



22S Theory of Constraints

Learning Objectives

- LO22S-1** Explain the Theory of Constraints (TOC).
- LO22S-2** Analyze bottleneck resources and apply TOC principles to controlling a process.
- LO22S-3** Compare TOC to conventional approaches.
- LO22S-4** Evaluate bottleneck scheduling problems by applying TOC principles.

Eli Goldratt's Theory of Constraints

LO22S-1

Explain the Theory of Constraints (TOC).

Around 1980, Dr. Eli Goldratt, an Israeli business management consultant, contended that manufacturers were not doing a good job in scheduling and controlling their resources and inventories. To solve this problem, Goldratt and his associates at a company named Creative Output developed software that scheduled jobs through manufacturing processes, taking into account limited facilities, machines, personnel, tools, materials, and any other constraints that would affect a firm's ability to adhere to a schedule.

This was called *optimized production technology (OPT)*. The schedules were feasible and accurate and could be run on a computer in a fraction of the time needed by an MRP system. This was because the scheduling logic was based on the separation of bottleneck and nonbottleneck operations. To explain the principles behind the OPT scheduling logic, Goldratt described nine production scheduling rules (see Exhibit 22S.1). After approximately 100 large firms had installed this software, Goldratt went on to promote the logic of the approach rather than the software.

In broadening his scope, Goldratt developed his "Theory of Constraints" (TOC), which has become popular as a problem-solving approach that can be applied to many business areas. Exhibit 22S.2 lists the "Five Focusing Steps of TOC." His Goldratt Institute (www.goldratt.com) teaches courses in improving production, distribution, and project management. The common thread through all of these courses is Goldratt's TOC concepts.

Before we go into detail on the theory of constraints, it is useful to compare it with two other popular approaches to continuous improvement: Six Sigma and lean manufacturing. Both Six Sigma and lean approaches focus on cost reduction through the elimination of waste and the reduction of variability at every step in a process or in view of every component of a system. In contrast, the TOC five-step approach is more focused in its application. It concentrates its improvement efforts only on the operation that is constraining a critical process or on the weakest component that is limiting the performance of the system as a whole. If these elements are effectively managed, then it follows that better overall performance of the system relative to its goal is more likely to be achieved.

In this supplement, we focus on Goldratt's approach to manufacturing. To correctly treat the topic, we decided to approach it in the same way that Goldratt did: that is, by first defining some basic issues about firms—purposes, goals, and performance measures—and then

Exhibit 22S.1 Goldratt's Rules of Production Scheduling

1. Do not balance capacity—balance the flow.
2. The level of utilization of a nonbottleneck resource is determined not by its own potential but by some other constraint in the system.
3. Utilization and activation of a resource are not the same.
4. An hour lost at a bottleneck is an hour lost for the entire system.
5. An hour saved at a nonbottleneck is a mirage.
6. Bottlenecks govern both throughput and inventory in the system.
7. The transfer batch may not, and many times should not, be equal to the process batch.
8. A process batch should be variable both along its route and in time.
9. Priorities can be set only by examining the system's constraints. Lead time is a derivative of the schedule.

Exhibit 22S.2 Goldratt's Theory of Constraints (TOC)

1. Identify the system constraint. (No improvement is possible unless the constraint or weakest link is found.)
2. Decide how to exploit the system constraint. (Make the constraint as effective as possible.)
3. Subordinate everything else to that decision. (Align every other part of the system to support the constraint even if this reduces the efficiency of nonconstraint resources.)
4. Elevate the system constraint. (If output is still inadequate, acquire more of this resource so it no longer is a constraint.)
5. If, in the previous steps, the constraint has been broken, go back to Step 1, but do not let inertia become the system constraint. (After this constraint problem is solved, go back to the beginning and start over. This is a continuous process of improvement: identifying constraints, breaking them, and then identifying the new ones that result.)

dealing with scheduling, providing buffer inventories, focusing on the influences of quality, and looking at interactions with marketing and accounting.

Underlying Goldratt's work is the notion of **synchronous manufacturing**, which refers to the entire production process working in harmony to achieve the profit goal of the firm. When manufacturing is truly synchronized, its emphasis is on total system performance, not on localized measures such as labor or machine utilization.

Synchronous manufacturing

A production process coordinated to work in harmony to achieve the goals of the firm.

The Goal of the Firm

Goldratt has a very straightforward idea about the goal of any firm:

THE GOAL OF A FIRM IS TO MAKE MONEY.

Goldratt argues that although an organization may have many purposes—providing jobs, consuming raw materials, increasing sales, increasing share of the market, developing technology, or producing high-quality products—these do not guarantee long-term survival of the firm. They are means to achieve the goal, not the goal itself. If the firm makes money—and only then—it will prosper. When a firm has money, it can place more emphasis on other objectives.

Performance Measurements

To adequately measure a firm's performance, two sets of measurements must be used: one from the financial point of view and the other from the operations point of view.

Financial Measurements We have three measures of the firm's ability to make money:

1. *Net profit*—an absolute measurement in dollars.
2. *Return on investment*—a relative measure based on investment.
3. *Cash flow*—a survival measurement.

All three measurements must be used together. For example, a *net profit* of \$10 million is important as one measurement, but it has no real meaning until we know how much investment was needed to generate that \$10 million. If the investment was \$100 million, this is a 10 percent *return on investment*. *Cash flow* is important because cash is necessary to pay bills for day-to-day operations; without cash, a firm can go bankrupt even though it is very sound in normal accounting terms. A firm can have a high profit and a high return on investment but still be short on cash if, for example, profit is invested in new equipment or tied up in inventory.

Operational Measurements Financial measurements work well at the higher level, but they cannot be used at the operational level. We need another set of measurements that will give us guidance:

Throughput

The rate at which money is generated by the system through sales (Goldratt's definition).

Inventory (Goldratt's definition)

All the money that the system has invested in purchasing things it intends to sell.

Operating expenses

All the money that the system spends to turn inventory into throughput (Goldratt's definition).

1. **Throughput**—the rate at which money is generated by the system through sales.
2. **Inventory**—all the money that the system has invested in purchasing things it intends to sell.
3. **Operating expenses**—all the money that the system spends to turn inventory into throughput.

Throughput is specifically defined as goods *sold*. An inventory of finished goods is not throughput, but inventory. Actual sales must occur. It is specifically defined this way to prevent the system from continuing to produce under the illusion that the goods *might* be sold. Such action simply increases costs, builds inventory, and consumes cash. Inventory that is carried (whether work-in-process or finished goods) is valued only at the cost of the materials it contains. Labor cost and machine hours are ignored. (In traditional accounting terms, money spent is called *value added*.)

Although this is often an arguable point, using only the raw material cost is a conservative view. When the value-added method (which includes all costs of production) is used, inventory is inflated and presents some serious income and balance sheet problems. Consider, for example, work-in-process or finished goods inventory that has become obsolete, or for which a contract was canceled. A management decision to declare large amounts of inventory as scrap is difficult because it is often carried on the books as an asset even though it may really have no value. Using just raw materials cost also avoids the problem of determining which costs are direct and which are indirect.

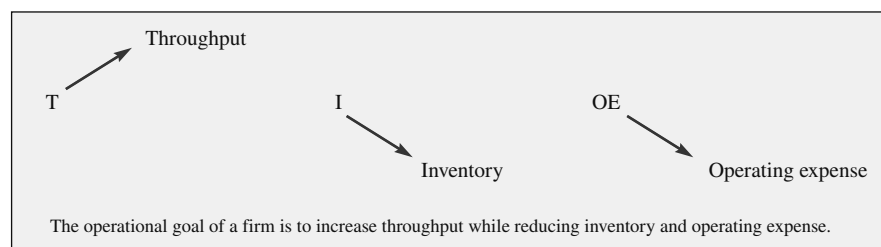
Operating expenses include production costs (such as direct labor, indirect labor, inventory carrying costs, equipment depreciation, and materials and supplies used in production) and administrative costs. The key difference here is that there is no need to separate direct and indirect labor.

As shown in Exhibit 22S.3, the objective of a firm is to treat all three measurements simultaneously and continually; this achieves the goal of making money.

From an operations standpoint, the goal of the firm is to

INCREASE THROUGHPUT WHILE SIMULTANEOUSLY REDUCING
INVENTORY AND REDUCING OPERATING EXPENSE.

Exhibit 22S.3 Operational Goal



Productivity Typically, **productivity** is measured in terms of output per labor hour. However, this measurement does not ensure that the firm will make money (for example, when extra output is not sold but accumulates as inventory). To test whether productivity has increased, we should ask these questions: Has the action taken increased throughput? Has it decreased inventory? Has it decreased operational expense? This leads us to a new definition:

PRODUCTIVITY IS ALL THE ACTIONS THAT BRING A COMPANY CLOSER TO ITS GOALS.

Productivity

A measure of how well resources are used. According to Goldratt's definition, all the actions that bring a company closer to its goals.

Unbalanced Capacity

Historically (and still typically in most firms), manufacturers have tried to balance capacity across a sequence of processes in an attempt to match capacity with market demand. However, this is the wrong thing to do—*unbalanced capacity* is better.

Consider a simple process line with several stations, for example. Once the output rate of the line has been established, production people try to make the capacities of all stations the same. This is done by adjusting machines or equipment used, workloads, skill and type of labor assigned, tools used, overtime budgeted, and so on.

In synchronous manufacturing thinking, however, making all capacities the same is viewed as a bad decision. Such a balance would be possible only if the output times of all stations were constant or had a very narrow distribution. A normal variation in output times causes downstream stations to have idle time *when upstream stations take longer to process*. Conversely, when upstream stations process in a shorter time, inventory builds up between the stations. The effect of the statistical variation is cumulative. The only way that this variation can be smoothed is by increasing work-in-process to absorb the variation (a bad choice because we should be trying to reduce work-in-process) or increasing capacities downstream to be able to make up for the longer upstream times. The rule here is that capacities within the process sequence should not be balanced to the same levels. Rather, attempts should be made to balance the flow of product through the system. When flow is balanced, capacities are unbalanced. This idea is further explained in the next section.

Dependent Events and Statistical Fluctuations The term *dependent events* refers to a process sequence. If a process flows from A to B to C to D, and each process must be completed before passing on to the next step, then B, C, and D are dependent events. The ability to do the next process is dependent on the preceding one.

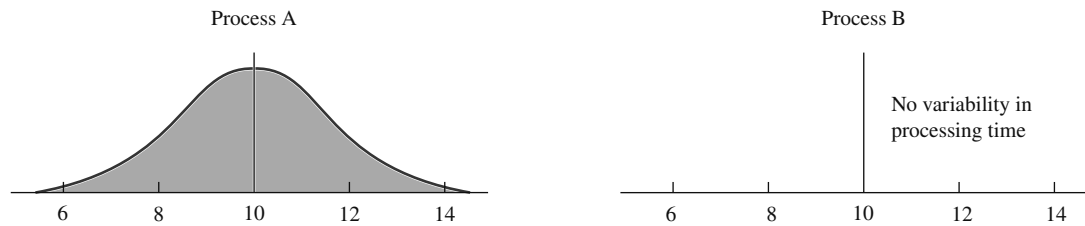
Statistical fluctuation refers to the normal variation about a mean or average. When statistical fluctuations occur in a dependent sequence without any inventory between workstations, there is no opportunity to achieve the average output. When one process takes longer than the average, the next process cannot make up the time. We follow through an example of this to show what could happen.

Suppose we wanted to process five items that could come from the two distributions in Exhibit 22S.4. The processing sequence is from A to B, with no space for inventory in between. Process A has a mean of 10 hours and a standard deviation of 2 hours. This means we would expect 95.5 percent of the processing time to be between 6 hours and 14 hours (plus or minus 2 sigma). Process B has a constant processing time of 10 hours.

We see that the last item was completed in 66 hours, for an average of 13.2 hours per item, although the expected time of completion was 60, for an average of 12 hours per item (taking into account the waiting time for the first unit by Process B).

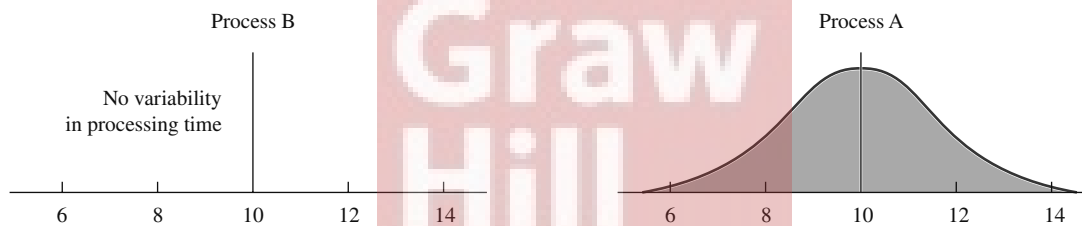
Suppose we reverse the processes—B feeds A. To illustrate the possible delays, we also reverse A's performance times. (See Exhibit 22S.5.) Again, the completion time of the last item is greater than the average (13.2 hours rather than 12 hours). Process A and Process B have the same average performance time of 10 hours, and yet performance is late. In neither case could we achieve the expected average output rate. Why? Because the time lost when the second process is idle cannot be made up.

