Multiple batch structures in throughput scheduling

P.J. Weeda

Department of Mechanical Engineering, University of Twente, P.O. Box 217, 7531 JE Enschede, The Netherlands

Abstract

Throughput scheduling deals with maximizing time-constrained throughput in batch production. Previous investigations have shown that time-constrained throughput increases by processing successive batch cycles in simple process configurations consisting of three processes and two facilities, even if set-up times are positive. This paper explores the extension of these results to more complex process configurations.

1. Introduction

This study deals with a theoretical approach to throughput scheduling in batch production. It focuses on the complex relationship existing between batch structure, utilization of facilities and maximum throughput. It pinpoints to process structures, in which minimization of the number of set-ups on a heavily utilized facility does not maximize throughput, thus contradicting the validity of a commonly used approach: On a bottleneck facility, one should minimize the number of set-ups in order to maximize time-constrained throughput.

The number of batches is determined by the batch structure. If a single batch structure is applied, one batch of each process is executed. If a multiple batch structure is applied then for at least one process at least two batches are executed. For process configurations involving facilities to which at least two successive processes with positive set-up times are assigned, a multiple batch structure implies additional set-ups. A remarkable consequence of applying a multiple rather than a single batch structure in these cases is an increase in time-constrained throughput in spite of the required additional set-up time.

A rationale for this phenomenon is the fact that the initial idle time of the succeeding facility depends on the size of the first batch on the foregoing multi-process facility. On the other hand,

by the reduction of the first batch size at least a second batch is needed in order to improve the throughput within the allowed time. An optimal balance is obtained by adopting a multiple batch structure.

A process configuration is defined by a number of processes, which are run through by products according to a specific routing. To each process a unique facility is assigned. Facilities may be assigned to more than one process. Each pair of process and facility assigned to it implies a set-up time and a unit processing time. Transfer between facilities is instantaneous and transfer batches have size one. A time interval is specified during which the throughput, i.e. the number of final products produced, should be maximized. An example process configuration, involving three raw materials, which run through eight processes on five facilities, resulting in three final products, is depicted in Fig. 1.

The topics treated here, have some history. In [1], one of the three possible serial process configurations has been discussed in relation to the nine so-called OPT-rules (cf. [2]). In [3], all three possible serial process configurations, involving two facilities and three processes have been analyzed. This analysis has been heuristic in the sense that the methods used have not been compared with truly optimal methods.

In [4], an informal treatment has been presented for all seven possible process configura-

Fig. 1

tions (including convergent and divergent configurations), consisting of two machines and three processes.

In [5], optimal methods have been developed. A set of integer programs has been formulated, which are based on the logical extension of the notion of bottleneck facility to critical utilization path. By means of these methods optimal batch sizes and maximum throughput can be obtained for a given batch structure.

This study exploits these integer programming formulations in order to explore the preference of a multiple batch structure to a single batch structure with respect to maximum throughput in more complex process configurations. After a section on definitions and notation, extensions of the serial process configurations studied in [3–5] are considered. These extensions involve the enlargement of respectively the number of processes on one facility, the number of single process facilities and the number of multi-process facilities.

2. Definitions and notation

Facilities are denoted by capital letters. If more than one process is executed on the same facility, its capital letter is followed by a natural number indicating the sequence of processes on that facility. For example, if on facility *A*, *n* processes have to be executed, the processes are represented by the sequence *A1*, *A2*, ..., *An*. Such a facility will be called a multi-process facility. In case of a single process facility, the number will be dropped. With each pair of facility and process a

set-up time $S(\cdot)$ and a unit processing time $P(\cdot)$ is associated.

If the processes of a configuration are executed sequentially, they are separated by commas. For example, configuration (*A1*, *A2*, *A3*, *B*, *C*) represents a serial process configuration with three successive processes to be executed on facility *A*, successively followed by one process on each of the facilities *B* and *C*.

Batch sizes are indicated by *Q*. The size of the *m*th batch of a process will be denoted by a subscript to *Q*: Q_m . For example, the notation $Q_3(A2)$ denotes the size of the third batch for process *A2*.

