Chapter 5
Managing Process Constraints

- Theory of Constraints
- Managing Bottlenecks
- Assembly Line Balancing

What is a Constraint?

Constraint: Any factor that limits the performance of a system and restricts its output.

Bottleneck: A capacity constraint resource (CCR) whose available capacity limits the organization's ability to meet the product volume, product mix, or demand fluctuations required by the marketplace

Operational Measures vs. Financial Measures

Operational Measures	TOC View	Relationship to Financial Measures
Inventory (I)	All the money invested in a system in purchasing things that it intends to sell	A decrease in I leads to an increase in net profit, ROI, and cash flow.
Throughput (T)	Rate at which a system generates money through sales	An increase in T leads to an increase in net profit, ROI, and cash flows.
Operating Expense (OE)	All the money a system spends to turn inventory into throughput	A decrease in OE leads to an increase in net profit, ROI, and cash flows.
Utilization (U)	The degree to which equipment, space, or workforce is currently being used; it is measured as the ratio of average output rate to maximum capacity, expressed as a percentage	An increase in U at the bottleneck leads to an increase in net profit, ROI, and cash flows. U↗ ⇒ T↗ ⇒ profit & ROI↗

Theory of Constraints

The focus should be on balancing flow, not on balancing capacity.

Maximizing the output and efficiency of every resource may not maximize the throughput of the entire system.

An hour lost at a bottleneck... is an hour lost for the whole system. An hour saved at a non-bottleneck resource is a mirage.

Inventory is needed only in front of bottlenecks and in front of assembly and shipping points.

Work should be released into the system only as frequently as needed by the bottlenecks. Bottleneck flows should be equal to market demand

Activating a non-bottleneck resource… doesn't increase throughput or promote better performance.

Every capital investment must be viewed from the perspective of the global impact on overall throughput, inventory, and operating expense.

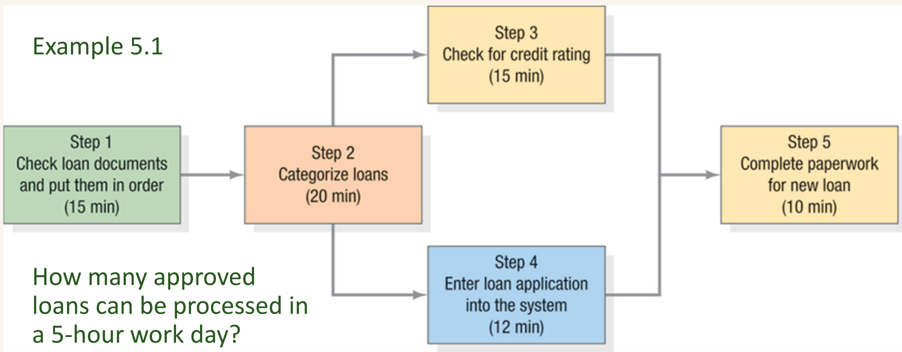
Implementation of The Theory of Constraints

1. Identify the System Bottleneck(s)
2. Exploit the Bottleneck(s): Maximize the throughput of the bottleneck(s).
3. Subordinate All Other Decisions to Step 2: Non-bottleneck resources should be scheduled to support the bottleneck.
4. Elevate the Bottleneck(s): Try to increase the capacity of the bottleneck
5. Do Not Let Inertia Set In: Repeat steps 1–4 in order to identify and manage the new set of constraints.

Managing Bottlenecks in Service Processes

- Throughput time: Total elapsed time from the start to the finish of a job or a customer being processed at one or more work centers