This example is intended to challenge the theory that capacities should be balanced to an average time. *Rather than capacities being balanced, the flow of product through the system should be balanced.*

Exhibit 22S.4 Dependent Events Analysis—Process A Feeding Process B

Item Number	Start Time	Processing Time	Finish Time	Item Number	Start Time	Processing Time	Finish Time
1	0 hrs	14 hrs	14 hrs	1	14 hrs	10 hrs	24 hrs
2	14	12	26	2	26	10	36
3	26	10	36	3	36	10	46
4	36	8	44	4	46	10	56
5	44	6	50	5	56	10	66
Average = 10 hours				Average = 10 hours			

Here the flow is from Process A (on the left) to Process B (on the right). Process A has a mean of 10 hours and a standard deviation of 2 hours; Process B has a constant 10-hour processing time.

Exhibit 22S.5 Dependent Events Analysis—Process B Feeding Process A

Item Number	Start Time	Processing Time	Finish Time	Item Number	Start Time	Processing Time	Finish Time
1	0 hrs	10 hrs	10 hrs	1	10 hrs	6 hrs	16 hrs
2	10	10	20	2	20	8	28
3	20	10	30	3	30	10	40
4	30	10	40	4	40	12	52
5	40	10	50	5	52	14	66
Average = 10 hours				Average = 10 hours			

This is similar to Exhibit 22S.4. However, the processing sequence has been reversed, as well as the order of the Process A times.

Bottlenecks, Capacity-Constrained Resources, and Synchronous Manufacturing

LO22S-2

Analyze bottleneck resources and apply TOC principles to controlling a process

A **bottleneck** is defined as any resource whose capacity is less than the demand placed upon it. A bottleneck is a constraint within the system that limits throughput. It is that point in the manufacturing process where flow thins to a narrow stream. A bottleneck may be a machine, scarce or highly skilled labor, or a specialized tool. Observations in industry have shown that most plants have very few bottleneck operations.

If there is no bottleneck, then excess capacity exists and the system should be changed to create a bottleneck (such as more setups or reduced capacity), which we will discuss later.

Capacity is defined as the available time for production. This excludes maintenance and other downtime. A **nonbottleneck** is any resource whose capacity is greater than the demand placed on it. A nonbottleneck, therefore, should not be working constantly because it can produce more than is needed. A nonbottleneck contains idle time.

A **capacity-constrained resource (CCR)** is one whose utilization is close to capacity and could be a bottleneck if it is not scheduled carefully. For example, a CCR may be receiving work in a job-shop environment from several sources. If these sources schedule their flow in a way that causes occasional idle time for the CCR in excess of its unused capacity time, the CCR becomes a bottleneck when the surge of work arrives at a later time. This can happen if batch sizes are changed or if one of the upstream operations is not working for some reason and does not feed enough work to the CCR.

Bottleneck

A resource that limits the capacity or maximum output of the process.

Nonbottleneck

Any resource whose capacity is greater than the demand placed on it (Goldratt's definition).

Capacity-constrained resource (CCR)

A resource whose utilization is close to capacity and could be a bottleneck if not scheduled carefully (Goldratt's definition).

Basic Manufacturing Building Blocks

All manufacturing processes and flows can be simplified to four basic configurations, as shown in Exhibit 22S.6. In configuration A of Exhibit 22S.6, product that flows through Process X feeds into Process Y. In configuration B, Y is feeding X. In C, Process X and Process Y are creating subassemblies, which are then combined, say, to feed the market demand. In D, Process X and Process Y are independent of each other and are supplying their own markets. The last column in the exhibit shows possible sequences of nonbottleneck resources, which can be grouped and displayed as Y to simplify the representation.

The value in using these basic building blocks is that a production process can be greatly simplified for analysis and control. Rather than track and schedule all of the steps in a production sequence through nonbottleneck operations, for example, attention can be placed at the beginning and end points of the building block groupings.

Methods for Synchronous Control

Exhibit 22S.7 shows how bottleneck and nonbottleneck resources should be managed.

Resource X and Resource Y are workcenters that can produce a variety of products. Each of these workcenters has 200 hours available per month. For simplicity, assume that we are dealing with only one product and we will alter the conditions and makeup for four different situations. Each unit of X takes one hour of production time, and the market demand is

Exhibit 22S.6 Recognizing the Bottleneck—Four Configurations

Description	Basic Building Blocks Simplified by Grouping Nonbottlenecks	Original Representation
A. Bottleneck feeding nonbottleneck	$X \rightarrow Y \rightarrow \text{Market}$	$X \rightarrow A \rightarrow B \rightarrow C \rightarrow D \rightarrow \text{Market}$
B. Nonbottleneck feeding bottleneck	$Y \rightarrow X \rightarrow \text{Market}$	$A \rightarrow B \rightarrow C \rightarrow D \rightarrow X \rightarrow \text{Market}$
C. Output of bottleneck and nonbottleneck assembled into a product	$X \rightarrow \text{Final Assembly} \rightarrow \text{Market}$ $Y \rightarrow \text{Assembly} \rightarrow \text{Market}$	$A \rightarrow B \rightarrow C \rightarrow D \rightarrow \text{Final Assembly} \rightarrow \text{Market}$
D. Bottleneck and nonbottleneck have independent markets for their output	$X \rightarrow \text{Market}$ $Y \rightarrow \text{Market}$	$A \rightarrow B \rightarrow C \rightarrow D \rightarrow \text{Market}$ $X \rightarrow \text{Market}$

X is a bottleneck.

Y is a nonbottleneck (has excess capacity).

200 units per month. Each unit of Y takes 45 minutes of production time, and the market demand is also 200 units per month.

Exhibit 22S.7A shows a bottleneck feeding a nonbottleneck. Product flows from Workcenter X to Workcenter Y. X is the bottleneck because it has a capacity of 200 units (200 hours/1 hour per unit) and Y has a capacity of 267 units (200 hours/45 minutes per unit). Because Y has to wait for X, and Y has a higher capacity than X, no extra product accumulates in the system. It all flows through to the market.

Exhibit 22S.7B is the reverse of A, with Y feeding X. This is a nonbottleneck feeding a bottleneck. Because Y has a capacity of 267 units and X has a capacity of only 200 units, we should produce only 200 units of Y (75 percent of capacity) or else work-in-process will accumulate in front of X.

Exhibit 22S.7C shows that the products produced by X and Y are assembled and then sold to the market. Because one unit from X and one unit from Y form an assembly, X is the bottleneck with 200 units of capacity and, therefore, Y should not work more than 75 percent or else extra parts will accumulate.

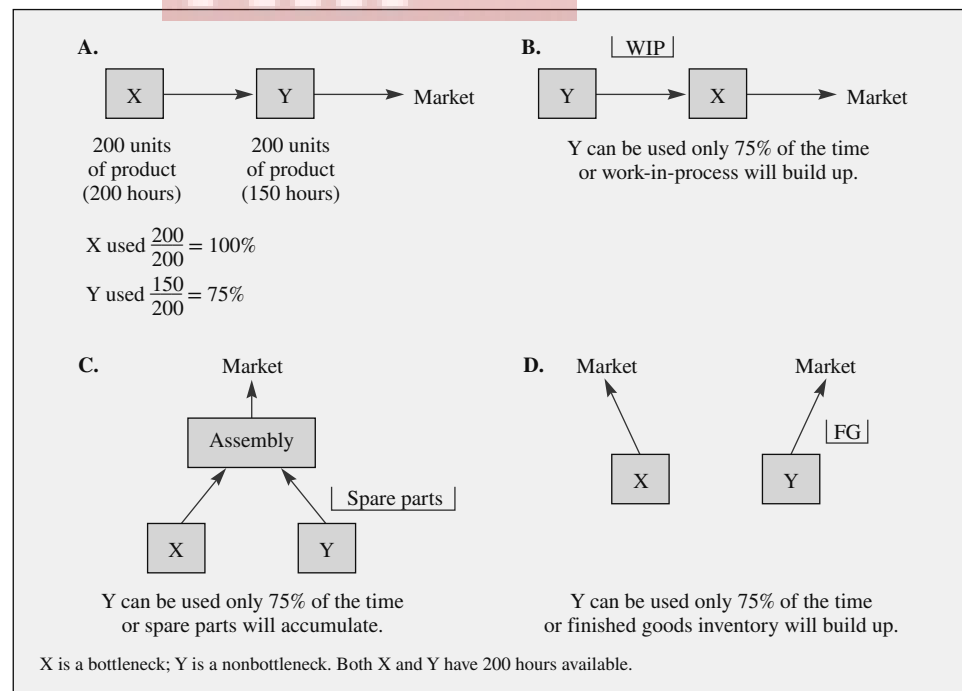
In Exhibit 22S.7D, equal quantities of product from X and Y are demanded by the market. In this case, we can call these products “finished goods” because they face independent demands. Here, Y has access to material independent of X and, with a higher capacity than needed to satisfy the market (in essence, the market is the bottleneck), it can produce more product than the market will take. However, this would create an inventory of unneeded finished goods.

The four situations just discussed demonstrate bottleneck and nonbottleneck resources and their relationships to production and market demand. They show that the industry practice of using resource utilization as a measure of performance can encourage the overuse of non-bottlenecks and result in excess inventories.

Time Components The following kinds of time make up production cycle time:

1. *Setup time*—the time that a part spends waiting for a resource to be set up to work on this same part.

Exhibit 22S.7 Synchronous Control Analysis



CASE
Tata Steel Rides on “Theory of Constraint”

Since 1868, Tata Group is now around 150 years old. Under its umbrella, there are 98 companies, which are supported by over 350,000 committed employees worldwide. It is undoubtedly the most global Indian business group and one of the best known Indian corporate brands across the globe. Tata Steel is its flagship company, which has provided leadership role in the corporate world of India. It is one of the leading steel producing companies of the world.

In the year 2001, mainly due to a slowdown in Indian economy, Tata Steel witnessed one of the toughest business environments in its history of existence. It was being called as the sunset industry. Indian steel companies were apparently at a brink of bankruptcy and Tata Steel was no exception.

Consequently, a new vision was created in 2002, which involved two aspects:

1. **EVA Positive Company:** EVA stands for “Economic-Value-Added”, which is the difference between “return on invested capital” and “weighted average cost of capital.” This part of new vision involved:

- Value creating partnerships with customers and suppliers,
- Moving from commodities to brands, and
- Developing EVA positive core business.

The challenge to become EVA-positive despite offering average steel prices was very difficult and no steel company has ever achieved it in the past. Tata Steel launched a new programme, called “ASPIRE”, to help it in making an EVA-positive growth.

2. **Improve the quality of life:** The second important aspect of the new vision was to improve the *quality of life* of its employees and the communities that are being served. Broadly, this involved:

- Enthused and happy employees, and
- Sustainable growth.

The Improvement Philosophy of Tata Steel [2]

The three major weaknesses in the operating practices of organizations on continuous improvement path were:

- (i) Approach towards continuous improvement was generally inward looking and relatively lesser attention was paid on distinctly differentiating the offerings to the customers,
- (ii) The objectives may not be aspirational (challenging),

- (iii) There was not much effort on testing the situations of causality during the journey of continuous improvement, i.e. the means are not checked for necessary and sufficiency conditions to meet the desired objective. Failure to test interventions against appropriate causality led to making the ‘one size fits all’ approach to target setting as well as the choice of tools and techniques.

To overcome these weaknesses, the ASPIRE (ASPIRational Initiatives to Retain Excellence) Program was designed to first set goals and then choose the appropriate tools for delivering these goals. Certain parts of the planned improvements were checked for causality, explored for means (causes) and then the implementations were carried out using differentiated tools and techniques. For the remaining parts of the aspirational goals, the means were innovated during the implementation journey [2].

In November 2006, Tata Steel came out with “ASPIRE-T3” initiative, which aimed at motivating the employees in dedicating themselves to 3 T’s, i.e., TOC (Theory of constraints), TQM (Total Quality Management), and Technology, for making the company a world leader in steel business [3]. TQM and Technology helped in achieving superior product quality through process consistency, process capability and technological capability. Similarly, TOC implementation helped it in improving the service quality [4,5,6]. To overcome many constraints in superior performance, ASPIRE-T3 is playing a major role.

“The focus on TOC, TQM and technology has been captured at Aspire with T3 (‘Teen se jeet’). The synergy and integration in the efforts of the employees through TOC, TQM and technology initiatives will help TATA Steel in becoming one of the top steel companies in the world...”

—B. Muthuraman, MD, Tata Steel on launch of Aspire T3, November 02, 2006

Opportunities

The ASPIRE-T3 evolves round the philosophy of setting aspirations (which may be beyond the current capability), then getting determined about reaching these aspirations, mustering required resources, which leads to achievement, and thus the self confidence to aspire for more. It was launched in the presence of Dr. Eliyahu M Goldratt, who explained the fundamentals of TOC to Tata Steel and explained that the TOC approach looks at

business problems at a system level, in contrast to some other methods that are focused on specific processes. Focusing on viable vision implementations, using TOC, industrial organizations could achieve unprecedented improvements in their bottom-line performances. He emphasized the need to identify and capitalize on those few elements that govern an organizational system and are constraining in the goal of the organizational performance indicators.

TOC approach has helped Tata Steel in its major restructuring of supply chain and replenishment process. Due to very high dependency on the demands from industrial clusters and other major institutional clients, its goal of making profit is adversely affected due to high inventory and stock-outs. In the language of TOC, major constraint for inferior SCM performance is demand uncertainty. To deal with this constraint, marketing and sales teams of Tata Steel are using replenishment-offers to distributors while vendor managed inventory (VMI)-offers to original equipment (OEM) customers, and rapid-response-offers to some other selected customers [2]. Backed by a product portfolio covered under these offers, reliable replenishments, warehousing, inventory management, and end-to-end communications for better planning and decision making, Tata Steel has started making considerable achievements from ASPIRE-T3. It has helped bringing down the average distributor inventory of cold rolled sheets and galvanized sheets by as much as 20 per cent [2, 3]. The ASPIRE program is thus successful in supporting an accelerated pace of growth and making Tata Steel an EVA positive on a sustainable basis.