A batch cycle consists of at most one batch for each process of the configuration. A batch cycle is homogeneous if all batch sizes are equal. For example, if the *i*th batch cycle for process configuration (*A*, *B*, *C*, *D*) is homogeneous, then

$$Q_i(A) = Q_i(B) = Q_i(C) = Q_i(D)$$

A multiple batch structure consists of $m > 1$ batch cycles by definition. For example, the batch sizes of a multiple batch structure consisting of three cycles for process configuration (*A1*, *A2*, *A3*, *B*, *C*), may be given by

<i>A1</i>	<i>A2</i>	<i>A3</i>	<i>B</i>	<i>C</i>
13	8	8	8	8
	5	3	3	3
		2	2	2

Note that this multiple batch structure is not homogeneous. However, for any batch structure the total number of units produced should be equal for each process. For a single batch structure, the only batch cycle is always homogeneous. A multiple batch structure is said to be repetitive, if all batch cycles have one common batch size.

3. Extensions of process configuration (*A1*, *A2*, *B*)

The extensions of process configuration (*A1*, *A2*, *B*) involve successively

- enlargement of the number of processes on facility *A*, i.e. configuration (*A1*, *A2*, ..., *An*, *B*)
- enlargement of the number of one-process facilities, i.e. configuration (*A1*, *A2*, ..., *An*, *B*, *C*, ..., *Z*)

- enlargement of the number of two-process facilities, i.e. configuration $(A1, A2, B1, B2)$.

Primarily a proof is given on the optimality of a multiple batch structure for the basic process configuration $(A1, A2, B)$. This proof is only of interest in case $P(B) > P(A2)$, since for $P(B) \leq P(A2)$, the optimality of a single batch structure is easily shown (cf. [3]).

In order to prove the optimality of a multiple batch structure it suffices to prove that the maximum throughput obtained for a double batch structure with homogeneous batch cycles exceeds the maximum throughput for a single batch structure. To this end, integer programming formulations for single and double batch structures are used.

The integer program corresponding to a single batch structure for configuration $(A1, A2, B)$ (cf. [5]), is given by

$$\text{Max MT}(1) = Q$$

Subject to

$$S(A) + Q \cdot P(A1) + P(A2) + Q \cdot P(B) \leq T$$

$$Q \geq 0 \text{ and integer}$$

It is easily verified, that the maximum throughput for the LP-relaxation is given by

$$\text{MT}(1)_{\text{LP}} = \frac{T - S(A) - P(A2)}{P(A1) + P(B)}$$

while the maximum throughput for the integer program is given by the integral part of $\text{MT}(1)_{\text{LP}}$ and denoted by $\text{MT}(1)_{\text{IP}}$. The integer program of the above type for a single batch structure will be referred to as IP(1).

In the discussion set-up on machine B has been neglected. In the sequel, a set-up on machine B is assumed to happen only once, i.e. during the initial idle time on machine B . This implies the following restriction on $S(B)$:

$$S(B) \leq S(A) + Q \cdot P(A1) + P(A2)$$

where Q denotes the size of the first batch of process $A1$.

For a double batch structure, two integer programs become relevant. They are respectively given by (cf. [5])

$$\text{Max MT}(2a) = Q_1 + Q_2$$

Subject to

$$2S(A) + Q_1 \cdot P(A) + Q_2 \cdot P(A1)$$

$$+ P(A2) + Q_2 \cdot P(B) \leq T$$

$$S(A) + Q_1 \cdot P(A2) + Q_2 \cdot P(A1) \geq Q_1 \cdot P(B)$$

$$Q_1, Q_2 \geq 0 \text{ and integer}$$

and

$$\text{Max MT}(2b) = Q_1 + Q_2$$

Subject to

$$S(A) + Q_1 \cdot P(A1) + P(A2)$$

$$+ (Q_1 + Q_2) \cdot P(B) \leq T$$

$$S(A) + Q_1 \cdot P(A2) + Q_2 \cdot P(A1) \leq Q_1 \cdot P(B)$$

$$Q_1, Q_2 \geq 0 \text{ and integer}$$

where Q_1 and Q_2 denote the respective batch sizes of the two homogeneous batch cycles. The two integer programs for a double batch structure will be referred to as IP(2a) and IP(2b).