Example 5.1



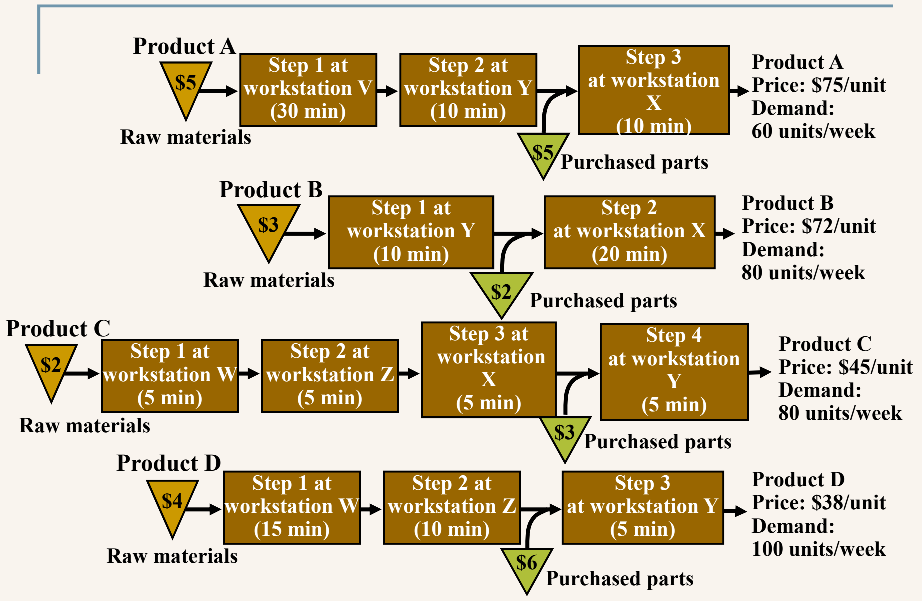
How many approved loans can be processed in a 5-hour work day?

Example 5.1

- The bottleneck is Step 2.
- The throughput time to complete an approved loan application is 15 + 20 + max(15, 12) + 10 = 60 minutes.
- The actual time taken for completing an approved loan will be longer due to non-uniform arrival of applications, variations in actual processing times, and the related factors.
- The capacity for loan completions is 3 customers per hour because the bottleneck step 2 can process only 1 customer every 20 minutes (60/3).
- The bank will be able to complete a maximum of 15 new loan accounts, in a 5-hour day.
- Management can increase the capacity of Step 2 up to the point where another step becomes the bottleneck.

Managing Bottlenecks in Manufacturing

- Manufacturing processes often pose complexities when identifying bottlenecks. If multiple products are involved, extra setup time at a workstation is usually needed to change over from one product to the next, which in turn increases the overload at the workstation.
- Example 5.2: Identifying Bottlenecks in a Batch Process
  - Diablo Electronics manufactures four unique products (A, B, C, D) that are fabricated and assembled in five different workstations (V, W, X, Y, Z) using a small batch process. Batch setup times have been reduced to such an extent that they are negligible.
  - Diablo can make and sell up to the limit of its demand per week, and no penalties are incurred for not meeting all the demand.



Example 5.2

Each week consists of 2,400 minutes of available production time.

Workstation	Load from Product A	Load from Product B	Load from Product C	Load from Product D	Total Load (min)
V	60×30 = 1800	0	0	0	1,800
W	0	0	80×5 = 400	100×15 = 1,500	1,900
X	60×10 = 600	80×20 =1,600	80×5 = 400	0	2,600
Y	60×10 = 600	80×10 =800	80×5 = 400	100×5 = 500	2,300
Z	0	0	80×5 = 400	100×10 = 1,000	1,400

Identifying bottlenecks becomes harder when setup times are lengthy and the degree of divergence in the process is greater. … floating bottlenecks

Drum-Buffer-Rope Systems

Drum-Buffer-Rope: A planning and control system that regulates the flow of work-in-process materials at the bottleneck or the capacity constrained resource (CCR) in a productive system

- The bottleneck schedule is the drum because it sets the beat or the production rate for the entire plant and is linked to market demand.
- The buffer is the time buffer that plans early flows into the bottleneck and thus protects it from disruption.
- The rope represents the tying of material release to the drum beat, which is the rate at which the bottleneck controls the throughput of the entire plant.

Applying TOC to Product Mix Decisions

Example 5.3: The management at Diablo Electronics wants to improve profitability by accepting the right set of orders (product mix).

They collected the following financial data:

- Variable overhead costs are $8,500 per week.
- Each worker is paid $18 per hour and is paid for an entire week, regardless of how much the worker is used.
- Labor costs are fixed expenses.
- The plant operates one 8-hour shift per day, or 40 hours each week.

Example 5.3

Traditional Method: accept as much of the highest contribution margin product as possible (up to the limit), followed by the next highest contribution margin product, and so on until no more capacity is available.

Step 1: Calculate the contribution margin per unit of each product.

	A	B	C	D
Price	$75.00	$72.00	$45.00	$38.00
Raw material and purchased parts	–10.00	–5.00	–5.00	–10.00
= Contribution margin	$65.00	$67.00	$40.00	$28.00

The order of the contribution margin per unit is B, A, C, D.