There are a few learning lessons (*Guru-mantras* (GMs)) from TOC and this case study.

GM 1: Three Ts make a formidable combination for operational success. These are: Theory of Constraints (TOC), Total Quality Management (TQM) and Technology.

GM 2: The goal of a firm is to make money.

GM 3: Three financial measures of firm help firm's ability to make money. These financial measures are: (i) Net profit, which is an absolute measure of money, (ii) Return On Investment (ROI), which is a relative measure, based

on investment, and (iii) Cash flow, which is a survival measure.

GM 4: For operational measurement in TOC, three sets of measurement are important. These are: (i) Throughput, which is the rate at which money is generated by the system through sales, (ii) Inventory, which represents all the money that the system has invested in purchasing things that it intends to sell, and (iii) Operating expenses, which is all the money that the system spends to turn inventory into throughput.

GM 5: Productivity is all the actions that bring a company closer to its goals.

GM 6: An hour saved at the bottleneck adds an extra hour to the entire production system.

GM 7: An hour saved at a non-bottleneck is a mirage and only adds an extra hour to its idle time.

Discussion Questions

1. Explain the approach followed by Tata Steel to implement Theory of Constraints.
2. Which other industry is interested in implementing Theory of Constraints?

References

- 1 http://www.tatasteel.com/investors/annual-report-2008-09/html/financial_highlights.html (Accessed on 12 May 2010).
- 2 http://www.tatasteel.com/investors/annualreport-2007-08/html/process_improvement.html (Accessed on 02 March 2014).
- 3 <http://www.tata.com/article.aspx?artid=cIIfRf+toMc=> (Accessed on 09 May 2010).
- 4 <http://www.indiaprwire.com/pressrelease/mining-metals/20061102908.htm> (Accessed on 02 March 2014).
- 5 Business Line Article on 'Tata Steel Revamping Supply Chain Management' by Santanu Sanyal, 22nd Jul 2006 available on <http://www.blonnet.com/2006/07/22/stories/2006072202080700.htm> (Accessed on 12 May 2010).
- 6 <http://www.tata.com/company/Articles/inside.aspx?artid=7VXdehaS0IY=> (accessed on 12 May 2010).

2. *Processing time*—the time that the part is being processed.
3. *Queue time*—the time that a part waits for a resource while the resource is busy with something else.
4. *Wait time*—the time that a part waits not for a resource but for another part so that they can be assembled together.

5. *Idle time*—the unused time; that is, the cycle time minus the sum of the setup time, processing time, queue time, and wait time.

For a part waiting to go through a bottleneck, queue time is usually the greatest. As we discuss later in this chapter, this is because the bottleneck has a fairly large amount of work to do in front of it (to make sure it is always working). For a nonbottleneck, wait time is usually the greatest. The part is just sitting there waiting for the arrival of other parts so that an assembly can take place.

Schedulers are tempted to save setup times. Suppose that the batch sizes are doubled to save half the setup times. Then, with a double batch size, all of the other times (processing time, queue time, and wait time) increase twofold. Because these times are doubled while saving only half of the setup time, the net result is that the work-in-process is approximately doubled, as is the investment in inventory.

Finding the Bottleneck There are two ways to find the bottleneck (or bottlenecks) in a system. One is to run a capacity resource profile; the other is to use our knowledge of the particular plant, look at the system in operation, and talk with supervisors and workers.

A capacity resource profile is obtained by looking at the loads placed on each resource by the products that are scheduled through them. In running a capacity profile, we assume that the data are reasonably accurate, although not necessarily perfect. As an example, consider that products have been routed through Resources M1 through M5. Suppose that our first computation of the resource loads on each resource caused by these products shows the following:

M1	130 percent of capacity
M2	120 percent of capacity
M3	105 percent of capacity
M4	95 percent of capacity
M5	85 percent of capacity

For this first analysis, we can disregard any resources at lower percentages because they are nonbottlenecks and should not be a problem. With this list in hand, we should physically go to the facility and check all five operations. Note that M1, M2, and M3 are overloaded; that is, they are scheduled above their capacities. We would expect to see large quantities of inventory in front of M1. If this is not the case, errors must exist somewhere—perhaps in the bill-of-materials or in the routing sheets. Let's say that our observations and discussions with shop personnel showed that there were errors in M1, M2, M3, and M4. We tracked them down, made the appropriate corrections, and ran the capacity profile again:

M1	110 percent of capacity
M2	115 percent of capacity
M3	105 percent of capacity
M4	90 percent of capacity
M5	85 percent of capacity

M1, M2, and M3 are still showing a lack of sufficient capacity, but M2 is the most serious. If we now have confidence in our numbers, we use M2 as our bottleneck. If the data contain too many errors for a reliable data analysis, it may not be worth spending time (it could take months) making all the corrections.

Saving Time Recall that a bottleneck is a resource whose capacity is less than the demand placed on it. Because we focus on bottlenecks as restricting *throughput* (defined as *sales*), a

bottleneck's capacity is less than the market demand. There are a number of ways we can save time on a bottleneck (better tooling, higher-quality labor, larger batch sizes, reduction in setup times, and so forth), but how valuable is the extra time? Very, very valuable!

AN HOUR SAVED AT THE BOTTLENECK ADDS AN EXTRA HOUR
TO THE ENTIRE PRODUCTION SYSTEM.

How about time saved on a nonbottleneck resource?

AN HOUR SAVED AT A NONBOTTLENECK IS A MIRAGE AND
ONLY ADDS AN HOUR TO ITS IDLE TIME.

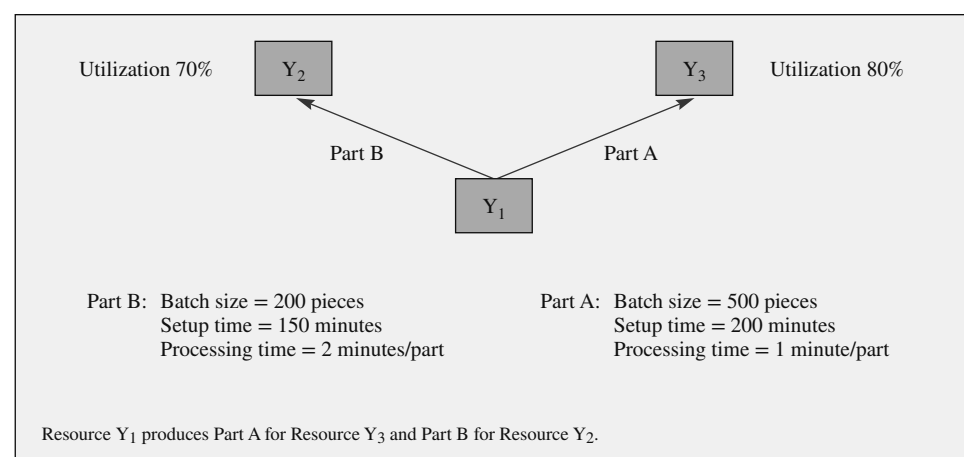
Because a nonbottleneck has more capacity than the system needs for its current throughput, it already contains idle time. Implementing any measures to save more time does not increase throughput but only serves to increase its idle time.

Avoid Changing a Nonbottleneck into a Bottleneck When nonbottleneck resources are scheduled with larger batch sizes, this action could create a bottleneck that we certainly would want to avoid. Consider the case in Exhibit 22S.8, where Y_1 , Y_2 , and Y_3 are nonbottleneck resources. Y_1 currently produces Part A, which is routed to Y_3 , and Part B, which is routed to Y_2 . To produce Part A, Y_1 has a 200-minute setup time and a processing time of 1 minute per part. Part A is currently produced in batches of 500 units. To produce Part B, Y_1 has a setup time of 150 minutes and 2 minutes' processing time per part. Part B is currently produced in batches of 200 units. With these parameters, it takes about 700 minutes to produce a batch of A ($500 + 200 \times 1$) and 550 minutes to produce a batch of B ($150 + 200 \times 2$). With the company's current volume Y_2 is utilized 70 percent of the time and Y_3 is utilized 80 percent of the time.

Because setup time is 200 minutes for Y_1 on Part A, both worker and supervisor mistakenly believe that more production can be gained if fewer setups are made. Let's assume that the batch size is increased to 1,500 units on Part A and see what happens. The illusion is that we have saved 400 minutes of setup. (Instead of three setups taking 600 minutes to produce three batches of 500 units each, there is just one setup with a 1,500-unit batch.)

The problem is that the 400 minutes saved served no purpose since this delay interferes with the production of Part B. Recall that Y_1 produces Part B for Y_2 . The sequence before any changes were made was Part A (700 minutes), Part B (550 minutes), Part A (700 minutes),

Exhibit 22S.8 Batch Size Impact On Utilization



Part B (550 minutes), and so on. Now, however, when the Part A batch is increased to 1,500 units (1,700 minutes), Y_2 and Y_3 could well be starved for work and have to wait more time than they have available. Now the new sequence is Part A (700 minutes), Part B (1,700 minutes), Part A (700 minutes), Part B (1,700 minutes), etc. Such an extended wait for Y_2 and Y_3 could be disruptive. Y_2 and Y_3 could become temporary bottlenecks and lose throughput for the system.

Drum, Buffer, Rope Every production system needs some control point or points to control the flow of product through the system. If the system contains a bottleneck, the bottleneck is the best place for control. This control point is called the *drum* because it strikes the beat that the rest of the system (or those parts that it influences) uses to function. Recall that a *bottleneck* is defined as a resource that does not have the capacity to meet demand. Therefore, a bottleneck is working all the time, and one reason for using it as a control point is to make sure that the operations upstream do not overproduce and build up excess work-in-process inventory that the bottleneck cannot handle.

If there is no bottleneck, the next-best place to set the drum would be a capacity-constrained resource (CCR). A capacity-constrained resource, remember, is one that is operating near capacity but, on average, has adequate capability as long as it is not incorrectly scheduled (for example, with too many setups, causing it to run short of capacity, or producing too large a lot size, thereby starving downstream operations).

If neither a bottleneck nor a CCR is present, the control point can be designated anywhere. The best position would generally be at some divergent point where the output of the resource is used in several downstream operations.

Dealing with the bottleneck is most critical, and our discussion focuses on ensuring that the bottleneck always has work to do. Exhibit 22S.9 shows a simple linear flow A through G. Suppose that Resource D, which is a machine center, is a bottleneck. This means the capacities are greater both upstream and downstream from it. If this sequence is not controlled, we would expect to see a large amount of inventory in front of Workcenter D and very little anywhere else. There would be little finished goods inventory because (by the definition of the term *bottleneck*) all the product produced would be taken by the market.

There are two things we must do with this bottleneck:

1. Keep a *buffer* inventory in front of it to make sure it always has something to work on. Because it is a bottleneck, its output determines the throughput of the system.
2. Communicate back upstream to A what D has produced so that A provides only that amount. This keeps inventory from building up. This communication is called the *rope*. It can be formal (such as a schedule) or informal (such as daily discussion).

The buffer inventory in front of a bottleneck operation is a *time buffer*. We want to make sure that Workcenter D always has work to do, and it does not matter which of the scheduled

Exhibit 22S.9 Drum, Buffer, Rope Bottleneck Example

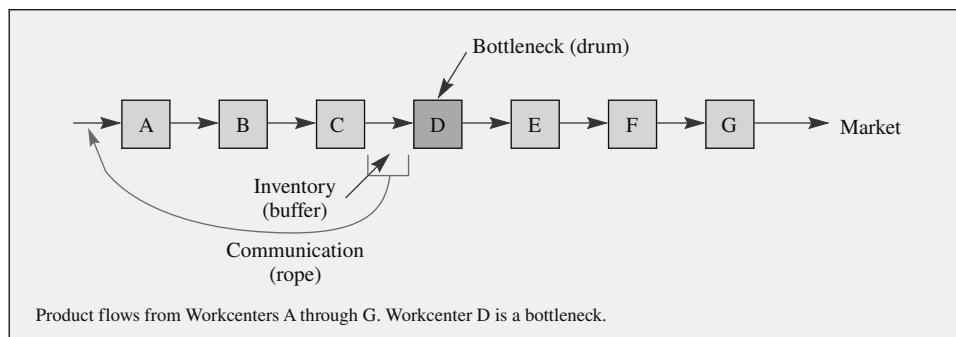
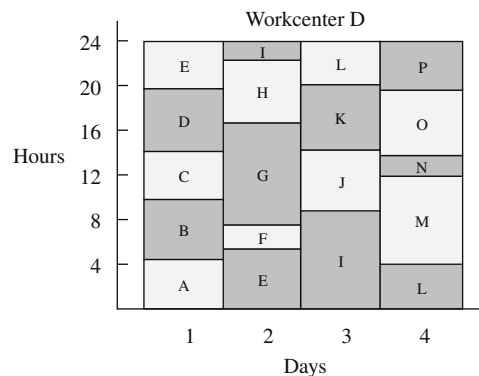


Exhibit 22S.10 Capacity Profile of Workcenter D (showing assigned Jobs A through P over a period of four 24-hour days)



products are worked on. We might, for example, provide 96 hours of inventory in the buffer as shown in the sequence A through P in Exhibit 22S.10. Jobs A through about half of E are scheduled during the 24 hours of Day 1; Jobs E through a portion of Job I are scheduled during the second 24-hour day; Jobs I through part of L are scheduled during the third 24-hour day; and Jobs L through P are scheduled during the fourth 24-hour day, for a total of 96 hours. This means that through normal variation, or if something happens upstream and the output has been temporarily stalled, D can work for another 96 hours, protecting the throughput. (The 96 hours of work, incidentally, include setups and processing times contained in the job sheets, which usually are based on engineering standard times.)

