BUFFER MANAGEMENT: A DIAGNOSTIC TOOL FOR PRODUCTION CONTROL

ELI SCHRAGENHEIM
BOAZ RONEN
The A. Y. Goldratt Institute, 442 Orange St., New Haven, CT 06511, Tel Aviv University, Faculty of Management, Tel Aviv 69978, Israel.

The shop-floor production control method described here assesses the actual state on the floor compared to the planned state and points out those areas that urgently need corrective action. Simple comparison between planned and actual state is not practical; in most cases, the data collection itself is so enormous that it cannot be done frequently. In other cases the task of the user, who has to draw conclusions out of a huge list of deviations from the plan, becomes very difficult.

Buffer management:

- Serves as an alarm system that spots serious and urgent problems which threaten to disrupt the plan and cause real damage.
- Provides control on lead time.
- Indicates the weak areas, thus prioritizing the necessary improvements in the shop floor.

Buffer management (BM) is a diagnostic tool which is a shop-floor control methodology. It can be used in an environment where the drum-buffer-rope (DBR) [2] method has been implemented. The DBR scheduling technique is described in [4] and a simulation of buffer management has been carried out on the example used in that article. BM has broad ramifications for manufacturing management by providing support for short-range and mid-range decisions.

CONCEPT OF BUFFER DURATION

The purpose of a buffer, according to DBR methodology, is to protect a schedule—meaning, to ensure that the scheduled parts will be where they are needed at the time they are needed.

The way to provide such protection to the schedule is to release the raw material some time (which constitutes the buffer) before its schedule. According to DBR methodology, only the capacity constraint resources and the shipping due dates are being properly scheduled. These schedules (the drum) have to be protected. Some assembly operations which assemble constraint and nonconstraint parts are being protected as well. The buffer duration is a fairly large, yet possible, lead time to the protected area. In the vast majority of the cases the processing times are negligible in the determination of the buffer durations. What counts the most are the statistical fluctuations which may take place and the level of excess capacity of the nonconstraint resources involved. The meaning of implementing a buffer of, say, 10 hours is that these parts are most certainly expected within the next 10 hours (provided the rope, the release schedule, has been correctly implemented.)

Every part, on its way from raw material release to the shipping area, passes through at least one scheduled area which is protected by a buffer. The total lead time of that part is determined by the sum of the buffer durations.

THE CONCEPT OF BUFFER MANAGEMENT

Basically, the idea behind buffer management is to monitor the inventory in front of the protected resources and compare the actual versus the planned performance.

Capacity Constraint Buffer

On one hand there is a detailed schedule for the constraint's operations (the drum), as opposed to the actual arrival time of the parts needed for the constraint's operations. Suppose the constraint is scheduled to start processing a certain batch at time T—that is, parts are expected to arrive at the constraint's area between T minus the buffer duration (on which the release of material is scheduled) and T.

Instead of tracking down the exact arrival times, it is simpler to check, at a certain frequency, the content of the inventory at the constraint's area, which is the buffer checkpoint. For the next buffer duration time, the schedule of the constraint dictates that it will need a certain list of parts. Marking on that list the parts which are still missing at the buffer checkpoint gives a picture of the system's actual deviations from the schedule. Table 1 displays the results of the inventory check.

C1 and C2 are names for the capacity constraint's jobs—part-number / operation in MRP terminology. The minus sign denotes that the parts are missing. Each column represents 1 hour. The duration of the buffer is 10 hours. Each column is divided into eight boxes of 7.5 minutes each. The drum (the constraint's schedule) is displayed for the next 10 hrs by the name of the operation in each 7.5-minute box.

According to the display the constraint has to work for the next 4.5 hrs on part-number / operation C1, but for about 1 hr of work the parts are still missing. The displayed area of missing parts is referred to as a "hole"; the constraint is actually protected for only about 3.5 hrs.

Notice that only the inventory at the buffer checkpoint is considered. The missing pieces are somewhere on their way, maybe just one or two operations behind.

Shipping Buffer

The BM layer for the shipping and assembly buffer have to differ in certain details from that of the constraint buffer. While the constraint buffer is easily expressed in time units the shipping buffer does not have any processing time of its own. Hence within a column, the shipping buffer layout may represent parts and not time. The columns still represent the bigger time unit, which is the significant unit for shipping due dates. Table 2 shows that 3 hrs from now, a shipment of 8 units of product A is due. Of these, about 4 are still missing. Shipment of product B is promised 6 hrs from now, and is already available. In 9 hrs a shipment of product A is due, but it is still not ready. An empty column means that no shipments are due at that time.