Example 5.3

Step 2: Allocate resources V, W, X, Y, and Z to the products in the order decided in Step 1. Satisfy each demand until the bottleneck resource (workstation X) is encountered.

B(80):Y10, X20 A(60): V30, Y10, X10 C(80):W5, Z5, X5, Y5 D(100):W15, Z10, Y5

Work Center	Minutes at the Start	Minutes Left After Making 80 B	Minutes Left After Making 60 A	Can Only Make 40 C	Can Only Make 100 D
V	2,400	2,400	600	600	600
W	2,400	2,400	2,400	2,200	700
X	2,400	800	200	0	0
Y	2,400	1,600	1,000	800	300
Z	2,400	2,400	2,400	2,200	1,200

Step 3: Profit=(80×67+60×65+40×40+100×28) – (3600+8500)=1560

Example 5.3

Bottleneck Method (TOC): Base on the dollar contribution margin per minute of processing time at the bottleneck station.

Step 1: Calculate the contribution margin/minute of processing time at bottleneck workstation X:

	Product A	Product B	Product C	Product D
Contribution margin	**$65.00**	**$67.00**	**$40.00**	**$28.00**
Time at bottleneck	**10 min.**	**20 min.**	**5 min.**	**0 min.**
Contribution margin per minute	**$6.50**	**$3.35**	**$8.00**	**Not defined**

Product D is scheduled first because it does not consume any resources at the bottleneck. The manufacturing sequence is D, C, A, B.

Example 5.3

Step 2: Allocate resources V, W, X, Y, and Z to the products in the order decided in step 1. Satisfy each demand until the bottleneck resource (workstation X) is encountered.

D(100):W15, Z10, Y5 C(80):W5, Z5, X5, Y5 A(60): V30, Y10, X10 B(80):Y10, X20

Work Center	Minutes at the Start	Minutes Left After Making 100 D	Minutes Left After Making 80 C	Minutes Left After Making 60 A	Can Only Make 70 B
V	2,400	2,400	2,400	600	600
W	2,400	900	500	500	500
X	2,400	2,400	2,000	1,400	0
Y	2,400	1,900	1,500	900	200
Z	2,400	1,400	1,000	1,000	1,000

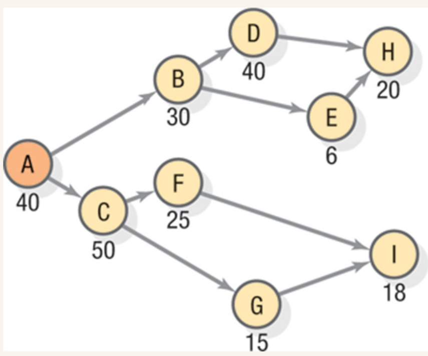
Step 3: Profit=(100×28+80×40+60×65+70×67) – (3600+8500)=2490

Managing Constraints in a Line Process

Example 5.4: Green Grass, Inc. is designing an assembly line to produce a new fertilizer spreader.

Work Element	Description	Time (sec)	Immediate Predecessor(s)
A	Bolt leg frame to hopper	40	None
B	Insert impeller shaft	30	A
C	Attach axle	50	A
D	Attach agitator	40	B
E	Attach drive wheel	6	B
F	Attach free wheel	25	C
G	Mount lower post	15	C
H	Attach controls	20	D, E
I	Mount nameplate	18	F, G
		Total 244	

Precedence Diagram for Example 5.4



One worker for each step

9 workers are needed.
Output rate = one unit every 50 sec. = 72 units/hour

One worker for all steps

Output rate = one unit every 244 sec. = 14.75 units/hour

Line Balancing 1/3

- The assignment of work to stations in a line so as to achieve the desired output rate with the smallest number of workstations.
- Assume r is the desired output rate matched to the production plan
- Cycle time: Maximum time allowed for process a unit at each station to achieve the desired output rate r. $c = \frac{1}{r}$
- Theoretical Minimum (TM): the smallest number of stations possible to achieve the desired output rate r.

$$TM = \frac{\Sigma t}{c}$$ where Σt = total time required to assemble each unit

Line Balancing 2/3

Create one station at a time. For the station now being created, identify the unassigned work elements that qualify for assignment: They are candidates if