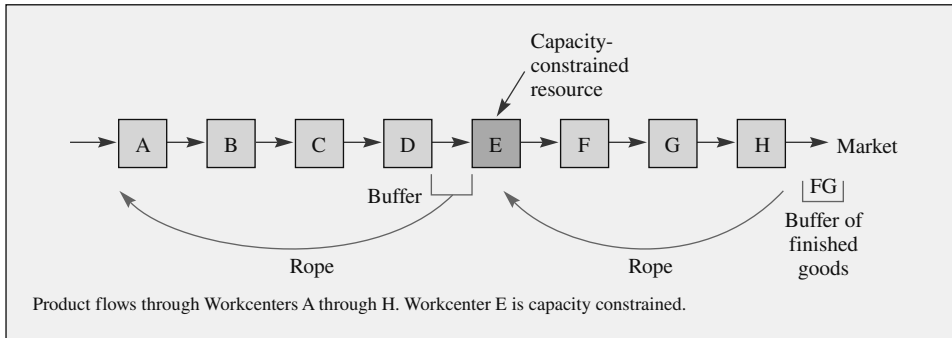
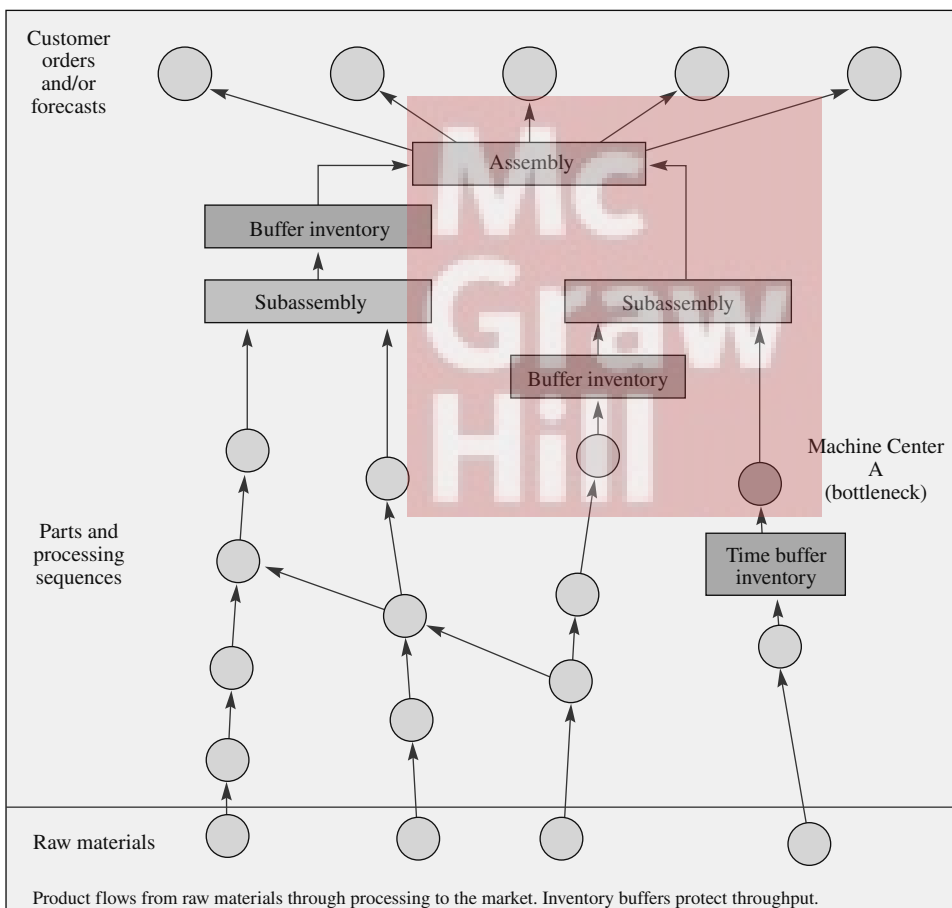
We might ask, How large should the time buffer be? The answer: As large as it needs to be to ensure that the bottleneck continues to work. By examining the variation of each operation, we can make a guess. Theoretically, the size of the buffer can be computed statistically by examining past performance data, or the sequence can be simulated. In any event, precision is not critical. We could start with an estimate of the time buffer as one-fourth of the total lead time of the system. Say, the sequence A to G in our example (Exhibit 22S.9) took a total of 16 days. We could start with a buffer of four days in front of D. If, during the next few days or weeks, the buffer runs out, we need to increase the buffer size. We do this by releasing extra material to the first operation, A. On the other hand, if we find that our buffer never drops below three days, we might want to hold back releases to A and reduce the time buffer to three days. Experience is the best determination of the final buffer size.

If the drum is not a bottleneck but a CCR (and thus it can have a small amount of idle time), we might want to create two buffer inventories: one in front of the CCR and the second at the end as finished goods. (See Exhibit 22S.11.) The finished goods inventory protects the market, and the time buffer in front of the CCR protects throughput. For this CCR case, the market cannot take all that we can produce, so we want to ensure that finished goods are available when the market does decide to purchase.

We need two ropes in this case: (1) a rope communicating from finished goods inventory back to the drum to increase or decrease output and (2) a rope from the drum back to the material release point, specifying how much material is needed.

Exhibit 22S.12 is a more detailed network flow showing one bottleneck. Inventory is provided not only in front of that bottleneck but also after the nonbottleneck sequence of processes that feed the subassembly. This ensures that the flow of product is not slowed down by having to wait after it leaves the bottleneck.

Importance of Quality An MRP system allows for rejects by building a larger batch than actually needed. A JIT system cannot tolerate poor quality because JIT success is based on a balanced capacity. A defective part or component can cause a JIT system to shut down, thereby losing throughput of the total system. Synchronous manufacturing, however, has

Exhibit 22S.11 Drum, Buffer, Rope Capacity-Constrained Resource Example**Exhibit 22S.12** Network Flow with One Bottleneck

excess capacity throughout the system except for the bottleneck. If a bad part is produced upstream of the bottleneck, the result is that there is a loss of material only. Because of the excess capacity, there is still time to do another operation to replace the one just scrapped. For the bottleneck, however, extra time does not exist, so there should be a quality control inspection just prior to the bottleneck to ensure that the bottleneck works only on good product. Also, there needs to be assurance downstream from the bottleneck that the passing product is not scrapped—that would mean lost throughput.

Batch Sizes In an assembly line, what is the batch size? Some would say “one” because one unit is moved at a time; others would say “infinity” because the line continues to produce the same item. Both answers are correct, but they differ in their point of view. The first answer, “one,” in an assembly line focuses on the *part* transferred one unit at a time. The second focuses on the *process*. From the point of view of the resource, the process batch is infinity because it is continuing to run the same units. Thus, in an assembly line, we have a *process batch* of infinity (or all the units until we change to another process setup) and a *transfer batch* of one unit.

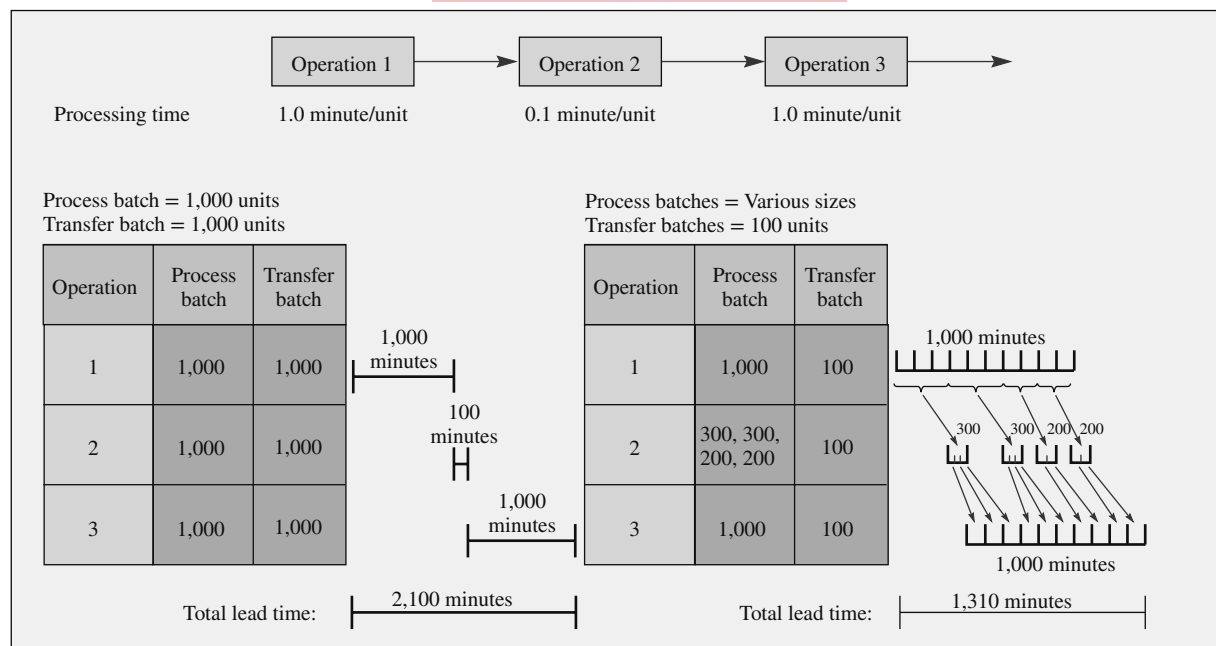
Setup costs and carrying costs were treated in depth in Chapter 20, “Inventory Management.” In the present context, setup costs relate to the process batch, and carrying costs relate to the transfer batch.

A process batch is of a size large enough or small enough to be processed in a particular length of time. From the point of view of a resource, two times are involved: setup time and processing run time (ignoring downtime for maintenance or repair). Larger process batch sizes require fewer setups and therefore can generate more processing time and more output. For bottleneck resources, larger batch sizes are desirable. For nonbottleneck resources, smaller process batch sizes are desirable (by using up the existing idle time), thereby reducing work-in-process inventory.

Transfer batches refer to the movement of part of the process batch. Rather than wait for the entire batch to be finished, work that has been completed by that operation can be moved to the next downstream workstation so it can begin working on that batch. A transfer batch can be equal to a process batch, but it should not be larger under a properly designed system. This could occur only if a completed process batch were held until sometime later when a second batch was processed. If this later time was acceptable in the beginning, then both jobs should be combined and processed together at the later time.

The advantage of using transfer batches that are smaller than the process batch quantity is that the total production time is shorter so the amount of work-in-process is smaller. Exhibit 22S.13 shows a situation where the total production lead time was reduced from 2,100 to 1,310 minutes by (1) using a transfer batch size of 100 rather than 1,000 and (2) reducing the process batch sizes of Operation 2.

Exhibit 22S.13 Transfer Batch Flow Analysis



How to Determine Process Batch and Transfer Batch Sizes Logic would suggest that the master production schedule (however it was developed) be analyzed as to its effect on various workcenters. In an MRP system, this means that the master production schedule should be run through the MRP and the CRP (capacity requirements planning program) to generate a detailed load on each workcenter. From this report, probable CCRs and bottlenecks can be identified. There should be only one (or a few), and they should be reviewed by managers so that they understand which resources are actually controlling their plant. These resources set the drumbeat.

Rather than try to adjust the master production schedule to change resource loads, it is more practical to control the flow at each bottleneck or CCR to bring the capacities in line. The process batch sizes and transfer batch sizes are changed after comparing past performances in meeting due dates.

Smaller transfer batches give lower work-in-process inventory and faster product flow (and consequently shorter lead time). More material handling is required, however. Larger transfer batches give longer lead times and higher inventories, but there is less material handling. Therefore, the transfer batch size is determined by a trade-off of production lead times, inventory reduction benefits, and costs of material movement.

When trying to control the flow at CCRs and bottlenecks, there are four possible situations:

1. A bottleneck (no idle time) with no setup time required when changing from one product to another.
2. A bottleneck with setup time required to change from one product to another.
3. A capacity-constrained resource with a small amount of idle time with no setup time required to change from one product to another.
4. A CCR with setup time required when changing from one product to another.

In the first case (a bottleneck with no setup time to change products), jobs should be processed in the order of the schedule so that delivery is on time. Without setups, only the sequence is important. In the second case, when setups are required, larger batch sizes combine separate similar jobs in the sequence. This means reaching ahead into future time periods. Some jobs will therefore be done early. Because this is a bottleneck resource, larger batches save setups and thereby increase throughput. (The setup time saved is used for processing.) The larger process batches may cause the early scheduled jobs to be late. Therefore, frequent small transfer batches are necessary to try to shorten the lead time.

Situations 3 and 4 include a CCR without a setup and a CCR with setup time requirements, respectively. Handling the CCR would be similar to handling a nonbottleneck, though more carefully. That is, a CCR has some idle time. It would be appropriate here to cut the size of some of the process batches so that there can be more frequent changes of product. This would decrease lead time, and jobs would be more likely to be done on time. In a make-to-stock situation, cutting process batch sizes has a much more profound effect than increasing the number of transfer batches. This is because the resulting product mix is much greater, leading to reduced work-in-process (WIP) and production lead time.