Assembly Buffer

The assembly buffer presents the same problem. Its own processing time is not the important one—it is the capacity constraint which feeds it that counts. (A layout like that of table 2 will make the assembly buffer more understandable.)

GENERATING BM INFORMATION

There are three types of results we expect from managing the buffers:

1. The alarm is sounded to handle situations where expediting is necessary to meet the schedule. Improper handling of the rope (like somebody releasing unnecessary material onto the floor) is also signaled.
2. The duration of the buffers is monitored (note that the buffer duration determines the lead time).
3. Decision support is provided for mid-range decisions, which indicates where to focus efforts to improve the system.

Whenever a hole in a buffer is spotted, three lines of action might be adopted: disregard, monitor, or expedite. The three lines of action define the regions in the BM layout:

Region 3: Any hole in this region which is in the buffer horizon but still far away from consumption will be disregarded. There is still enough time until these parts are needed by the drum.

Region 2: Any hole in this middle region calls for locating the missing parts. They do not yet justify expediting but they may cross the time limit for action. In most cases the manager should refrain from direct interference.

Region 1: Those are the holes, very near the scheduled consumption time, which threaten the output of the

TABLE 1: Typical Constraint's Buffer Checkpoint Inventory Layout

	Time (hr)									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Time (min) 60	C1	C1	C1	−C1	C2	C2	C2	C2	−C1	−C1
	C1	C1	C1	−C1	C2	C2	C2	C2	−C1	−C1
45	C1	C1	C1	−C1	C2	C2	C2	C2	−C1	−C1
	C1	C1	C1	−C1	C2	C2	C2	C2	−C1	−C1
30	C1	C1	C1	C1	−C1	C2	C2	C2	−C1	−C1
	C1	C1	C1	C1	−C1	C2	C2	C2	−C1	−C1
15	C1	C1	C1	C1	−C1	C2	C2	C2	C2	−C1
	C1	C1	C1	C1	−C1	C2	C2	C2	C2	−C1
	Region 1			Region 2			Region 3			

TABLE 2: Typical Layout of the Shipping Buffer Checkpoint

	Time (hr)									
	1	2	3	4	5	6	7	8	9	10
Units 16										
14										
12						B				
10						B				
8			−A			B			−A	
6			−A			B			−A	
4			A			B			−A	
2			A			B			−A	
	Region 1			Region 2			Region 3			

plant. The first thing to do is to expedite. The warning is sound and clear. The hole should have been noticed when it was still in Region 2; when it moves into Region 1, it becomes a first priority to handle. The optimal division of the buffer into the three regions depends on the characteristics of the actual lead-time distribution function. Without losing generality we've assumed that they are of equal size. That assumption is convenient, but doesn't have an effect on the analysis.

In Table 1 there is a Region 2 hole between 3.5 and 4.5. The line of action is to locate the missing parts and monitor the location in case it penetrates into Region 1. There is a hole in Region 3 as well, but no action is needed.

The BM layout accumulates all the disruptions, delays, and faster-than-usual operations into one map, with the correct distance from the critical areas scheduled. Region 1 has to be large enough so that in most cases Region-1 hole parts can be rushed in time. The buffer duration should be large enough so that the probability of several Region-1 holes competing for the same resources at the same time will be very small.

Monitoring the number of Region-2 and Region-1 holes provides a very good evaluation of the effect of the buffers as a protective mechanism. Should the BM layout show very few Region-2 holes, it means that the buffer duration can be reduced. Should the layout show many Region-1 holes, the buffer duration should be increased immediately. Determining the "right" buffer durations reflects the balance between lead time and protection of the key areas of the whole organization.

The areas in which improvement will affect the whole system are of two kinds:

The constraint itself. Notice that improvement on the constraint will cause one of the following results: (1) The old constraint will still remain a system constraint, but the improved throughput of the system will result in increased load on the nonconstraints. This may require increasing the buffer durations. (2) The old constraint will cease to be the constraint and a new constraint will emerge. The whole scheme of the buffers will then have to change.

A nonconstraint resource that regularly causes Region-2 holes. Improving the performance of such a resource will not increase the throughput of the system, assuming management is satisfied with the current balance between protection and lead time, but the improvement of a troublemaker (by such means as setup reduction, etc.) may lead to reduced buffer durations and thereby to reduced lead time. (A troublemaker is any cause of frequent holes.)

From analyzing the characteristics of both kinds of improvements, it clearly becomes imperative to identify the troublemakers in the system. If there are causes of frequent holes in Region 1, the buffer duration should be increased, and, if this does not help, it means that the troublemaker is actually a constraint in itself. (Notice that a troublemaker is a potential future constraint.)