Both integer programs have two constraints, which can be derived by using the notion of a critical utilization path (cf. [5]). A critical utilization path is defined by a path of full utilization covering the full time allowance T . As the Gantt chart in Fig. 2 reveals, the associated critical utilization path for IP(2a) is largely on facility A . Only at the second batch it switches to facility B . For IP(2b) the associated critical utilization path is largely on facility B as Fig. 3 reveals. Only during the first batch it is on facility A . In both programs the first of the two constraints reflects the critical utilization path, while the second constraint represents the condition for the shape of the first constraint. Note in this respect the connection between the two batches on facility B .

The main tool to establish the optimality of a multiple batch structure for configuration $(A1, A2, B)$ and its extensions is the following theorem.

Theorem. Suppose integer program IP(2b) for configuration $(A1, A2, B)$ is feasible for $P(B) > P(A2)$ and possesses integer solutions with positive Q_2 , then a double batch structure with homogeneous batch cycles matches or ex-

Fig. 2.

Fig. 3.

ceeds a single batch structure in time-constrained throughput.

Without loss of generality, the two integer programs IP(2a) and IP(2b) can be joined to one integer program IP(2). The set of feasible solutions of integer program IP(2) is given by the union of the set of feasible solutions of IP(2a) and IP(2b). Integer program IP(2) is given by

$$\text{Max MT}(2) = Q_1 + Q_2$$

Subject to

$$2S(A) + Q_1 \cdot P(A) + Q_2 \cdot P(A1) + P(A2) + Q_2 \cdot P(B) \leq T$$

$$S(A) + Q_1 \cdot P(A1) + P(A2) + (Q_1 + Q_2) \cdot P(B) \leq T$$

$$Q_1, Q_2 \geq 0 \text{ and integer}$$

The maximum throughput of IP(2) satisfies

$$\text{MT}(2) = \text{Max}[\text{MT}(2a), \text{MT}(2b)]$$

Enlargement of the number of processes on facility A

This extension of configuration (A1, A2, B) is denoted by (A1, A2, ..., An, B). Integer program IP(1) for this configuration is given by

$$\text{Max MT}(1) = Q$$

Subject to

$$S(A) + Q \cdot \sum_{j=1}^{n-1} P(Aj) + P(An) + Q \cdot P(B) \leq T$$

$$Q \geq 0 \text{ and integer}$$

where

$$S(A) = \sum_{j=1}^n S(Aj)$$

Obviously this integer program is identical to IP(1) for configuration (A1, A2, B) if $\sum_{j=1}^{n-1} P(Aj)$ and $P(An)$ are replaced by $P(A1)$ and $P(A2)$ respectively. It is easily verified that this replacement also holds for IP(2). Hence the same proof remains valid, implying the optimality of a multiple batch structure compared with a single batch structure for process configuration (A1, A2, ..., An, B).

It should be noted, that nothing is said about the optimality of a multiple batch structure with homogeneous batch cycles. A better solution may be obtained for a batch structure with two or more batch cycles, which are not homogeneous. The following example demonstrates this.

$n = 3, T = 1000$

$S(A1) = 50$	$P(A1) = 10$	$P(B) = 50$
$S(A2) = 20$	$P(A2) = 30$	
$S(A3) = 20$	$P(A3) = 30$	

This problem is equivalent for homogeneous batch cycles to the numerical problem

$n = 2, T = 1000$

$S(A) = 90$	$P(A1) = 40$	$P(B) = 50$
	$P(A2) = 30$	

The latter problem yields a throughput of 10 for the following two homogeneous batch cycles:

$A1$	$A2$	$A3$	B
8	8	8	8
2	2	2	2

while a throughput of 11 is obtained by the following batch cycles:

$A1$	$A2$	$A3$	B
11	7	7	7
	3	3	3
	1	1	1

which are obviously not homogeneous.