How to Treat Inventory The traditional view of inventory is that its only negative impact on a firm's performance is its carrying cost. We now realize that inventory's negative impact also comes from lengthening lead times and creating problems with engineering changes. (When an engineering change on a product comes through, which commonly occurs, product still within the production system often must be modified to include the changes. Therefore, less work-in-process reduces the amount of product rework necessary.)

From a constraint management perspective, inventory is a loan given to the manufacturing unit. The value of the loan is based only on the purchased items that are part of the inventory. As we stated earlier, inventory is treated in this chapter as material cost only, without any accounting-type value added from production. If inventory is carried as a loan to manufacturing, we need a way to measure how long the loan is carried. One measurement is dollar days.

OSCM at Work

Critical Chain Project Management

Critical Chain Project Management is the name of the approach that Eli Goldratt developed for scheduling and managing projects. The approach borrows many ideas from those used for manufacturing processes. The conventional critical path method was covered in Chapter 4, and Goldratt goes beyond those ideas by considering resource constraints and special time buffers in the project. The following are specific ideas included in his Critical Chain Project Management approach:

1. Schedules are level loaded based on the limitations of available resources (constraints). This produces the “critical chain”—the longest set of sequential tasks (due to both task dependency and resource contention)—which dictates the shortest overall project duration.
2. Time buffers are inserted at strategic locations in the plan—at the end of the critical chain and at every point where a task enters the critical chain—to absorb the adverse effects of uncertainty without damaging performance. To create the buffers, some of the slack time built into tasks in planning is repositioned to these strategic locations.
3. Projects are “pipelined” or staged based on resource availability to combat the cascade effect of shared resources across projects and create viable multiproject plans.
4. Buffer management is used to dynamically set task priorities in execution. As uncertainty changes the original plan, tasks are prioritized based on the buffer burn rate (the amount of buffer consumed versus the percentage of the work complete). Tasks with critical buffer penetration take precedence over those with lower burn rates.

Dollar Days A useful performance measurement is the concept of *dollar days*, a measurement of the value of inventory and the time it stays within an area. To use this measure, we could simply multiply the total value of inventory by the number of days that inventory spends within a department.

Suppose Department X carries an average inventory of \$40,000, and, on average, the inventory stays within the department five days. In dollar days, Department X is charged with \$40,000 times five days, or \$200,000 dollar days of inventory. At this point, we cannot say whether the \$200,000 is high or low, but it does show where the inventory is located. Management can then see where it should focus attention and determine acceptable levels. Techniques can be instituted to try to reduce the number of dollar days while being careful that such a measure does not become a local objective (that is, minimizing dollar days) and hurt the global objectives (such as increasing return on investment, cash flow, and net profit).

Dollar days could be beneficial in a variety of ways. Consider the current practice of using efficiencies or equipment utilization as a performance measurement. To get high utilization, large amounts of inventory are held to keep everything working. However, high inventories would result in a high number of dollar days, which would discourage high levels of work-in-process. Dollar day measurements also could be used in other areas:

- Marketing—to discourage holding large amounts of finished goods inventory. The net result would be to encourage the sales of finished products.
- Purchasing—to discourage placing large purchase orders that on the surface appear to take advantage of quantity discounts. This would encourage just-in-time purchasing.
- Manufacturing—to discourage large work-in-process and producing earlier than needed. This would promote rapid flow of material within the plant.
- Project management—to quantify a project’s limited resource investments as a function of time. This promotes the proper allocation of resources to competing projects. See the OSCM at Work box titled “Critical Chain Project Management” for Goldratt’s ideas on scheduling projects.

Comparing Synchronous Manufacturing (TOC) to Traditional Approaches

LO22S-3

Compare TOC to conventional approaches.

MRP and JIT

MRP uses *backward scheduling* after having been fed a master production schedule. MRP schedules production through a bill-of-materials explosion in a backward manner—working backward in time from the desired completion date. As a secondary procedure,

MRP through its capacity resource planning module develops capacity utilization profiles of workcenters. When workcenters are overloaded, either the master production schedule must be adjusted or enough slack capacity must be left unscheduled in the system so that work can be smoothed at the local level (by workcenter supervisors or the workers themselves). Trying to smooth capacity using MRP is so difficult and would require so many computer runs that capacity overloads and underloads are best left to local decisions, such as at the machine centers. An MRP schedule becomes invalid just days after it was created.

The synchronous manufacturing approach uses *forward scheduling* because it focuses on the critical resources. These are scheduled forward in time, ensuring that loads placed on them are within capacity. The noncritical (or nonbottleneck) resources are then scheduled to support the critical resources. This can be done backward to minimize the length of time that inventories are held. This procedure ensures a feasible schedule. To help reduce lead time and work-in-process, the process batch size and transfer batch size in synchronous manufacturing are varied—a procedure that MRP is not able to do.

Comparing JIT to synchronous manufacturing, JIT does an excellent job of reducing lead times and work-in-process, but it has several drawbacks:

1. JIT is most often used in repetitive manufacturing environments.
2. JIT requires a stable production level (usually about a month long).
3. JIT does not allow very much flexibility in the products produced. (Products must be similar with a limited number of options.)
4. JIT still requires work-in-process when used with kanbans so that there is “something to pull.” This means that completed work must be stored on the downstream side of each workstation to be pulled by the next workstation.
5. Vendors need to be located nearby because the system depends on smaller, more frequent deliveries.

Because synchronous manufacturing uses a schedule to assign work to each workstation, there is no need for more work-in-process other than that being worked on. The exception is for inventory specifically placed in front of a bottleneck to ensure continual work, or at specific points downstream from a bottleneck to ensure flow of product.

Concerning continual improvements to the system, JIT is a trial-and-error procedure applied to a real system. In synchronous manufacturing, the system can be programmed and simulated on a computer because the schedules are realistic (can be accomplished) and computer run time is short.

Relationship with Other Functional Areas

The production system must work closely with other functional areas to achieve the best operating system. This section briefly discusses accounting and marketing—areas where conflicts can occur and where cooperation and joint planning should occur.

Accounting's Influence Sometimes we are led into making decisions to suit the measurement system rather than to follow the firm's goals. Consider the following example: Suppose that two old machines are currently being used to produce a product. The processing time for each is 20 minutes per part and because each has the capacity of three parts per hour, they have the combined capacity of six per hour, which exactly meets the market demand of six parts per hour. Suppose that engineering finds a new machine that produces parts in 12 minutes rather than 20. However, the capacity of this one machine is only five per hour, which does not meet the market demand. Logic would seem to dictate that the supervisor should use an old machine to make up the lacking of one unit per hour. However, the system does not allow this. The standard has been changed from 20 minutes each to 12 minutes each and performance would look very bad on paper because the variance would be 67 percent $[(20 - 12)/12]$ for units made on the old machines. The supervisor, therefore, would work the new machine on overtime.

Problems in Cost Accounting Measurements Cost accounting is used for performance measurement, cost determinations, investment justification, and inventory valuation. Two sets of accounting performance measurements are used for evaluation: (1) global measurements, which are financial statements showing net profit, return on investment, and cash flow (with which we agree) and (2) local cost accounting measurements showing efficiencies (as variances from standard) or utilization rate (hours worked/hours present).

From the cost accounting (local measurement) viewpoint, then, performance has traditionally been based on cost and full utilization. This logic forces supervisors to activate their workers all the time, which leads to excess inventory. The cost accounting measurement system also can instigate other problems. For example, attempting to use the idle time to increase utilization can create a bottleneck, as we discussed earlier in this chapter. Any measurement system should support the objectives of the firm but not stand in the way. Fortunately, the cost accounting measurement philosophy is changing.

Marketing and Production Marketing and production should communicate and conduct their activities in close harmony. In practice, however, they act very independently. There are many reasons for this. The difficulties range from differences in personalities and cultures to unlike systems of merits and rewards in the two functions. Marketing people are judged on the growth of the company in terms of sales, market share, and new products introduced. Marketing is sales oriented. Manufacturing people are evaluated on cost and utilization. Therefore, marketing wants a variety of products to increase the company's position, whereas manufacturing is trying to reduce cost.

Data used for evaluating marketing and manufacturing are also quite different. Marketing data are "soft" (qualitative); manufacturing data are "hard" (quantitative). The orientation and experiences of marketing and production people also differ. Those in marketing management have likely come up through sales and a close association with customers. Top manufacturing managers have likely progressed through production operations and therefore have plant performance as a top objective.

Cultural differences also can be important in contrasting marketing and manufacturing personnel. Marketing people tend to have a greater ego drive and are more outgoing. Manufacturing personnel tend to be more meticulous and perhaps more introverted (at least less extroverted than their marketing counterparts).

The solution to coping with these differences is to develop an equitable set of measurements to evaluate performance in each area and to promote strong lines of communication so they both contribute to reaching the firm's goals.

Theory of Constraints—Problems About What to Produce

LO22S-4

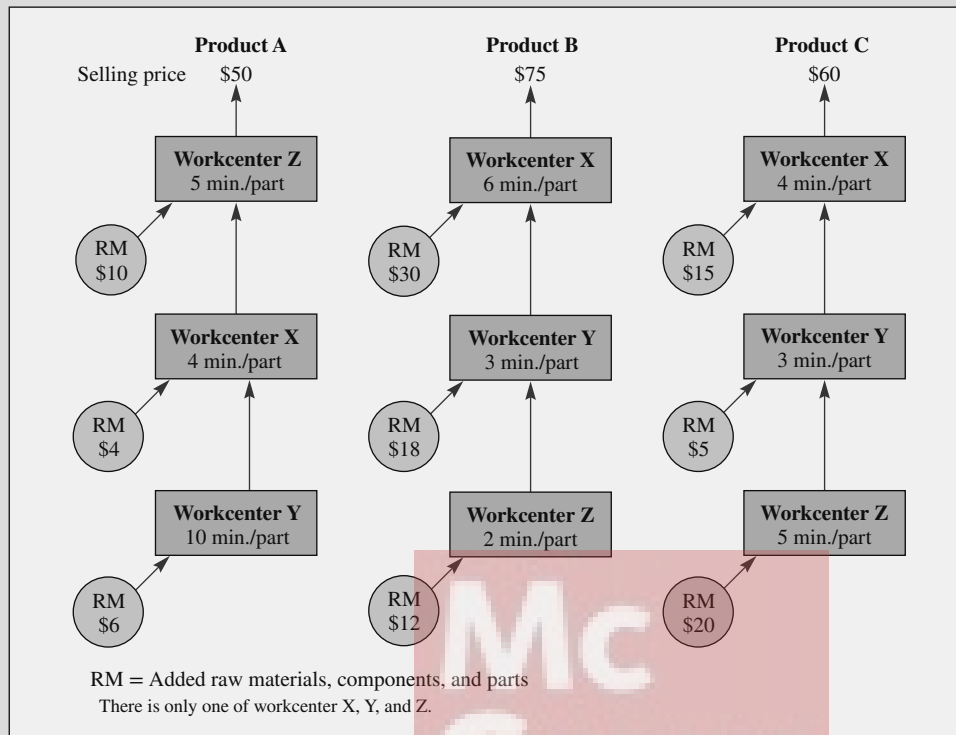
Evaluate bottleneck scheduling problems by applying TOC principles.

We now present three examples to show that different objectives and measurement criteria can lead to the wrong decisions. These examples also show that, even though you may have all the data required, you still may not be able to solve the problem—unless you know how!

EXAMPLE 22S.1: What to Produce?

In this first example, three products (A, B, and C) are sold in the market at \$50, \$75, and \$60 per unit, respectively. The market will take all that can be supplied. The marketing personnel at the company are paid a sales commission which is based only on the sales price of the product.

Three workcenters (X, Y, and Z) process the three products as shown in Exhibit 22S.14. Processing times for each workcenter also are shown. Note that each workcenter works on all three products. Raw materials, parts, and components are added at each workcenter to produce each product. The per unit cost of these materials is shown as RM.

Exhibit 22S.14 Multiple Criteria Economic Analysis —Using Theory Of Constraints

Which product or products should be produced? Consider the following three objectives: (1) maximize the commission made by the marketing personnel, (2) maximize the per unit gross profit, and (3) maximize total gross profit.

SOLUTION

Three different objectives could exist that lead to different conclusions:

1. Maximize sales revenue because marketing personnel are paid commissions based on total revenue.
2. Maximize per unit gross profit.
3. Maximize total gross profit.

In this example, we use gross profit as selling price less materials. We also could include other expenses such as operating expenses, but we left them out for simplicity. (We include operating expenses in our next example.)

Objective 1: Maximize sales commission. Sales personnel in this case are unaware of the processing time required so, therefore, they will try to sell only B at \$75 per unit and none of A or C. Maximum revenue is determined by the limiting resource as follows:

Product	Limiting Resource	Time Required	Number Produced per Hour	Selling Price	Sales Revenue per Hour
A	Y	10 min	6	\$50	\$300
B	X	6 min	10	75	750
C	Z	5 min	12	60	720

Objective 2: Maximize per unit gross profit.

(1) Product	(2) Selling Price	(3) Raw Material Cost	(4) Gross Profit per Unit (2) – (3)
A	\$50	\$20	\$30
B	75	60	15
C	60	40	20

The decision would be to sell only Product A, which has a \$30 per unit gross profit.