Accumulated BM information about causes of holes provides a list of the few resources that affect system performance. To get that information, the buffer checkpoints should be checked at a fixed frequency and the following data concerning holes in Regions 2 and 1 should be collected:

- Buffer type: constraint, shipping, or assembly
- The cause of the hole (this might be a resource, scarce raw material, setup crew, etc.). If the cause is not clear, it is suggested that the name of the resource which holds the missing part be entered.

- The size of the hole in time units
- The Region (2 or 1).

The outcome of the accumulated information is the base for mid-range improvement-plan decisions.

THE SIMULATED ENVIRONMENT

The example used here is the same as in [4] in resource layout, routing layout for two products and drum (Tables 1, 2, and 4). There are different types of machines: V, N, Y, R, M, and W. Machine R has two identical units, all the rest have one. M is a capacity constraint resource. The two products are I and J. The routing operations are denoted by A1 to G3 (the grid is defined between A–G and 1–5).

The constraint buffer protecting M's operations is 9 hrs long, the shipping buffer protecting the due dates is 6 hrs long, and the assembly buffer protecting the F4 assembly is 9 hrs long. The rope schedule (material release) is derived by the drum schedule and the F4 assembly schedule (also derived by the drum) minus the appropriate buffer duration. For details concerning the buffers, the derivation of the schedule, the rope schedule, the buffer durations and the nonconstraint resource activation, see [4].

The simulator has full interactive capabilities, so that even when the above policy doesn't yield the best solution, the user is able to override a decision. Such intervention occurred only according to the BM line of action—either a hole in Region 1 or a Region 2 hole which would certainly penetrate to Region 1. Any intervention was collected and is a part of the results.

According to transfer batch policy, parts are moved in quantities of 1. (See [2] for transfer/process batch discussion.) The stochastic environment is provided by the following:

- Downtime on machines. The occurrence of a break is linear at times of work, while the length of a break is exponential.
- Setup times are exponential.
- Time-per-part is normally distributed.

No scrap is used in these runs.

THE SIMULATION RESULTS

Two independent runs of two weeks each were conducted. Data entry for the holes was done every

TABLE 3: The Buffer Management Layout for the Initial State

Week 1, day 1, time: 0:0

Capacity Constraint's Buffer

	1	2	3	4	5	6	7	8	9
60	E3	E3	E3	E3	E3	E3	E3	E3	−c3
	E3	E3	E3	E3	E3	E3	E3	E3	−c3
45	E3	E3	E3	E3	E3	E3	E3	E3	−c3
	E3	E3	E3	E3	E3	E3	E3	E3	SU
30	E3	E3	E3	E3	E3	E3	E3	E3	E3
	E3	E3	E3	E3	E3	E3	E3	E3	E3
15	E3	E3	E3	E3	E3	E3	E3	E3	E3
	SU	E3	E3	E3	E3	E3	E3	E3	E3
	Region 1			Region 2			Region 3		

Shipping Buffer

	1	2	3	4	5	6
8						
7						
6						
5						
4						
3						
2						−J
1						−J

Assembly Buffer

	1	2	3	4	5	6	7	8	9
8									
7									
6									
5									
4									
3				F4		−f4			
2	F4	F4	F4	F4	−f4	−f4	−f4	−f4	−f4
1	F4	F4	F4	F4	F4	−f4	−f4	−f4	−f4

four hours. The computerized simulator draws the BM layout every 15 minutes. Table 3 shows the initial state of the BM layout. This corresponds to the plant's initial state and the drum. The material in the constraint buffer checkpoint (for operation E3) provides an actual protection of 500 minutes.

The buffer layout was drawn in accordance with the schedule of resource M, noting whether the parts have already reached the buffer checkpoint.

The initial state of the assembly buffer checkpoint is 10 pieces, and it will need 20 parts in the next 9 hrs for the F4 assembly, when 20 pieces from E3 are scheduled to arrive. It appears that within 5 hrs, F4 will need the first part, which is now missing. This means that there is a hole in Region 2 (Region 2 is 4–6 hrs from now). The parts missing are stuck behind the G3 operation, which is a W operation.

The shipping buffer layout shows that at 6 hrs from now two product units are expected. This was derived by the drum which starts with E3 in accordance with the setup and processing times. The 6-hr shipping buffer added to that schedule. Data entry should record that W is responsible for a Region-2 hole in the assembly buffer checkpoint.

