspotentiale erkennen und bewerten** (in German). Beitrag zum 9. Östereichischen Logistik-Dialog, Wien.

Panwalkar, SS; Iskander, W; (1977); **A survey of scheduling rules**. Operations research, Vol 25, No 1, 45-61.

Park, PS; Bobrowski, PM; (1989); **Job release and labor flexibility in a dual resource constrained job shop**. Journal of operations management, Vol 8, No 3, 230-249.

Ramasesh, R; (1990); **Dynamic job shop scheduling: A survey of simulation research**. Omega, Vol 18, No 1, 43-57.

Shimoyashiro, S; Isoda, K; Awane, H; (1984); **Input scheduling and load balance control for a job shop**. Int. journal of production research, Vol 22, No 4, 597-605.

Tatsiopoulos, IP; (1983); **A microcomputer-based interactive system for managing production and marketing in small component-manufacturing firms using a hierarchical backlog control and lead time management methodology**. Unpublished PhD dissertation, University of Lancaster.

Tatsiopoulos, IP; (1993); **Simplified production management software for the small manufacturing firm**. Production planning & control, Vol 4, No 1, 17-26.

Wein, LM; (1990); **Scheduling networks of queues: heavy traffic analysis of a two-station network with controllable inputs**. Operations Research, Vol 38, No 6, 1065-1078.

Wein, LM; (1992); **Scheduling networks of queues: heavy traffic analysis of a multistation network with controllable inputs**. Operations Research, Vol 40, No 2, S312-S334.

Wein, LM; Chevalier, PB; (1992); **A broader view of the job-shop scheduling problem**. Management science, Vol 38, No 7, 1018-1033.

Wiendahl, H-P; (1987); **Belastungsorientierte Fertigungssteuerung, Grundlagen, Verfahrensaufbau, Realisierung** (in German). Hanser Verlag, Munich.

Wiendahl, H-P; (1992); **Anwendung der belastungsorientierten Fertigungssteuerung** (in German). Hanser Verlag, Munich.

Wiendahl, H-P; (1995); **Load-oriented manufacturing control**. Springer Verlag, Berlin Heidelberg.

Wight, O; (1970); **Input/output control: a real handle on lead time**. Production and Inventory Management, Vol 11, No 3, 9-31.

Zäpfel, G; Missbauer, H; (1993); **Production planning and control (ppc) systems including load-oriented order release - problems and research perspectives**. Int. journal of production economics, Vol 30/31, 107-122.