Objective 3: Maximize total gross profit. We can solve this problem by finding either total gross profit for the period or the rate at which profit is generated. We use rate to solve the problem both because it is easier and because it is a more appropriate measure. We use profit per hour as the rate.

Note that each product has a different workcenter that limits its output. The rate at which the product is made is then based on this bottleneck workcenter.

(1) Product	(2) Limiting Work Center	(3) Processing Time per Unit (Minutes)	(4) Product Output Rate (per Hour)	(5) Selling Price	(6) Raw Material Cost	(7) Profit per Unit	(8) Profit per Hour (4) × (7)
A	Y	10	6	\$50	\$20	\$30	\$180
B	X	6	10	75	60	15	150
C	Z	5	12	60	40	20	240

From our calculations, and if we consider only a single product, Product C provides the highest profit of \$240 per hour. Note that we get three different answers:

1. We choose B to maximize sales revenue.
2. We choose A to maximize profit per unit.
3. We choose C to maximize total profit.

Choosing Product C is obviously the correct answer for the firm if we restrict ourselves to making a single product. Profit can be improved to \$280/hour by producing a mix of 3 units of product A, 2 units of product B, and 8 units of product C each hour. This solution can be obtained by solving the “product mix” problem described in Supplement 19S (see Example 19S.1).

In this example, all workcenters were required for each product and each product had a different workcenter as a constraint. We did this to simplify the problem and to ensure that only one product would surface as the answer. If there were more workcenters or the same workcenter constraint in different products, the problem could still easily be solved using linear programming (as in Supplement 19S).

EXAMPLE 22S.2: How Much to Produce?

In this example shown in Exhibit 22S.15, two workers are producing four products. The plant works three 8-hour shifts per day, five days a week. The market demand is unlimited and takes all the products that the workers can produce. The only stipulation is that the ratio of products sold cannot exceed 10 to 1 between the maximum sold of any one product and the minimum of another. For example, if the maximum number sold of any one of the products is 100 units, the minimum of any other cannot be fewer than 10 units. Workers 1 and 2 on each shift are not cross-trained and can work only on their own operations. The time and raw material (RM) costs are shown in the exhibit, and a summary of the costs and times involved is on the lower portion of the exhibit. Weekly operating expenses are \$3,000.

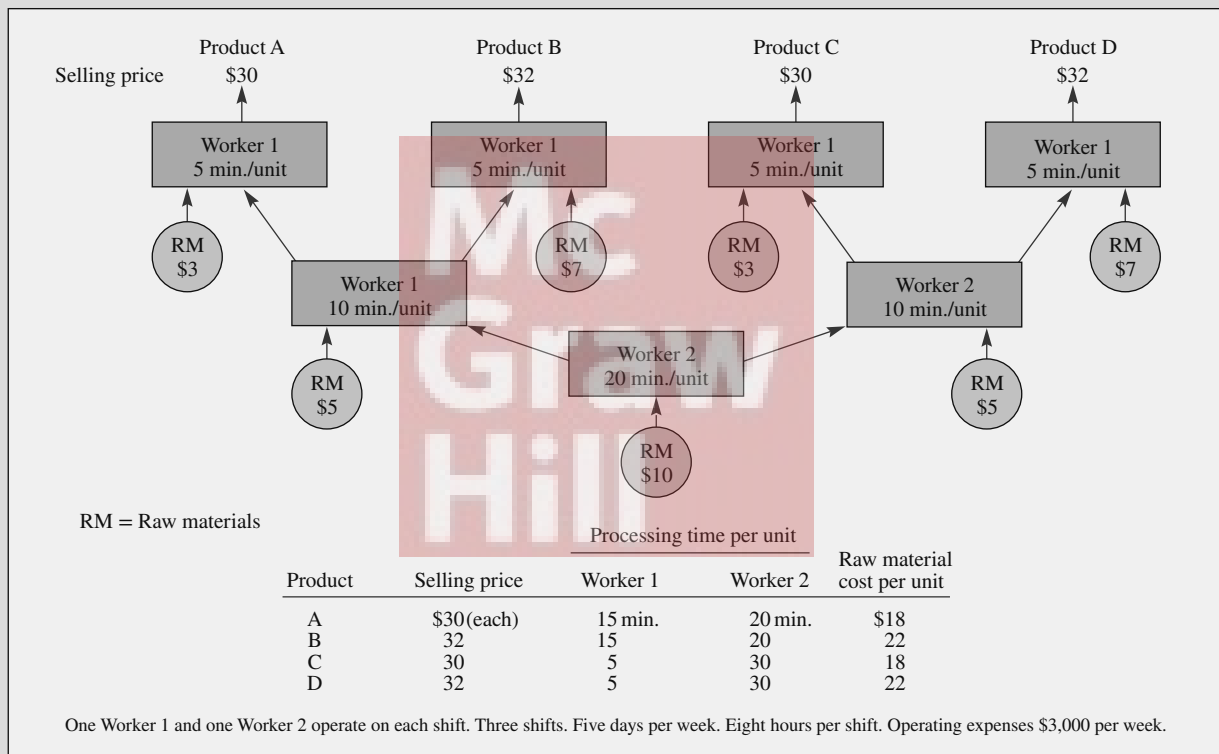
What quantities of A, B, C, and D should be produced given the following three scenarios: (1) maximize the revenue for sales personnel who are paid a commission based on total revenue, (2) maximize per unit gross profit, (3) maximize the total gross profit?

SOLUTION

As in the previous example, there are three answers to this question, depending on each of the following objectives:

1. Maximize revenue for sales personnel who are paid on commission.
2. Maximize per unit gross profit.
3. Maximize the utilization of the bottleneck resource (leading to maximum gross profit).

Exhibit 22S.15 The Production Requirements and Selling Price of Four Products



Objective 1: Maximize sales commission on sales revenue. Sales personnel prefer to sell B and D (selling price \$32) rather than A and C (selling price \$30). Weekly operating expenses are \$3,000.

The ratio of units sold will be 1A : 10B : 1C : 10D.

Worker 2 on each shift is the bottleneck and therefore determines the output. Note that if this truly is a bottleneck with an unlimited market demand, this should be a seven-day-per-week operation, not just a five-day workweek.

5 days per week × 3 shifts × 8 hours × 60 minutes = 7,200 minutes per week available

Worker 2 spends these times on each unit:

A 20 minutes B 20 minutes C 30 minutes D 30 minutes

The ratio of output units is 1 : 10 : 1 : 10. Therefore,

$$\begin{aligned} 1x(20) + 10x(20) + 1x(30) + 10x(30) &= 7,200 \\ 550x &= 7,200 \\ x &= 13.09 \end{aligned}$$

Therefore, the numbers of units produced are

$$A = 13 \quad B = 131 \quad C = 13 \quad D = 131$$

Total revenue is

$$13(30) + 131(32) + 13(30) + 131(32) = \$9,164 \text{ per week}$$

For comparison with Objectives 2 and 3, we will compute gross profit per week.

Gross profit per week (selling price less raw material less weekly expenses) is

$$\begin{aligned} 13(30 - 18) + 131(32 - 22) + 13(30 - 18) + 131(32 - 22) - 3,000 \\ = 156 + 1,310 + 156 + 1,310 - 3,000 \\ = (\$68) \text{ loss} \end{aligned}$$

Objective 2: Maximize per unit gross profit.

	Gross Profit	=	Selling Price	–	Raw Material Cost
A	12	=	30	–	18
B	10	=	32	–	22
C	12	=	30	–	18
D	10	=	32	–	22

A and C have the maximum gross profit, so the ratio will be 10 : 1 : 10 : 1 for A, B, C, and D. Worker 2 is the constraint and has

$$5 \text{ days} \times 3 \text{ shifts} \times 8 \text{ hours} \times 60 \text{ minutes} = 7,200 \text{ minutes available per week}$$

As before, A and B take 20 minutes, while C and D take 30 minutes. Thus,

$$\begin{aligned} 10x(20) + 1x(20) + 10x(30) + 1x(30) &= 7,200 \\ 550x &= 7,200 \\ x &= 13 \end{aligned}$$

Therefore, the number of units produced is

$$A = 131 \quad B = 13 \quad C = 131 \quad D = 13$$

Gross profit (selling price less raw materials less \$3,000 weekly expense) is

$$\begin{aligned} 131(30 - 18) + 13(32 - 22) + 131(30 - 18) + 13(32 - 22) - 3,000 \\ = 1,572 + 130 + 1,572 + 130 - 3,000 \\ = \$404 \text{ profit} \end{aligned}$$

Objective 3: Maximize the use of the bottleneck resource, Worker 2. For every hour Worker 2 works, the following numbers of products and gross profits result:

(1) Product	(2) Production Time	(3) Units Produced per Hour	(4) Selling Price Each	(5) Raw Material Cost per Unit	(6) Gross Profit per Hour (3) × [(4) – (5)]
A	20 minutes	3	\$30	\$18	\$36
B	20	3	32	22	30
C	30	2	30	18	24
D	30	2	32	22	20

Product A generates the greatest gross profit per hour of Worker 2 time, so the ratio is 10 : 1 : 1 : 1 for A, B, C, and D.

Available time for Worker 2 is the same as before:

$$3 \text{ shifts} \times 5 \text{ days} \times 8 \text{ hours} \times 60 \text{ minutes} = 7,200 \text{ minutes}$$

Worker 2 should produce 10 As for every 1 B, 1 C, and 1 D. Worker 2's average production rate is

$$\begin{aligned} 10x(20) + 1x(20) + 1x(30) + 1x(30) &= 7,200 \\ 280x &= 7,200 \\ x &= 25.7 \end{aligned}$$

Therefore, the number of units that should be produced is

$$A = 257 \quad B = 25.7 \quad C = 25.7 \quad D = 25.7$$

Gross profit (price less raw materials less \$3,000 weekly expenses) is

$$\begin{aligned} 257(30 - 18) + 25.7(32 - 22) + 25.7(30 - 18) + 25.7(32 - 22) - 3,000 \\ = 3,084 + 257 + 308.4 + 257 - 3,000 \\ = \$906.40 \end{aligned}$$

In summary, using three different objectives to decide how many of each product to make gave us three different results:

1. Maximizing sales commission resulted in a \$68 loss in gross profit.
2. Maximizing gross profit gave us a profit of \$404.
3. Maximizing the use of the capacity-constrained worker gave us the best gross profit: \$906.40.

Both examples demonstrate that production and marketing need to interact. Marketing should sell the most profitable use of available capacity. However, to plan capacity, production needs to know from marketing what products could be sold.

EXAMPLE 22S.3: TOC Applied to Bank Loan Application Processing

In this example, Goldratt's Theory of Constraints five-step approach (see Exhibit 22S.2) to removing bottlenecks is applied to a bank loan application process. As can be seen through the example, the ideas can be applied to all types of applications, including service processes.

Step 1: Identify the system constraint. Assume that the bank is a private-sector institution and that its goal is to make more money now and in the future. Furthermore, suppose that the initial constraint is internal, namely, the loan officers are unable to carry out all of their responsibilities in a timely manner. That is, given the current demand for bank loan application processing, the loan officers are unable to perform all of the steps in the loan approval process in a responsive manner that is viewed as satisfactory by its customers.

Step 2: Decide how to exploit the system constraint. Once a constraint is identified, management must effectively maximize the usage of the constraint's capacity and capability to fulfill the system's goal. By calculating the throughput yield per unit of time at the constraining resource, management has the information necessary to prioritize the work performed at the constraint. For example, the loan department manager could measure the throughput yield associated with each hour spent working on each type of loan request, such as home mortgage, automobile, and small business. The sequence of loans processed at the constraint would then be established by the "profitability" of the different types of loans so that the bank's goal can be expeditiously met. An optional approach to exploitation that complements

prioritization is ensuring that the constraint is always being effectively utilized. Thus, it may be possible to redesign the loan approval process so that some of the loan officers' current workload is offloaded to available personnel who are currently being only partially utilized.

Step 3: Subordinate everything else to the preceding decision. Subordination involves aligning all of the nonconstraint resources in support of maximizing the performance of the constraint resource. In this case, the bank management would want to schedule appointments for potential customers seeking to complete their loan applications with bank agents so that there was always an abundant supply of completed loan requests waiting for the loan officers to process. Also, the manager of the bank's loan approval process would control the release of loan applications into the approval process so that the loan officers were not overwhelmed. Finally, the bank would have a clerk who is not fully occupied to ensure that each application was complete and met process quality standards prior to being given to the loan officers. (Note that having a supply of finished applications available ensured a highly productive usage of loan officer time; this approach to subordination would produce only a small increase in throughput. It would move the bank toward its goal; the constraint would remain with the loan officers.)

Step 4: Elevate the constraint. Elevating the constraining resource means adding enough new capacity so that the current constraint no longer limits system throughput. In contrast to the previous two steps, elevation often requires a monetary outlay or investment for new resources or capabilities. In the bank's loan subsystem illustration, despite increases in loan officer productivity that were presumed to result from Steps 2 and 3, the system constraint has remained with the bank's loan officers. Thus, because these improvements were insufficient to break the constraint, it is necessary to address the constraining factor directly. The obvious step is to hire an additional loan officer. This action elevates the existing constraint by providing more than sufficient capacity to meet existing demand for processing loan applications. While this decision would produce a sizable increase in operating expenses, it could be justified by management as the best approach to meeting their process goal as well as the bank's overall goal.