RESULTS ANALYSIS

A compilation of results obtained from two simulation runs is shown in Table 4. Run 2 results show more Region-1 holes caused by W than run 1. Run 2 suggests that the 9-hr buffer duration for the constraint may be too short—the trade-off is between protection and lead time.

The assembly buffer is clearly the less critical from the protection point of view: as long as the shipping buffer does not have too many Region-1 holes, there is not much point in increasing it.

The list of troublemakers which emerges from these results is very clear: W is certainly a troublemaker far more than any other nonconstraint resource. This can be explained by the total load on W (excluding the current WIP): 1420 min processing time + 720 min setup time (8 setups per week) = 2140 min. On top of this there is an average 10% downtime which brings the total load to 2354 min out of 2400 min of available time. In a real-life situation the data, as stored in the data base, may be wrong, but BM points to the problematic area based on actual performance.

Y may be the next troublemaker, but the holes it caused in the shipping buffer were partially due to downtime (and the other fluctuations) on the capacity constraint itself. The shipping buffer should take care of the capacity constraint's fluctuations as well. This finding and the fact that R has provided several of Region 2's holes lead to the conclusion that the shipping buffer is too short.

The message for mid-range improvements should be to concentrate on W. It looks as if M's performance is improved, and W will soon become the next constraint. This is of utmost importance when calculating how much more throughput the system will have after improving M's performance. Any improvement in W will have a positive effect on the protection of M (the system constraint). Reduction of either setup time or time per part are just two possibilities.

TABLE 4: Location and Causes of Holes in the Two Simulation Runs

	RUN #1			RUN #2		
Machine	*Region*	*#Holes*	*Total Size (hr)*	*Region*	*#Holes*	*Total Size (hr)*
			Constraint Buffer			
W	1	1	.25	1	3	4.75
W	2	13	23.5	2	15	23.7
			Shipping Buffer			
Y	1	4	3.75	1	1	1
Y	2	4	6.5	2	6	6.5
R	1	0	0	1	1	1
R	2	3	4.5	2	4	4
			Assembly Buffer			
W	1	4	4.3	1	6	9.6
W	2	8	13.5	2	9	16.
N	2	3	4.	2	2	3.
V	1	0	0	1	1	1
V	2	4	5	2	5	4.5

Very small holes were ignored. Each Region 1 hole would trigger expediting in a real-life situation. Twice in the first run there was a major mistake in assigning the W resource to a job. There was one manual intervention in the second run.

CONCLUSIONS

Buffer management enables management to focus on the right corrective actions to keep the system performance intact, monitor the trade-off between protection and lead time, and assess the impact of major changes and/or improvements which are to be implemented.

Since it focuses only on what is crucial [1], BM uses only a very limited number of data items which greatly simplifies the procedure while maintaining its full effectiveness. There are, nevertheless, some information-system and organizational ramifications. The information system, computerized or not, should be able to plot the drum so that adding the layout of the buffer checkpoint can be done manually (or through an on-line system which records the arrival of materials to the buffer checkpoint). A buffer manager should be appointed to check the buffer checkpoints and spread the information (Region-1 and -2 holes) to the relevant work centers. In case of an expediting mode (Region-1 hole), his task will be similar to that of an expeditor.

REFERENCES

1. Goldratt, E. M., *Theory of Constraints,* North River Press, Croton-on-Hudson, NY (1990).
2. ———, and Fox, R. E., *The Race,* North River Press, Croton-on-Hudson, NY (1986).
3. Ronen, B. and Palley, M., "A Topology of Financial versus Manufacturing Management Information Systems," *Human Systems Management* (April 1988).
4. Schragenheim, E. M. and Ronen, B., "The Drum-buffer-rope Shop Floor Control," *Production and Inventory Management Journal,* Vol. 31, No. 3 (1990), pp. 18–23.

About the Authors—

ELI SCHRAGENHEIM is a partner with A. Y. Goldratt Institute, New Haven, Connecticut. He is in charge of developing computerized educational tools and teaching workshops for management in manufacturing. He holds a BS in mathematics and a MBA from Tel-Aviv University, where he is now a PhD student in the Faculty of Management.

BOAZ RONEN is a senior lecturer at Tel-Aviv University in the Faculty of Management. He holds a BS in electronics engineering from Technion, Haifa, and MS and PhD degrees from Tel-Aviv University. Prior to his academic career he worked for 10 years in the high-tech electronics industry. He has also been a visiting professor in the business schools of New York University and Columbia University. The main areas of his research, teaching and consulting are MIS in operations management, JIT, TQM and TOC.