Enlargement of the number of one-process facilities

This extension of configuration $(A1, A2, \dots, An, B)$ is denoted by $(A1, A2, \dots, An, B, C, \dots, Z)$. Let $P(X)$ denote the process with the largest unit processing time with respect to the processes $B, C, \dots, Z$. The integer program IP(1) for this configuration is given by

Max $MT(1) = Q$

Subject to

$$S(A) + Q \cdot \sum_{j=1}^{n-1} P(Aj) + P(A2) + Q \cdot P(X) \\ + (P(B) + \dots + P(Z)) - P(X) \leq T$$

$Q \geq 0$ and integer

where $\{P(B) + \dots + P(Z)\}$ denotes the sum of all unit processing times of the one-process facilities. Note that $P(X) > P(An)$ is assumed, similar to the condition $P(B) > P(A2)$ for configuration $(A1, A2, B)$. The critical utilization path establishing the time constraint in IP(1) is depicted in Fig. 4. It starts and remains on machine A until the first unit of process $A2$ is produced. Next it runs through exactly one unit of each process of the sequence before process X , followed by the batch of process X with size Q and exactly one unit of each process of the sequence of processes after process X . Also for this configuration, it is assumed that set-up times of the one-process facilities can be neglected. For each of the one-process facilities $B, C, \dots, Z$ this assumption means that set-up can be done during the initial idle time.

Integer program IP(2) for two successive homogeneous batch cycles reflects the two possible utilization paths and is given by

Max $MT(2) = Q_1 + Q_2$

Subject to

$$2S(A) + Q_1 \cdot P(A) + Q_2 \cdot \sum_{j=1}^{n-1} P(Aj) + P(An) \\ + Q_2 \cdot P(X) + \{P(B) + \dots + P(Z)\} - P(X) \leq T \\ S(A) + Q_1 \cdot \sum_{j=1}^{n-1} P(Aj) + P(An) + (Q_1 + Q_2) \cdot P(X) \\ + 2\{P(B) + \dots + P(Z)\} - 2P(X) \leq T$$

$Q_1, Q_2 \geq 0$ and integer

Replacing respectively $T - \{P(B) + \dots + P(Z)\} + P(X)$ by T , $P(X)$ by $P(B)$, $\sum_{j=1}^{n-1} P(Aj)$ by $P(A1)$, and $P(An)$ by $P(A2)$ integer program IP(2) for configuration $(A1, A2, B)$ are reobtained with exception of the second constraint, which is given by

Fig. 4.

Fig. 5.

$$S(A) + Q_1 \cdot P(A1) + P(A2)$$

$$+ (Q_1 + Q_2) \cdot P(B) + \Theta \leq T$$

where $\Theta = \{P(B) + \dots + P(Z)\} - P(X)$. Note that Θ only depends on the data of the problem and not on a batch size. Because of this term $\Theta > 0$, the set of feasible solutions of integer program IP(2) is clearly reduced. Consequently the advantage of a multiple batch structure is also reduced.

Enlargement of the number of two-process facilities

Primarily process configuration (A1, A2, B1, B2) is analyzed. As the Gantt chart of Fig. 5 reveals, the integer program IP(1) for this process configuration is given by

$$\text{Max MT}(1) = Q$$

Subject to

$$S(A) + Q \cdot P(A1) + P(A2)$$

$$+ Q \cdot P(B) + S(B2) \leq T$$

$$Q \geq 0 \text{ and integer}$$

Integer program IP(2) is given by

$$\text{Max MT}(2) = Q_1 + Q_2$$

Subject to

$$2S(A) + Q_1 \cdot P(A) + Q_2 \cdot P(A1) + P(A2)$$

$$+ Q_2 \cdot P(B) + S(B2) \leq T$$

$$S(A) + Q_1 \cdot P(A1) + P(A2)$$

$$+ (Q_1 + Q_2) \cdot P(B) + S(B) + S(B2) \leq T$$

$$Q_1, Q_2 \geq 0 \text{ and integer}$$

Replacing $T - S(B2)$ by T , the integer programs for configuration (A1, A2, B) are reobtained with the exception of a quantity $\Theta > 0$ depending only on the data. For this configuration $\Theta = S(B)$. The remarks made for the preceding case also apply here.

The analysis of process configuration (B, A1, A2) is quite similar to the analysis of process configuration (A1, A2, B) because of the inherent symmetry. Also its extensions can be treated along the lines described above. They are respectively denoted by (B, A1, A2, ..., An) and (B, C, ..., Z, A1, A2, ..., An).