Step 5: Return to Step 1, but do not allow inertia to cause a system constraint. After the original constraint has been overcome in Step 4, it is necessary to revisit all of the changes made in Steps 2 and 3 to determine if they are still appropriate to effective process and system performance. In the loan example, a review of the implemented changes in Step 2 might show that offloading the responsibilities for assembling the loan package and some of the credit-checking activities to bank clerical personnel was working well and that there is no need to go back to the original procedure. With regard to Step 3, although the bank might still seek to aggressively schedule bank agents to meet with customers to help them complete their loan applications, it might not be possible to have a large inventory of loan requests in progress because the constraint in the loan application and approval process had shifted to the marketplace. Thus, it is appropriate to return to Step 1 of the five-step focusing process.

Extending the process. Exhibit 22S.16 shows how the application of the five-step focusing process might realistically unfold in managing the bank's loan application process over the next couple of years. Elevating the capacity of the original approval process constraint by hiring a new loan officer leads to a new constraint. This time it resides in the marketplace. Suppose this new constraint turns out to be a policy constraint, namely, bank management does not extend consumer loans to clients who do not use the bank's credit card services. Reconsideration of this policy leads to an exemption for a bank customer who has had any type of an account at the bank for at least the past year. Next, because there are insufficient monetary reserves to fund all of the approved loads, the new system constraint resides in the supply of capital. To address this new constraint, assume that the bank negotiates for additional funds from a wholesale lender and is now able to provide more loans than customers are currently demanding. Now, a new market constraint develops because the monetary supply of funds is greater than the demand in the marketplace. With some effort, the bank marketing team is able to break this constraint by creating a special loan product-service bundle to serve the needs of local university students. Finally in this example, the constraint shifts back inside the bank's loan approval process, where the loan officers and bank clerks are unable to process

applications fast enough to keep up with demand. Bank management purchases a new software package that has been designed to augment loan application processing and fully trains the loan officer staff and clerical assistants on its use.

Exhibit 22S.16 Sequential Application of the Five-Step Focusing Process in Managing the Bank's Loan Subsystem

Constraint Location	Constraint Type	Constraint Identification	Approach to Constraint Alleviation
Bank loan application process	Physical	Loan officers and bank clerks are unable to process all customer loan applications in a timely manner.	Some loan officer tasks offloaded to clerks and additional loan officers hired. Now sufficient capacity exists in the loan application process.
Marketplace	Policy	Current bank policy: If loan applicants do not have a credit card account with the bank, they are not eligible to apply for a consumer loan.	New bank policy: Every loan applicant must have some type of active account with the bank. Now demand for loans increases because more potential applicants qualify.
Supply	Physical	The availability of funds is insufficient to meet all approved customer loan requests.	Bank negotiates for additional funds from wholesale lenders. Now capital reserves are greater than customers are demanding.
Marketplace	Policy	Loan markets are saturated relative to current loan products, and excess funds are available to loan to qualified customers.	Bank develops new loan product designed for local college students. Now total demand for loans in the marketplace increases.
Bank loan application process	Physical	Loan officers and bank clerks are unable to process all customer loan applications in a timely manner.	Bank invests in the acquisition of a new software package to facilitate loan application processing. Now process capacity exceeds demand.

Source: Richard A. Reid, "Applying the TOC Five-Step Focusing Process in the Service Sector: A Banking Subsystem," *Managing Service Quality* 17, no. 2 (2007), pp. 209–234. Copyright © 2007 Emerald Group Publishing Ltd.

Concept Connections

LO22S-1 Explain the Theory of Constraints (TOC).

Summary

- Eli Goldratt developed his Theory of Constraints as an alternative way to think about improving processes. His ideas have stimulated thought by practitioners due to their applicability to many areas, including production, distribution, and project management.
- His underlying philosophy is that it is essential to concentrate on system limitations imposed by capacity-constrained resources, and for a firm to make money it must systematically remove these limitations.
- He argues that, to do this, the firm must simultaneously increase throughput, reduce inventory, and reduce operating expenses. He argues that improving labor productivity will not necessarily make money for the firm and will do so only when it increases throughput, reduces inventory, or reduces operating expenses.
- Goldratt argues that trying to maintain perfectly balanced capacity leads to many problems because this makes every resource dependent on every other. Since statistical fluctuations are inherent in any process, perfect balance leads to disruptions. He argues that not capacity, but flow through the process, should be balanced.

Key Terms

Synchronous manufacturing A production process coordinated to work in harmony to achieve the goals of the firm.

Throughput The rate at which money is generated by the system through sales (Goldratt's definition).

Inventory All the money that the system has invested in purchasing things it intends to sell (Goldratt's definition).

Operating expenses All the money that the system spends to turn inventory into throughput (Goldratt's definition).

Productivity A measure of how well resources are used. According to Goldratt's definition, all the actions that bring a company closer to its goals.

LO22S-2 Analyze bottleneck resources and apply TOC principles to controlling a process.

Summary

- Managing the flow through the bottlenecks is essential to the TOC synchronous manufacturing approach.
- Bottlenecks are identified by calculating the expected utilization (percentage of capacity used) for each resource.
- Saving time on a bottleneck resource is the only way to increase throughput.
- A technique that paces work through the system according to the speed of the bottleneck is used to manage flow through the system.

Key Terms

Bottleneck A resource that limits the capacity or maximum output of the process.

Nonbottleneck Any resource whose capacity is greater than the demand placed on it (Goldratt's definition).

Capacity-constrained resource (CCR) A resource whose utilization is close to capacity and could be a bottleneck if not scheduled carefully (Goldratt's definition).

LO22S-3 Compare TOC to conventional approaches.

Summary

- MRP uses a backward scheduling approach and is oriented toward meeting due dates and maximizing the use of capacity.
- JIT pulls material based on need but does not allow much flexibility for changes, particularly when capacity is tight.
- Synchronous manufacturing (TOC) is more flexible and focused on maximizing flow through the system while minimizing cost.
- In order for TOC to be successfully applied, the firm must recognize how it can be in conflict with conventional accounting and marketing/sales thought. Traditional cost accounting is based on a goal of fully utilizing resources and minimizing cost. This can be in direct conflict with TOC goals, which include maximizing profit by improving throughput.

LO22S-4 Evaluate bottleneck scheduling problems by applying TOC principles.

Summary

- TOC can be used to schedule production. The solutions are often very different compared to those using conventional rules such as discussed in Chapter 22.

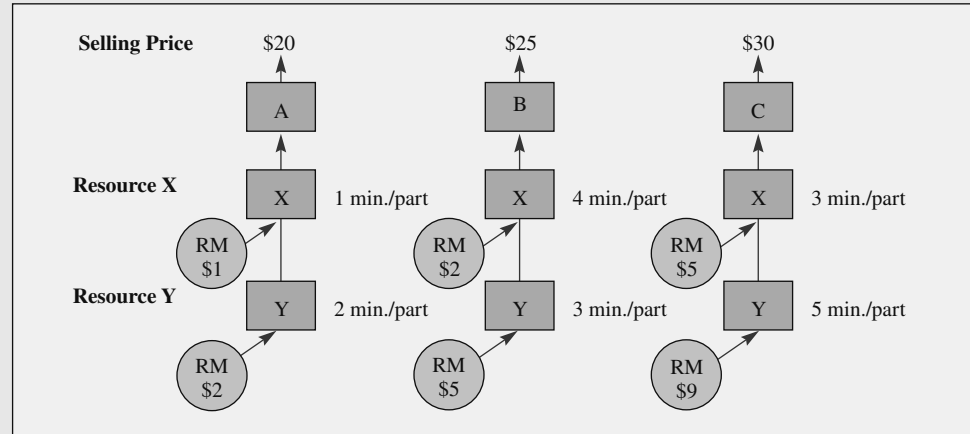
Solved Problem

LO22S-4 Here is the process flow for Products A, B, and C: Products A, B, and C sell for \$20, \$25, and \$30, respectively. Resource X and Resource Y are used to produce A, B, and C. The number of minutes needed for each unit produced is shown in the diagram. Resources X and Y are available 24 hours per day and 7 days a week. Raw materials (RM) are needed at the process steps as shown with the costs in dollars per unit of raw material. (One unit is used for each product.)

The market will take all that you can produce.

a. Which product would you produce to maximize gross margin per unit?

- b. If sales personnel are paid on commission, which product or products would they sell and how many could they sell?
- c. Which and how many product or products should you produce to maximize gross profit for one week?
- d. From part (c), how much gross profit would there be for the week?



Solution

- a. Maximizing gross margin per unit:

	Gross Margin	=	Selling Price	–	Raw Material Cost
A	17	=	20	–	3
B	18	=	25	–	7
C	16	=	30	–	14

Product B will be produced.

- b. Maximizing sales commission: Sales personnel would sell the highest-priced product, C (unless they knew the market and capacity limitations). If we assume the market will take all that we can make, then we would work 7 days/week, 24 hours/day. Resource Y is the constraint in producing C. The number of C we can make in a week is

$$C = \frac{24 \text{ hours/day} \times 7 \text{ days/week} \times 60 \text{ minutes/hour}}{5 \text{ minutes/part}} = 2,016 \text{ units}$$

- c. To maximize profit, we need to compare profits per hour for each product:

(1) Product	(2) Constraint Resource	(3) Production Time on Resource	(4) Number of Units Output per Hour	(5) Selling Price (\$)	(6) RM Cost (\$)	(7) Gross Profit per Hour (4) × (5 – 6)
A	Y	2	30	20	3	\$510
B	X	4	15	25	7	270
C	Y	5	12	30	14	192

If the constraining resource were the same for all three products, our problem would be solved and the answer would be to produce just A, and as many as possible. However, X is the constraint for B, so the answer could be a combination of A and B. To test this, we can see that the value of each hour of Y while producing B is

$$\frac{60 \text{ minutes/hour}}{3 \text{ minutes/unit}} \times (\$25 - 7) = \$360/\text{hour}$$

This is less than the \$510 per hour producing A, so we would produce only A. The number of units of A produced during the week is

$$\frac{60 \text{ minutes/hour} \times 24 \text{ hours/day} \times 7 \text{ days/week}}{2 \text{ minutes/unit}} = 5,040$$

d. Gross profit for the week is $5,040 \times \$17 = \$85,680$.

Discussion Questions

- LO22S-1**
1. How does Goldratt's Theory of Constraints (TOC) differ from other approaches to continuous improvement in organizations? How is it similar?
 2. State the global performance measurements and operational performance measurements and briefly define each. How do these differ from traditional accounting measurements?
 3. Most manufacturing firms try to balance capacity for their production sequences. Some believe this is an invalid strategy. Explain why balancing capacity does not work.
- LO22S-2**
4. Individually or in a small group, examine your own experiences either working in a company or as a customer of a company. Describe an instance where TOC was successfully applied to improve a process, or where you saw the potential for TOC to improve the process.
 5. Discuss why transfer batches and process batches often may not and should not be equal.
 6. Discuss process batches and transfer batches. How might you determine what the sizes should be?
 7. Define and explain the cause or causes of a moving bottleneck.
 8. Explain how a nonbottleneck can become a bottleneck.
 9. Discuss the concept of "drum-buffer-rope."
- LO22S-3**
10. Compare and contrast JIT, MRP, and synchronized manufacturing, stating their main features, such as where each is or might be used, amounts of raw materials and work-in-process inventories, production lead times and cycle times, and methods for control.
 11. Compare the importance and relevance of quality control in JIT, MRP, and synchronous manufacturing.
 12. Discuss how a production system is scheduled using MRP logic, JIT logic, and synchronous manufacturing logic.
 13. Discuss what is meant by forward loading and backward loading.
 14. Define *process batch* and *transfer batch* and their meaning in each of these applications: MRP, JIT, and bottleneck or constrained resource logic.
 15. From the standpoint of the scheduling process, how are resource limitations treated in an MRP application? How are they treated in a synchronous manufacturing application?
 16. What are operations people's primary complaints against the accounting procedures used in most firms? Explain how such procedures can cause poor decisions for the total company.
- LO22S-4**
17. As an individual or small group, visit your favorite fast-food restaurant during lunch period. As you order, and while you are eating, observe as much of the process as you can. Where are the bottlenecks in the system? What recommendations can you come up with to relieve them?
 18. When making decisions about what product(s) to produce in a manufacturing system, why is it not enough to simply consider the selling prices and demands for different items?

Objective Questions

- LO22S-1**
1. What is the name of the software Goldratt developed to implement his idea of TOC?
 2. What are the three financial measurements necessary to adequately measure a firm's performance?
 3. What is the term that refers to the entire production process working in harmony to achieve the profit goals of the firm?
 4. What classic operational measurement does Goldratt redefine as "all the actions that bring a company closer to its goals"?

LO22S-2

5. The following production flow shows Parts O, Q, and T; Subassembly U; and the final assembly for Product V:

M to N to O

P to Q

R to S to T

O and Q to U

U and T to V

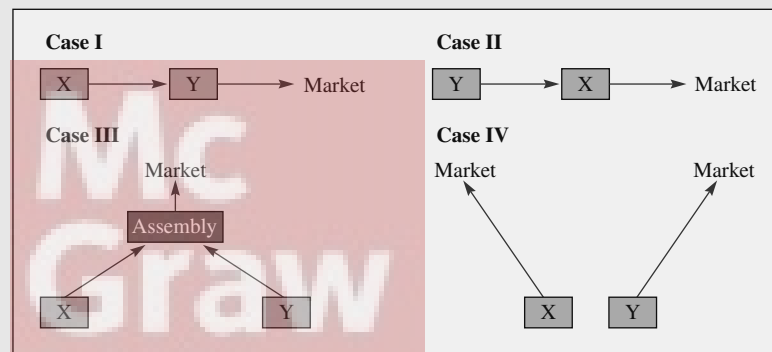
N involves a bottleneck operation, and S involves a capacity-constrained resource.