1. All of their predecessors have been assigned to this station or stations already created.

2. Adding them to the workstation being created will not create a workload that exceeds the cycle time.

Decision Rule	Logic
Longest work element	Picking the candidate with the longest time to complete is an effort to fit in the most difficult elements first, leaving the ones with short times to "fill out" the station.
Shortest work element	This rule is the opposite of the longest work element rule because it gives preference in workstation assignments to those work elements that are quicker. It can be tried because no single rule guarantees the best solution. It might provide another solution for the planner to consider.
Most followers	When picking the next work element to assign to a station being created, choose the element that has the most *followers* (due to precedence requirements). In Figure 5.5, item C has three followers (F, G, and I) whereas item D has only one follower (H). This rule seeks to maintain flexibility so that good choices remain for creating the last few workstations at the end of the line.
Fewest followers	Picking the candidate with the fewest followers is the opposite of the most followers rule.

Line Balancing 3/3

- Idle time $= nc - \Sigma t$

  where n =number of stations, c = cycle time,
  Σt =total time required to assemble each unit

- Efficiency: the ratio of productive time to total time

  Efficiency (%) = $\frac{\Sigma t}{nc} \times (100)$

- Balance Delay: the amount by which efficiency falls short of 100%

  Balance delay (%) = 100 – Efficiency

Example 5.5

Green Grass's plant manager just received marketing's latest forecasts of Big Broadcaster sales for the next year. She wants its production line to be designed to make 2,400 spreaders per week. The plant will operate 40 hours per week.

a. What should be the line's cycle time?
b. What is the smallest number of workstations that she could hope for in designing the line for this cycle time?
c. Suppose that she finds a solution that requires only five stations. What would be the line's efficiency?

Example 5.5

a. First convert the desired output rate (2,400 units per week) to an hourly rate of 60. Then the cycle time is

$c = 1/r =$ 1/60 (hr/unit) = 1 minute/unit = **60 seconds/unit**

b. Calculate the theoretical minimum for the number of stations by dividing the total time, Σt, by the cycle time, c = 60 sec.

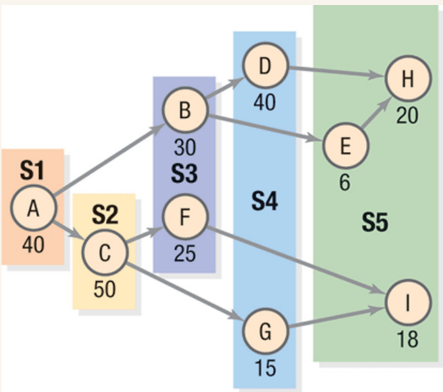
$$\text{TM} = \frac{\Sigma t}{c} = \frac{244 \text{ seconds}}{60 \text{ seconds}} = 4.067 \text{ or } \mathbf{5 \text{ stations}}$$

The number of stations is at least 5.

Example 5.5

The precedence and cycle-time requirements can not be violated.

1. Start with A. → Station 1
2. Use longest work element rule to select C. → Station 2
3. Use most followers rule to select B. → Station 3
   3.1 Add F to Station 3.
4. Use longest work element rule to select D. → Station 4
   4.1 Add G to Station 4.
5. Use most followers rule to select E. → Station 5
   5.1 Add H and I to Station 5.

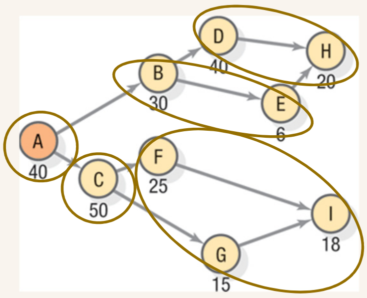


Example 5.5

c. The theoretical minimum number of workstations is 5. So this represents an optimal solution to the problem.

$$\text{Efficiency} = \frac{\Sigma t}{nc} \times (100) = \frac{244}{5(60)} \times (100) = \mathbf{81.3\%}$$

Balance delay (%)
= 100 – Efficiency = 18.7%



Rebalancing the Assembly Line

- Managerial Considerations
  - Pacing is the movement of product from one station to the next as soon as the cycle time has elapsed
  - Behavioral factors such as absenteeism, turnover, and grievances can increase.
  - The number of models produced complicates scheduling and necessitates good communication.
  - Cycle times are dependent on the desired output rate or sometimes on the maximum workstations allowed.