4. Convergent process configurations

In this section the convergent process configuration (A1/A2, B) and its extensions are considered. This configuration represents for example the situation in which two types of parts are processed on one facility A. The two types of parts are subsequently assembled to final products on facility B. The main difference with serial configuration (A1, A2, B) is the fact that the two processes on facility A can be executed in arbitrary order. Without loss of generality, one may fix the

Fig. 6.

order of the processes $A1$ and $A2$. The integer programs $IP(1)$ and $IP(2)$ are identical to those for the corresponding serial structure $(A1, A2, B)$. The following extensions require an identical analysis:

- enlargement of the number of processes on facility A , i.e. configuration $(A1/A2/\dots/An, B)$,
- enlargement of the number of one-process facilities, i.e. the configuration $(A1/A2/\dots/An, B, C, \dots, Z)$.

The structure corresponding to the latter extension is depicted in Fig. 6. The results are similar to those in the preceding section. For the first extension $\Theta=0$ while for the second $\Theta>0$.

Conclusions

In this paper immediate extensions of the serial process configurations $(A1, A2, B)$ and $(B, A1, A2)$ and convergent process configuration $(A1/A2, B)$ have been discussed. It has been shown that the same proof could be used in order to show that time-constrained throughput is maximized by a multiple rather than a single batch structure. However the advantage of a multiple batch structure is reduced if a quantity Θ depending on the process configuration is positive.

In spite of this reduction for the latter two extensions of process configuration $(A1, A2, B)$ it has been shown in this paper that minimizing the number of set-ups on heavily utilized machines does not necessarily lead to a maximum time-constrained throughput for all process configurations investigated. Other process configurations can be explored by the same line of thought.

Appendix

Proof of Theorem. Consider the LP-relaxation of $IP(2b)$. The intersection of the straight lines in two-dimensional space corresponding to the constraints of the LP-relaxation with the axes are respectively given by

$$(Q'_1, Q'_2) = \left(\frac{T - S(A) - P(A2)}{P(A1) + P(B)}, \frac{T - S(A) - P(A2)}{P(B)} \right)$$

and

$$(Q''_1, Q''_2) = \left(\frac{S(A)}{P(B) - P(A2)}, -\frac{S(A)}{P(A1)} \right)$$

The first relation implies $Q'_2 > Q'_1 = MT(1)_{LP}$. A feasible solution $(Q^*_1, Q^*_2) > 0$, satisfying the first constraint of the LP-relaxation with equality sign, implies $Q^*_2 > Q'_1 - Q^*_1$ or equivalently $Q^*_1 + Q^*_2 > Q'_1 = MT(1)_{LP}$. Hence there exists a feasible solution to the LP-relaxation of $IP(2b)$ with $MT(2b)_{LP} > MT(1)_{LP}$. Let Y_1 denote the integral part of $MT(1)_{LP} = Q'_1$ and let Y_2 be the point on the first constraint of the LP-relaxation with abscissa Y_1 . If $Y_2 \geq 1$ then $(Y_1, 1)$ is a feasible integer solution to $IP(2b)$ with throughput equal to $Y_1 + 1 > Y_1 = MT(1)_{IP}$. If $Y_2 < 1$ then the throughput equals at least $MT(1)_{IP}$. $\square$

References

- 1 Huisman, H.H., Polderman, G. and Weeda, P.J., 1990. Maximizing throughput in some simple time-constrained scheduling situations. *Eng. Costs Prod. Econ.*, 18: 293-299.
- 2 Goldratt, E.M. and Fox, R.E., 1986. *The Race*. North River Press, New York.

- 3 Weeda, P.J., 1990. On the choice of batch mode in order to maximize throughput. In: Proc. 5th Internat. Symp. on Inventories. Elsevier, Amsterdam.
- 4 Weeda, P.J., 1990. On the relation between batch mode, bottleneck machines and maximum throughput. In: Proc. 10th Internat. Conf. on Production Research, Nottingham.
- 5 Weeda, P.J., 1990. Maximizing throughput by evaluating critical utilization paths. Preprints 6th Seminar on Production Economics, Igls.