Draw the process flow.

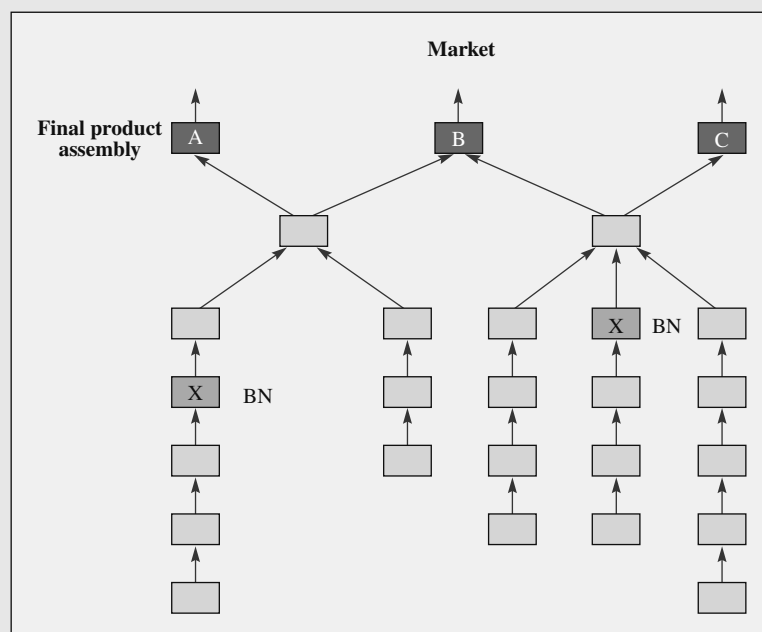
6. For the four basic configurations that follow, assume that the market is demanding product that must be processed by both Resource X and Resource Y for Cases I, II, and III. For Case IV, both resources supply separate but dependent markets; that is, the number of units of output from both X and Y must be equal.

Plans are being made to produce a product that requires 40 minutes on Resource X and 30 minutes on Resource Y. Assume that there is only one of each of these resources and that market demand is 1,400 units per month.

How many hours of production time would you schedule for X and Y? What would happen if both were scheduled for the same number of hours? (Answer in Appendix E.)



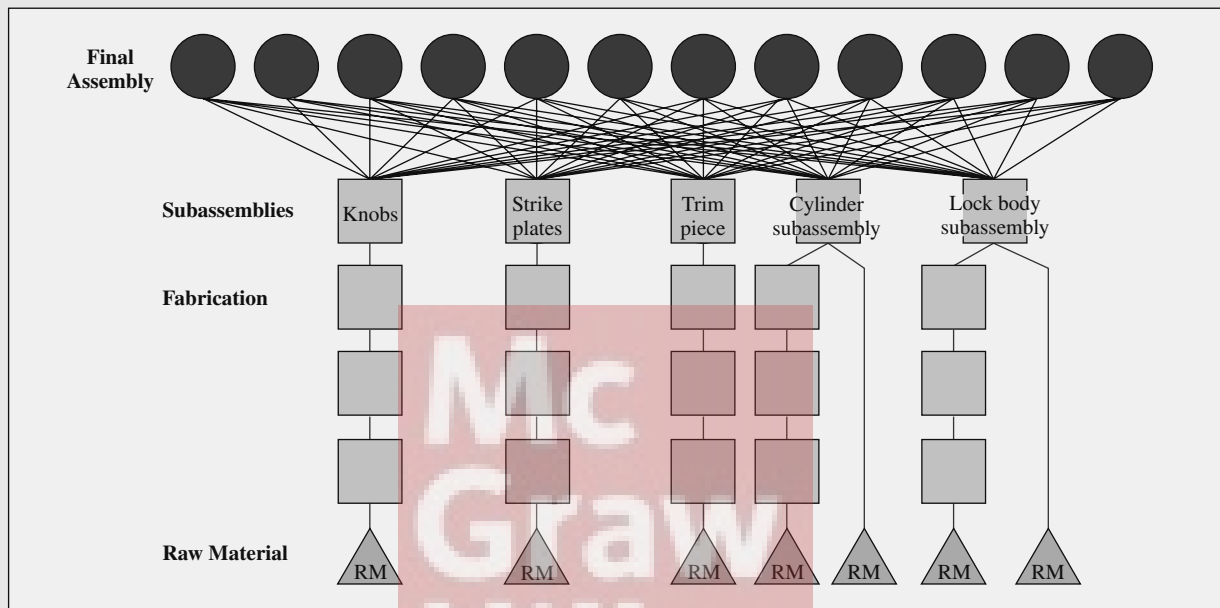
7. Following are the process flow sequences for three products: A, B, and C. There are two bottleneck operations—on the first leg and fourth leg—marked with an X. Boxes represent processes, which may be either machine or manual. Suggest the location of the drum, buffer, and ropes.



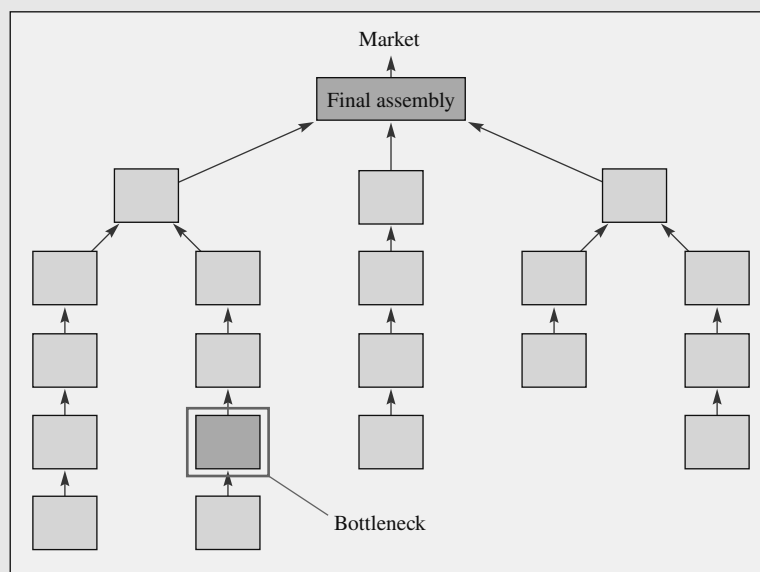
8. Willard Lock Company is losing market share because of horrendous due date performance and long delivery lead times. The company's inventory level is high and includes

many finished goods that do not match the short-term orders. Material control analysis shows that purchasing has ordered on time, the vendors have delivered on time, and the scrap/rework rates have been as expected. However, the buildable mix of components and subassemblies does not generally match the short-term and past due requirements at final assembly. End-of-month expediting and overtime are the rule, even though there is idle time early in the month. Overall efficiency figures are around 70 percent for the month. These figures are regarded as too low.

You have just been hired as a consultant and must come up with recommendations. Help the firm understand its problems. State some specific actions it should take. The product flow is shown in the following diagram.



9. The accompanying figure shows a production network model with the parts and processing sequences. State clearly on the figure (1) where you would place inventory; (2) where you would perform inspection; and (3) where you would emphasize high-quality output.



10. The following production flow shows Parts E, I, and N; Subassembly O; and the final assembly for Product P.
 A to B to C to D to E
 F to G to H to I

J to K to L to M to N

E and I to O

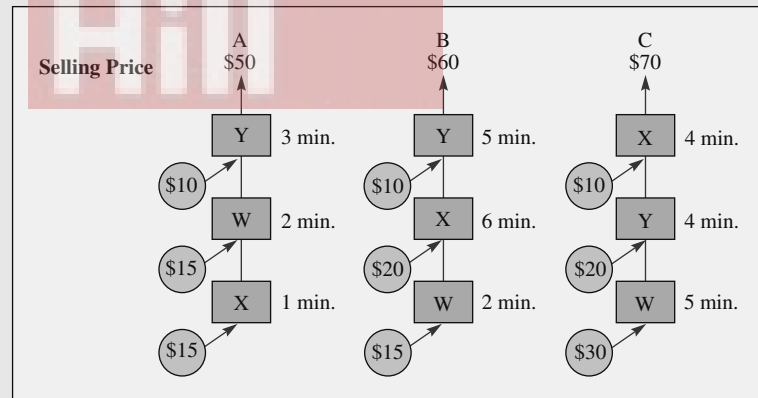
N and O to P

B involves a bottleneck operation, and M involves a CCR.

- Draw the process flow.
 - Where would you locate buffer inventories?
 - Where would you place inspection points?
 - Where would you stress the importance of quality production?
11. Here are average process cycle times for several workcenters. State which are bottlenecks, nonbottlenecks, and capacity-constrained resources.

Processing time		Setup time
Processing time	Setup	Idle
Processing time	Setup	Idle
Processing time	Setup	Idle
Processing time	Setup	Idle
Processing time	Setup	Idle

12. The following diagram shows the flow process, raw material costs, and machine processing time for three products: A, B, and C. There are three machines (W, X, and Y) used in the production of these products; the times shown are in required minutes of production per unit. Raw material costs are shown in cost per unit of product. The market will take all that can be produced.
- Assuming that sales personnel are paid on a commission basis, which product should they sell?
 - On the basis of maximizing gross profit per unit, which product should be sold?
 - To maximize total profit for the firm, which product should be sold?

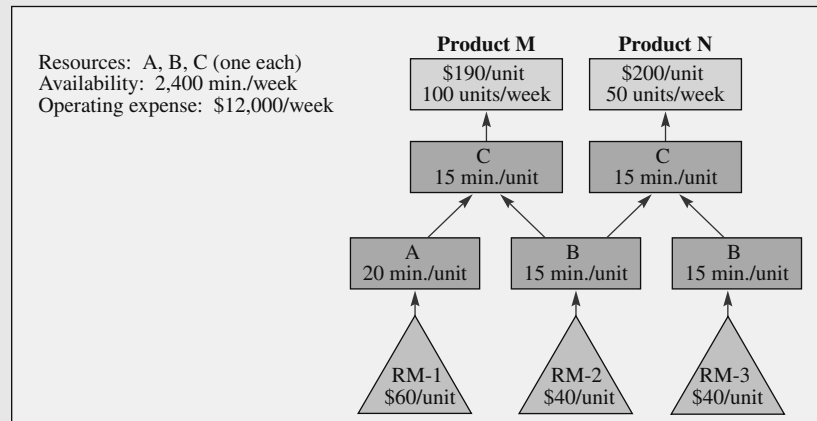


- LO22S-3**
- How does synchronous manufacturing differ from MRP with respect to scheduling?
 - JIT is limited to what type of manufacturing environment?
 - Cost accounting logic can lead managers to keep their resources busy all the time, increasing productivity with no regard to demand. What is the negative result from this effect?
 - The solution to coping with natural differences between marketing and production functions is to do what two things?
- LO22S-4**
- The M–N plant manufactures two different products: M and N. Selling prices and weekly market demands are shown in the following diagram. Each product uses raw materials with costs as shown. The plant has three different machines: A, B, and C. Each performs different tasks and can work on only one unit of material at a time.

Process times for each task are shown in the diagram. Each machine is available 2,400 minutes per week. There are no “Murphys” (major opportunities for the system to foul up). Setup and transfer times are zero. Demand is constant.

Operating expenses (including labor) total a constant \$12,000 per week. Raw materials are not included in weekly operating expenses. (Answers in Appendix E.)

- Where is the constraint in this plant?
- What product mix provides the highest profit?
- What is the maximum weekly profit this plant can earn?



18. A steel product is manufactured by starting with raw material (carbon steel wire) and then processing it sequentially through five operations using machines A to E, respectively (see following table). This is the only use that the five machines are put to. The hourly rates for each machine are given in the table.

Operation:	1	2	3	4	5
Machine:	A	B	C	D	E
Hourly unit output rate:	100	80	40	60	90

Consider the following questions:

- What is the maximum output per hour of the steel product?
- By how much would the output be improved if the output of B was increased to 90 per hour?
- By how much would the output be improved if the output for C was increased to 50 per hour?
- By how much would the output be improved if the output for C was increased to 70 per hour?
- What is the effect on the system if machine A can manage an output of only 90 in one hour?
- What is the effect on the system if machine C can manage an output of only 30 in one hour?
- What is the effect on the system if machine B is allowed to drop to an output of 30 in one hour?

Practice Exam

In each of the following, name the term defined or answer the question. Answers are listed at the bottom.

- According to Goldratt, the goal of a firm is to do what?
- At the operational level, Goldratt suggests that these three measures should guide decisions.
- The goal related to these three measures is this.
- Goldratt argues that, rather than capacity, this should be balanced.
- This is any resource whose capacity is less than the demand placed on it.
- Goldratt suggests that a production system should be controlled using these three mechanisms.
- A bottleneck should have this placed in front of it to ensure it never runs out of work.
- A rope is used for this purpose in controlling a production system.
- To pace the production system, this is used.
- This is a measure of the value of inventory and the time it stays within an area.

Answers to Practice Exam 1. Make money 2. Throughput, inventory, operating expenses 3. Increase throughput while simultaneously reducing inventory and reducing operating expense 4. Flow 5. Bottleneck 6. Drum, buffer, and rope 7. Buffer 8. Communication 9. Drum 10. Dollar days